Fall 2022

Bridging Psychometric and Cognitive Models of (General) Intelligence: An Investigation Based on Process Overlap Theory

Han Hao
Claremont Graduate University

Bridging Psychometric and Cognitive Models of (General) Intelligence:

An Investigation Based on Process Overlap Theory

By

Han Hao

Claremont Graduate University

2022

Approval of the Dissertation Committee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Han Hao as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Psychology with a concentration in Cognitive Psychology.

Andrew Conway, Chair
New Mexico State University
Professor of Psychology

Kathy Pezdek
Claremont Graduate University
Professor of Psychology

Gabriel Cook
Claremont McKenna College
Associate Professor of Psychology

Michael J. Kane
University of North Carolina at Greensboro
Professor of Psychology

Abstract

Bridging Psychometric and Cognitive Models of (General) Intelligence:

An Investigation Based on Process Overlap Theory

by

Han Hao

Claremont Graduate University: 2022

Human intelligence has been scientifically investigated as a psychological construct for over a century but there has not been a universally accepted definition or theory. One cause of this problem is that traditional theories attempt to explain the robust findings in cognitive ability testing, such as the positive manifold, from two different perspectives: psychometric or cognitive. Both approaches have their own limitations and are sometimes incompatible with each other. Therefore, contemporary theories of intelligence have been developed to provide a more unified perspective by combining both types of approaches, allowing the psychometric structure of cognitive abilities to be represented and explained by cognitive mechanisms. In other words, inter-individual differences in intelligence are explained in terms of intra-individual psychological processes.

This dissertation investigated a contemporary theoretical framework of intelligence, the process overlap theory (POT; Kovacs & Conway, 2016; 2019), that attempts to bridge the gap between psychometric and cognitive theories of human intelligence. POT proposed a novel psychometric structure and cognitive architecture to explain individual differences in higher-order cognition. This dissertation consisted of three studies that illustrated and investigated the POT framework using a combination of latent factor models, item response models, and psychometric network models of both simulated data and real-world cognitive testing data.

These exploratory studies provide an account of the psychometric structure and the cognitive mechanisms of higher-order cognition from a POT perspective. Results of the studies demonstrated that, based on the POT algorithm, the positive manifold of intelligence can emerge at the psychometric level in the absence of a general mental ability at the cognitive level. This dissertation therefore provides critical supportive evidence for POT and illustrates an alternative theoretical and statistical framework for contemporary research of human cognition that combines psychometric and cognitive theories of intelligence.

Table of Contents

1 Introduction

Human intelligence as a psychological construct has been a topic of scientific investigation for more than a hundred years. Ever since, in both psychology and education, the concept of human intelligence has been controversial, which has resulted in various operational definitions across different fields of research (Sternberg, 2013). To date, there are numerous definitions of intelligence, each of which attempts to explain the pattern and variation in human intellectual performance "on different occasions, in different domains, as judged by different criteria" (Neisser et al., 1996, p.77); yet, there has not been a universally accepted definition, theory, or model. In fact, most definitions of human intelligence largely rely on observations of human performance that are dependent on the specific intellectual activities involved.

For instance, Boring (1923) provided his circular definition of intelligence as being what the intelligence tests test. This definitional predicament has been recognized from the very beginning of intelligence research. However, despite lacking agreement on a definition, the community of intelligence research has widely reached a consensus on several robust findings in cognitive ability testing. For example, one of the most robust patterns found in intelligence testing is the positive manifold, which refers to the finding that performance on subtests of different mental abilities tend to be positively correlated (Carroll, 1993; Neisser et al., 1996; Spearman, 1927). This pattern of communality among diverse intelligence tests prompted the discussion of a "general intelligence", or *g*, that could statistically represent this positive manifold and plays a central role as a psychological property in all intellectual activities. Since the pioneering work of Spearman (1904), different explicit theories have emerged around the construct of general intelligence to measure, explain, and predict individual differences in human intellectual behaviors at different levels of analysis.

Traditionally, some attempts in intelligence research, known as psychometric approaches or psychometric theories of intelligence, have been made to investigate the correlational relationship among behavioral performance in different cognitive activities. Specifically, these attempts focused on understanding how these relationships reflect the role of the general factor(s) in potential structural mechanisms of human cognitive activities (McGrew, 2009; Neisser et al., 1996). Other attempts, known as cognitive approaches or cognitive theories of intelligence, have been focused on identifying specific cognitive processes and associating such processes with general intelligence, or with different ranges of cognitive activities (Sternberg, 1985). These two types of approaches are not necessarily exclusive but have conventionally employed distinct operationalizations, hypotheses, and methods to ask different kinds of questions about human intelligence.

Contemporary intelligence theories have been developed to provide a more unified perspective by combining both types of approaches, allowing the correlational associations among measures of cognitive abilities to be explained by information-processing mechanisms (Conway & Kovacs, 2015; Kaufman et al., 2013; Savi et al., 2019). These more recent theories connect psychometric and cognitive perspectives by proposing a quantitatively defined theoretical framework with a variety of statistical approaches, including structural equation modeling (MacCallum & Austin, 2000), item response theory (Embretson & McCollam, 2000), and network analysis (van der Maas et al., 2017). For instance, a recent framework, namely the process overlap theory (POT; Kovacs & Conway, 2019), proposed a modern sampling theory of cognitive processes that provides an alternative explanation to a range of primary findings in psychometric and cognitive intelligence research. The theory is motivated by both psychometric theories and cognitive theories, providing a bridge between these two traditionally divided approaches. POT, along with other interdisciplinary theories of intelligence, rejects the notion of a psychological *g* in intelligence, suggesting alternative explanations to general intelligence and its underlying cognitive mechanisms, which could better harmonize intelligence research evidence found in psychometrics, cognitive psychology, and neuroscience.

This dissertation investigates contemporary theoretical frameworks that attempt to bridge the gap between psychometric and cognitive models of human intelligence, using a combination of computational approaches such as data simulations, latent factor modeling, item response theory modeling, and psychometric network analysis. Specifically, the current studies focus on a series of investigations on process overlap theory (POT; Kovacs & Conway, 2019), which proposes a novel psychometric structure and cognitive architecture to explain individual differences in intelligence. The POT framework provides an alternative explanation of intelligence and connects psychometric and cognitive approaches using formal computational models of cognitive processes and mechanisms.

This dissertation consists of three studies. The first study illustrates a simulation based on the sampling algorithm of POT to demonstrate that, based on the POT algorithm, the positive manifold can emerge at the psychometric level in the absence of a general cognitive ability at the cognitive level. The second study focuses on extending this simulation by the application of psychometric network analysis. Furthermore, a latent network modeling approach is also applied to the simulated data. The latent network model combines the traditional latent factor model and psychometric network, which is more compatible with process overlap theory than either latent factor models or psychometric networks on its own. In the third study, the latent psychometric network approach is applied to an existing dataset of cognitive ability tests to investigate the correlational associations among working memory tasks and reasoning tasks in the absence of the assumption of a domain-general cause. This exploratory study attempts to explain the psychometric properties as well as the potential cognitive mechanisms in working memory and reasoning tasks from a POT perspective.

2 Conventional Theories of (General) Intelligence

This introduction consists of a brief review of a few classic theories of human intelligence. Depending on the differences in their emphases, these theories are summarized in two major groups: psychometric theories and cognitive theories. These two types of theories attempted to explain the observed convergence and divergence of human intellectual performance but have taken different approaches. Theoretical and methodological problems of the conventional psychometric and cognitive theories, which are provoked by the differences between the two approaches, are also discussed.

2.1 Psychometric Theories

Psychometric theories of intelligence investigate correlational relationships and individual differences in performance of (cognitive) tests that require cognitive activities to understand the “map of mind” (Sternberg, 2012, p.19). Factor analysis was developed to find the underlying dimension(s) of cognitive ability among correlations of cognitive tests in early psychometric work, and has been one of the most influential statistic approaches in the field of psychometrics. With factor analysis, underlying dimensions, known as the common factors, in correlations are extracted to estimate the latent structure of the cognitive tests selected to measure individual differences in complex human cognition. The psychometric theories are mostly built upon the underlying dimension(s) resulting from such analysis. Latent common factors in psychometric theories are conceptualized as, or associated with, specific psychological constructs involved in intelligence. In this section, three broad types of psychometric theories are introduced: general-factor theory, multiple-factor theory, and hierarchical-factor theory. The three types of theories are roughly presented in chronological order of their initial formulation because each of these types complements the preceding theory, providing alternative descriptions of the factorial structure of intelligence.

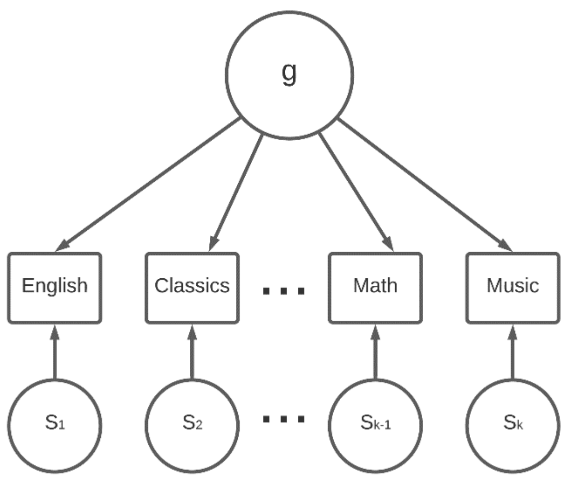
2.1.1 Spearman's Theory of General Intelligence

Based on his advocation of a "correlational psychology" (Spearman, 1904; p. 205) to investigate the psychical tendencies among mental tests, Spearman conducted his famous study on children's school performance and sensory discrimination abilities. This was one of the first attempts to develop a theory of intelligence using a psychometric approach. A primary goal of the study was to propose a method to represent the intercorrelation, or "functional uniformity" (p.219), emerging from diverse mental measures such as school efficiency in different subjects. In this study, Spearman assumed a latent common cause among a set of mental measures, which he referred to as "general intelligence", and developed (an early form of) factor analysis to estimate the association between the latent general intelligence factor and the observed measures based on the raw correlations among the measures. Later, Spearman (1927) formally introduced his factor theory of intelligence, in which intellectual performance on any task was assumed to be determined by two types of factors: a general intelligence (g), that is universal to all mental tests, and specific abilities (s), that are particular to specific tests.

Since Spearman's work, the positive intercorrelations among diverse mental tests have been observed across different batteries of tests and samples (Jensen, 1998; van der Maas et al., 2006). This positive manifold exhibits a robust pattern of positive correlations among cognitive tests, which Spearman interpreted as a reflection of a general mechanism (e.g., "mental energy"; Spearman, 1927) that causes performance on different ability tests to correlate. Spearman's theory proposed a single-factor model of intelligence in which *g*, as the "general intelligence", is regarded as the common cause of all correlations among mental tests, while *s*, as the specific abilities, account for the unique test content in addition to general intelligence. In his theory (see Figure 1 for a conceptual illustration), general intelligence is represented by the latent common factor, and specific abilities are represented by the residual variances after the latent factor is accounted for. The *g* factor is involved in all types of intellectual tasks, although the degree to which different tasks invoke the *g* factor differs, depending on the nature of the tasks. Therefore, as a reflection, tasks with high *g*-to-*s* ratios will tend to show a higher correlation among each other than tasks with low *g*-to-*s* ratios.

Figure 1

The Conceptual Illustration of Spearman's Theory of General Intelligence



Note. Spearman's theory of general intelligence indicated a one-factor model of intelligence in which intellectual tasks/activities (English, Classics, Math, Music, etc.) were influenced by a common factor of intelligence ("*g*" as the latent factor) and task-specific factors ("*s*" as the residuals).

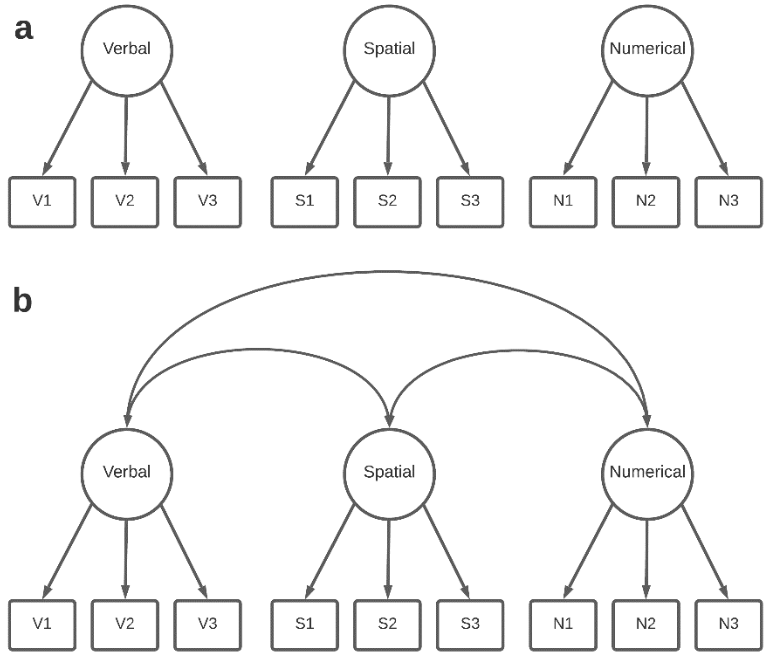
2.1.2 Thurstone's Primary Mental Abilities

Spearman's theory of general intelligence was later challenged by alternative explanations of the positive manifold that do not rely on a unitary factor. For example, Thurstone (1935) argued that the general factor Spearman extracted was not necessarily sufficient to account for the intercorrelations of a set of tests. Based on his extension of Spearman's original factor analytic method, namely, multiple factor analysis and factor rotation technique (Thurstone, 1935, 1947), and his own sampling of tests, Thurstone (1938) proposed a multiple-factor model of intelligence. Instead of being regarded as a unitary trait, Thurstone suggested that human intelligence is represented by a set of uncorrelated primary mental abilities, including verbal comprehension, verbal ability, numerical and arithmetic abilities, spatial visualization, inductive and deductive reasoning, memory, and perceptual speed (see Brody & Brody, 1976; Gardner, 2011). Each of these primary mental abilities constitutes certain aspects of mental activities involved in different tasks.

On a continuum, Thurstone's theory of primary mental abilities may be perceived as being at the opposite extreme of Spearman's theory (Walrath et al., 2020), yet this multiple-factor model of intelligence was actually an extension of Spearman's theory of intelligence both statistically and theoretically. Statistically, the multiple factor analysis, as well as other similar statistical approaches (e.g., principal components analysis, Hotelling, 1933), projected the intercorrelations of test performances to multiple dimensions instead of a single dimension, which accounted for the entire matrix of the observed correlation patterns more effectively. Theoretically, Thurstone's theory placed more attention on the specific roles of intelligence in different areas and argued that the specific factors determined how intelligence is expressed in different situations (Walrath et al., 2020).

Figure 2

A Conceptual Illustration of Thurstone's Primary Mental Abilities



Note. Panel a illustrated an uncorrelated multi-factor model, as Thurstone's theory of primary mental abilities initially presented, in which the latent factors (circles that represent the primary mental abilities) were not correlated to each other. Panel b illustrated a correlated multi-factor model, as Thurstone's later work indicated, in which the latent factors were rotated (correlated). For a proof of concept, only three primary mental abilities, with each accounting for the common variance of three tasks, were included in these illustrations.

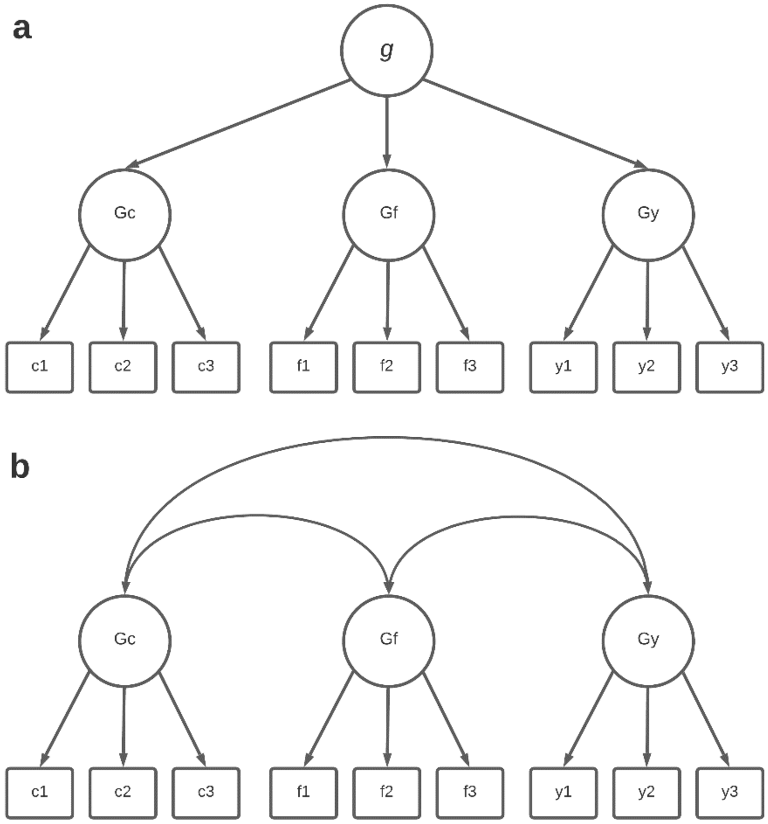
Initially, Thurstone developed the factor analytic method and intelligence theory under the assumption that primary mental abilities represented by multiple factors are orthogonal (Thurstone, 1938). However, his later works with different samples (e.g., Thurstone & Thurstone, 1941) indicated that the simple structure of the rotated factors was best represented when these primary mental ability factors were correlated (see Figures 2a and b). Therefore, it appeared that the positive manifold was still reflected among the correlated factors, providing common ground between Spearman's general factor theory and Thurstone's multiple factor theory.

2.1.3 The CHC Theory

Various attempts have been made to combine Spearman and Thurstone's perspectives, with several of these attempts converging on the idea to conceptualize the correlations among sets of tests by a hierarchical structure of intelligence (e.g., Cattell, 1963; Vernon, 1950). Although the details of these conceptualizations differ, the hierarchical models all agree that intelligence consists of multiple tiers of factors. At the bottom of the hierarchical structure, there are test scores from diverse ability tests, which are organized by (and assumed to be causally influenced by) a number of narrow ability factors. Similar to raw test scores, these narrow ability factors are also intercorrelated, and thus, are assumed to be causally influenced by several broad ability factors at a higher order above the narrow abilities. Depending on the specific theory, the number of broad ability factors ranged from two (e.g., Vernon, 1950) to sixteen (e.g., McGrew, 2005). Some of the hierarchical theories went further by proposing a general factor (as Spearman's *g*) at the top level as a latent common cause of intercorrelation among all broad abilities (see Figure 3a), while other theories retained a correlated broad abilities level (see Figure 3b). These models provided not only accommodations between the generalist view and the specificity view but also better accountability of data across batteries of tests and samples.

Figure 3

A Conceptual Illustration of the Two Types of Hierarchical Models of Intelligence



Note. Panel a illustrated a higher-order multi-factor model in which a higher-order g was proposed as a latent common cause of all second-order broad ability factors. Panel b illustrated the correlated multi-factor model with the same set of broad ability factors, in which the broad ability factors were correlated and no higher-order g was posited. For a proof of concept, only three broad ability factors, with each of them accounting for the common variance of three narrow abilities, were included in this figure.

To date, one of the most influential hierarchical models is the Cattell-Horn-Carroll (CHC) model, which integrated two of the most widely recognized hierarchical theories of intelligence: Cattell-Horn's (Extended) Gf-Gc theory (Cattell & Horn, 1978; Horn & Blankson, 2012) and Carroll's Three-Stratum theory (Carroll, 1997; McGrew, 2009).

Cattell (1943, 1963) posited that performance in intelligence tests is impacted by two types of intelligence, fluid intelligence (Gf) and crystallized intelligence (Gc), which are identified with different properties for educational and clinical prediction. In this theory (Horn & Cattell, 1966; see also Hunt, 2010), fluid intelligence represents the ability to solve novel problems by reasoning, and is associated with individual differences in biological and neurological factors; crystalized intelligence represents the ability to apply acquired knowledge to current problems, and is associated with personal experience, education, and culture. This two-factor structure was supported by the age difference observed in the two types of intelligence where fluid intelligence declines steadily after late adolescence, while crystalized intelligence increases with age (Horn & Cattell, 1967). The Gf-Gc theory was later extended with additional broad abilities (Horn, 1968; Horn & Noll, 1997), such as Visual-spatial ability (Gv), Short-term memory (Gsm), Long-term storage and retrieval (Glr), and Cognitive processing speed (Gs), based on evidence from psychometric, neurocognitive, and developmental research (Kaufman et al., 2013). In the extended Gf-Gc theory, narrow abilities were identified and categorized under 8 to 10 broad abilities, but a higher-order *g* was not proposed above them (see McGrew, 2009).

Carroll (1993) took another path and applied a systematic exploratory factor analysis on a collection of cognitive ability testing datasets and summarized a three-stratum hierarchy of intelligence. Based on his “meta-factor analysis” results, the three-stratum hierarchical factor model illustrated a 3-tier structure of diverse intelligence tests such that a higher-order general factor (Stratum III) was reflected by the intercorrelations of eight broad factors (Stratum II; Sattler, 2008; McGrew, 2009): Fluid intelligence (Gf), Crystalized intelligence (Gc), General memory and learning (Gy), Broad visual perception (Gv), Broad auditory perception (Gu), Broad retrieval ability (Gr), Broad cognitive speediness (Gs), and Processing speed (Gt). The eight broad ability factors organized approximately 70 narrow abilities at Stratum I, which were measured by specific intelligence tests.

Given the similarity in their factor structures and identified broad abilities, the Gf-Gc theory and the Three-Stratum theory led to the development of the Cattell-Horn-Carroll (CHC) theory (Flanagan et al., 2007). The CHC theory formally introduced a psychometric framework that extended both the generalist view of general intelligence and the specialist view of multiple aspects of intelligence. However, it is important to note that the two theories deviate with respect to the concept of *g* (McGrew, 2005). While the extended Gf-Gc theory regards *g* as a statistical artifact but not necessarily a psychological construct, the Three-Stratum theory emphasizes *g* and its superiority in the hierarchical structure. This deviation remains in the integration of the two theories, which eventually led to a de-emphasis of a psychological *g* in the contemporary CHC theory.

In its contemporary form (Schneider & McGrew, 2012), CHC theory proposes a total of 16 broad abilities of intelligence at Stratum II, and over 81 narrow abilities at Stratum I. However, Stratum III with a higher-order *g* factor is not explicitly specified but remains "an option for any assessor who wishes to use it" (Kaufman et al., 2013; Ortiz, 2015, p. 226; see also Schrank & Wendling, 2018). The 16 broad abilities are Gf (fluid intelligence), Gq (quantitative knowledge), Gc (crystallized intelligence), Grw (reading and writing), Gsm/Gwm (short-term memory or working memory), Gv (visual processing), Ga (auditory processing), Glr (long-term storage and retrieval), Gs (processing speed), Gt (decision speed/reaction time), Gkn (domain-specific knowledge), Go (olfactory ability), Gh (tactile ability), Gp (psychomotor ability), Gk (kinesthetic ability), and Gps (psychomotor speed). Recently, the CHC theory has been influential in the construction of many current IQ tests such as the Stanford-Binet Intelligence Scale (5th Edition) and Woodcock-Johnson III and IV tests (Roid & Pomplun, 2012; Schrank, 2011; Schrank et al., 2016), and has been recognized as a comprehensive account of the correlational structure of diverse intelligence tests across test batteries (Conway & Kovacs, 2015; Flanagan et al., 2013).

2.1.4 Summary

The positive manifold is one of the most robust findings in psychology and it is widely accepted, across any sample or test battery, that intelligence tests are all positively correlated. Intuitively, the positive manifold is often interpreted as a common cause in all intellectual abilities. Psychometric theories rely primarily on latent factor approaches to define and investigate this “common cause” of human intelligence. After 100 years of debate on general and/or specific abilities of intelligence, the current version of traditional psychometric models, namely the CHC model, has been widely accepted as a comprehensive account of a descriptive structure of human intelligence. However, despite a good descriptive account that summarizes variations in intelligence, the psychometric theories are not sufficient to explain the structure of intelligence they describe.

2.2 Cognitive Theories of Intelligence

Psychometric theories are focused on the structure and dimensions of human intelligence but are limited in their capability to interpret individual differences in terms of the exact cognitive processes that influence cognitive performance (Hunt, 2010). Thus, based on evidence from experimental studies, cognitive theories have been developed along with psychometric theories to understand the psychological properties of *g* and/or broad ability factors from a perspective of information processing. Conventional cognitive theories either focused on (a) identifying the information processing abilities that are associated with psychometric factors of intelligence by correlational analyses (the cognitive correlates approach) or (b) directly decomposing the processes involved in typical tasks of intelligence by disassociating task performances under different experimental manipulations (the cognitive component approach; Pellegrino & Glaser, 1979; Sternberg, 1985). Both approaches rely on evidence from elementary cognitive tasks (ECTs) that are relatively knowledge-free (Anderson, 1992), and have pinned down a variety of cognitive processes or mechanisms that are essential in different intellectual activities. To date, three of the most frequently discussed cognitive processes/mechanisms are processing speed, working memory, and executive functions (Conway et al., 2003; Friedman et al., 2006; Jensen, 1987). While it is commonly agreed that these three mechanisms are all important predictors of individual differences in intelligence (Frischkorn et al., 2019), there is still limited consensus on their exact roles and relationships to intelligence and between each other.

2.2.1 Processing Speed

The idea of individual differences in brain speed and efficiency of information processing being the basis of individual differences in intelligence is simple but plausible: if the brain has a limited capacity to process and store concurrent information, slower processing would result in more information loss compared to faster processing, especially for complex tasks (Nettelbeck et al., 2020). It has been argued that, given that all mental actions depend on neural processing, individual differences in processing speed (PS) at the neural level would cause individual differences in processing speed at the behavior level regardless of task type (Hunt, 2010; Jensen, 2006). Consequently, this would lead to individual differences in psychometric measures of intelligence. Furthermore, PS has also been associated with a decline in fluid ability from an aging perspective (Salthouse, 1996). That is, the slowing of processing speed in older age results in a decline of the intellectual functions. The functions that hinge more on processing speed, such as fluid reasoning, are more influenced than other functions, such as verbal ability.

Processing speed is operationalized and measured by the minimum time required to respond correctly to a specific target of simple item identification or discrimination tasks, such as inspection time tasks (Nettelbeck, 2003) and reaction time tasks (Jensen, 2006). PS parameters are estimated from subjects' timed performance in different task conditions that are designed to reflect the cognitive processes involved. For example, in a typical inspection time paradigm (e.g., Burns et al., 1998), stimuli to discriminate (e.g., two vertical lines with different lengths) are presented to subjects, followed by masks that overlay the original figures. The amount of time lapse between the target stimulus and the mask, defined as stimulus onset asynchrony (SOA), varies across trials. Inspection time is therefore estimated by a function of SOA across trials and response accuracy, and it is considered to represent mental processing speed that is free from psychomotor and other confounds (Jensen, 2006). In a typical reaction time paradigm (Jensen, 1987; Jensen & Munro, 1979), subjects react to colored lights presented as stimuli by pressing buttons in corresponding colors as fast as they can. The number of stimuli and types of action change across conditions: in the simple reaction condition, subjects only need to watch for one possible colored light to be lit; in the choice reaction conditions, subjects need to watch for multiple possible lights (e.g., from two to eight) with different colors. Mean reaction times for different conditions are calculated and divided up based on the types of conditions. For example, reaction time in the simple reaction condition represents basic movement time for execution, while the time delays in choice reaction conditions, compared to the simple reaction condition, represent the time required for decision-making given a variety of possible choices.

These performance estimates of PS have been found to be consistently correlated with intelligence regardless of the measures; subjects with faster processing speed generally perform better on intelligence tests, but these correlations are mostly moderate in magnitude and typically range from .20 to .50 (Deary, 2003; Grudnik & Kranzler, 2001). Therefore, although PS is clearly associated with psychometric intelligence, it cannot explain the majority of variance in human intelligence. It is argued that parameters from these speed tasks may reflect different cognitive processes for subjects from different ability levels and are psychologically complex (Lally & Nettelbeck, 1980; Larson & Alderton, 1990; Nettelbeck, 2001). However, recent studies have applied more advanced statistical and experimental methods to investigate PS and have made further progress in understanding this association. For example, the drift rate from the diffusion model was suggested as a better indicator of PS, which increases the estimated association between processing speed and intelligence (Schmiedek et al., 2007; Schubert et al., 2015).

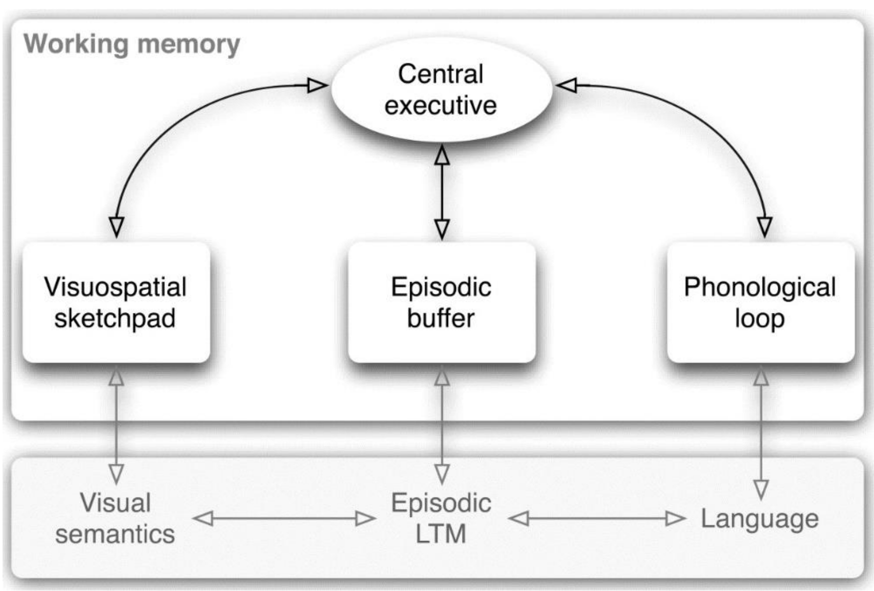
2.2.2 Working Memory

Working memory (WM) is conceptualized as a system of cognitive processes that maintain access to a limited amount of temporary information to be available for ongoing information processing (Cowan, 2017). The research interest in WM and its relationship with intelligence is closely associated with that of processing speed, in which the information processing system is generally considered to have limited capacity. Thus, individuals with a higher ability to maintain, update, and manipulate information will reveal higher performance on intelligence tests, especially in the face of distraction and competing information (Kaufman et al., 2013).

The early conceptualization of WM, namely the multiple-component model of working memory (Baddeley & Hitch, 1974; Baddeley & Logie, 1999), was proposed as an alternative to the short-term store to explain performance in a range of complex cognitive behaviors (Conway & Kovacs, 2020). The multiple-component model describes working memory as a multi-component system (Figure 4) that includes a domain-general mechanism responsible for attention control, and several "slave" subsystems for domain-specific storage. Contemporary theoretical accounts of working memory have refined the initial multiple-component model and focused more on clarifying the definition of the central executive and its information-processing mechanism (for a review, see Conway & Kovacs, 2020).

Figure 4

Baddeley's Multi-Component Model of Working Memory (from Repovš & Baddeley, 2006, p. 6)



Various tasks have been developed to measure individual differences in working memory, quantified and referred to as working memory capacity (WMC). These WM tasks usually engage memory recall and processing of other information simultaneously, such as the complex span tasks (Conway et al., 2005), scope of attention tasks (Luck & Vogel, 1997), coordination and transformation tasks (Oberauer, 2005), and N-Back tasks (Owen et al., 2005).

Similar to the positive manifold in intelligence tests, WMC measures of all types are observed to be highly intercorrelated (Oberauer et al., 2018). Therefore, this WM positive manifold is analogous to the positive manifold of intelligence, such that the WMC measures are correlated because of a domain-general mechanism (as was proposed in Baddeley's model), namely the central executive. The central executive serves as an attentional bottleneck that coordinates information from domain-specific storages, causing intercorrelations among different WMC measures. Thus, the psychometric general factor of WM represents an individual differences construct reflecting the cognitive property of the central executive (Conway et al., 2003). However, correlational evidence has also shown that WMC is not best accounted for by a unitary factor but by multiple tiers of lower-level factors or a hierarchical structure (Kane et al., 2004; Kovacs et al., 2019), indicating that individual differences in WM may also reflect domain-specific properties. Despite the debate on whether individual differences in WMC reflect more domain-general control or domain-specific maintenance, WMC measures have been demonstrated to strongly correlate with psychometrically measured intelligence, especially with fluid intelligence (Gf), with correlations ranging from r = .70 to .90 (Frischkorn et al., 2019; Oberauer et al., 2018), indicating that working memory tasks and intelligence tests tap into similar, if not identical, cognitive functions.

Overall, WM has been shown to play a central role in higher-order cognitive abilities (Cowan, 2005). Cognitive theories of working memory attempt to explain the relationship between WMC and Gf, with one or several central cognitive processes that are shared among WMC measures and intelligence tests. For example, executive attention theory (Engle & Kane,

2004) emphasized that individual differences in attention control resources drive the intercorrelations among WMC measures and the relationship between WMC and Gf. Additionally, Oberauer (2009, 2019a) emphasized that working memory is an attention selection system to memory representations, with temporary bindings of objects/features, and individual differences in WMC reflecting individual differences in a specific limit on establishing and maintaining these bindings (Oberauer, 2019b). Although cognitive theories of WM still do not agree upon an exact emphasis on a cognitive mechanism that explains the association between WM and intelligence, it is plausible that both WM and fluid intelligence tests require multiple domain-general cognitive mechanisms that are essential to the active maintenance and rapid retrieval of information (Conway & Kovacs, 2020).

2.2.3 Executive Functions

The cognitive theories of working memory have already emphasized the essential role of the domain-general central executive in complex human behaviors. A more recent cognitive construct, compared to PS and WM, executive functions (EFs) were initially presented as the attention control mechanisms in the central executive (Baddeley & Hitch, 1974). From broad to narrow, EFs have been defined as either a range of higher-order cognitive processes involved in goal-directed cognition (Barbey et al., 2012; Diamond, 2013), such as reasoning, problem-solving, and working memory, or the cognitive processes that specifically supervise attention during goal-directed cognition (von Bastian et al., 2020).

Current cognitive research on EFs tends to use the term to refer to its narrower definition, which is the supervisory attention control processes that are identified from experimental effects of tasks that require attention regulation. Three executive functions have often been recognized in empirical works: inhibition, shifting, and updating (Friedman et al., 2006; Miyake et al., 2000). Per the descriptions outlined in Miyake et al. (2000), inhibition refers to the ability to focus attention on relevant information while suppressing irrelevant information, shifting refers to the ability to switch between different tasks, and updating refers to the ability to incorporate new information and remove outdated information in working memory. These different EFs are distinguished by different tasks and/or experimental conditions that contrast between higher or lower requirements of these attention regulation mechanisms. Empirical results have indicated that these measures are correlated to some extent, but it is still not clear the degree to which these functions overlap (Friedman & Miyake, 2017; Karr et al., 2018).

While the idea is plausible that EFs are a set of essential, domain-general cognitive mechanisms underlying WM and are thus required to engage in a broad range of complex cognitive activities, empirical evidence for the relationship between EFs and intelligence is mixed. Some studies have found moderate to strong associations between EFs and intelligence (e.g., Friedman et al., 2006; Kane & Engle, 2003), but other studies failed to support such associations (e.g., Frischkorn et al., 2019; Rey-Mermet et al., 2019). Recent research has started to investigate the reliability of current EF measures and to develop "process-pure" measures of EFs by further isolating the components within these EFs. For example, Ecker et al. (2014) introduced a method for isolating encoding and removing speed from the original updating measures. However, there is not yet a consensus on the strength nor the mechanisms associated with the relationship between EF and intelligence.

2.2.4 Summary

In empirical research, it is challenging to associate the differences in cognitive task performance with specific cognitive processes, as most cognitive tasks are complex in nature. Contrary to most psychometric theories that attempt to explain complex intellectual activities with one or a few latent factors as the common cause, cognitive theories narrow down to simpler forms of intellectual activities, in which specific cognitive processes/mechanisms are engaged as basic information-processing components (Gardner, 2011; Sternberg, 1985). Most cognitive theories had been influenced by reductionism. These theories have attempted to deconstruct intellectual behaviors into their basic information-processing components and find the "most influential" cognitive processes/mechanisms of intelligence. From processing speed to working memory to executive functions, cognitive theories have been investigating the cognitive processes/mechanisms of interest with more specific measures and clearer models. However, the cognitive processes/mechanisms of interest are still quite dependent on the tasks and measures developed based on theoretical conceptualization.

3 Problems with the Traditional Theories

3.1 Problems with Psychometric Theories

While psychometric theories have tried to assemble an exhaustive list of representative cognitive abilities that represent the structure of human intelligence, there are several problems with this approach. Most traditional psychometric theories are based on a broad premise: common variance among different measures reflects a general source that causes the communality of the individual differences, with unique variances reflecting task/measure specific resources. This premise has led to a potential ambiguity between "statistical *g*" and "psychological *g*" in early theories and has, in fact, fueled the debate associated with the psychological interpretation of general intelligence. Also, the traditional psychometric theories usually explain interpersonal individual differences well but are not as good at explaining intrapersonal individual differences, because most of the psychometric theories are based on task-level correlations, which are usually estimated from composite scores of items with invariant formats and properties within each task. These items are regarded as equivalent in a task and the variation among items in the same task is usually perceived as measurement error, as is the variation among tasks of a similar type.

3.2 Problems with Cognitive Theories

On the other hand, the traditional cognitive approach has been investigating exact processes and mechanisms in intellectual behaviors but is also limited by its own issues. One of the fundamental challenges for the cognitive theories/models is the measurement problem (Frischkorn & Schubert, 2018). The cognitive models are derived from empirical research using elementary cognitive tasks (ECTs) that are designed to reflect individual differences in specific simple cognitive processes. However, ECTs in the real world can hardly be “process pure”, such that one task only reflects individual differences in a specific cognitive process. Two analytic approaches are heavily involved in cognitive research to address this problem. The “cognitive-correlates” approach investigates individual differences in cognitive processes by latent factor models on correlational data from multiple cognitive measures on the same process. The latent factors that reflect “pure” estimates of the target process are associated with latent factors in psychometric models of intelligence. This approach relies on the same premises as the psychometric models and requires large batteries of tasks. On the other hand, the “cognitive-components” approach estimate “pure” individual differences in cognitive processes by comparing performance in different experimental conditions of a task, in which the conditions are designed to require different combinations of processes. This approach usually assumes that the different processes in different experimental conditions are affecting performance without any interactions with each other, or with other confounding processes such as top-down strategies (Deary, 2003). It also often requires more trials and multiple experimental conditions to estimate reliable intrapersonal individual differences.

3.3 Problems with The Gap Between the Two Approaches

In addition to the aforementioned problems, traditional psychometric and cognitive approaches are hardly compatible with each other. As aforementioned, due to the complex nature of intelligence, traditional psychometric and cognitive theories use different goals to understand intelligence. Psychometric theories, especially recent hierarchical theories such as the CHC theory, offer comprehensive descriptions of the structure of intelligence. However, psychometric theories cannot specify the exact cognitive processes and operational mechanisms that prompt the correlational structures of intelligence. For instance, the CHC theory includes processing speed (Gs) as a broad ability factor of intelligence. In the CHC theory, the role of processing speed in the CHC theory is only represented by its correlations with the other broad abilities or its estimated loading with higher-order *g*. However, a latent factor of processing speed does not really describe how the corresponding cognitive processes or mechanisms of processing speed that are engaged in the information processing workflow produce intellectual behaviors. On the other hand, cognitive theories deconstruct cognitive behavior by isolating essential cognitive processes but are not good at explaining how these “isolated” cognitive processes (especially across different theories) cooperate with each other and contribute to intelligence as a system. Taken together, there is still a gap between traditional psychometric and cognitive theories in explaining how basic information processing abilities function in the face of different external problems and how they can be systematically translated into responses in intellectual behaviors. Thus, new intelligence research perspectives, with updated theoretical inference and advanced computational realization, are needed to bridge the gap.

3.4 Modern Instruments to the Rescue?

The development of modern statistical instruments may help to address the aforementioned problems of traditional intelligence theories by providing alternative modeling techniques with different assumptions and benefits. These modern instruments could also lead to alternative interpretations of correlational/experimental data of intelligence research, for both exploratory and confirmatory purposes. For example, item response theory (IRT) is proposed as an alternative psychometric approach to classic test theory (CTT). IRT estimates a respondent's level of ability on underlying attributes, or latent traits, by creating a corresponding psychometric estimation based on maximized information from individual items (Brown & Croudace, 2015), independent of the actual test. This differs from classical test theory (CTT) in that the focus of CTT is to estimate a respondent's level of ability based on a composite score of all items from the test. IRT, on the other hand, incorporates more information about each item within the test to produce a more accurate depiction of the target traits and constructs. In other words, IRT investigates not only person properties, but also item properties. In fact, IRT provides a detailed examination of differences in respondents' performance and test item properties, such as item difficulty and item discrimination, in the same model. Furthermore, a confirmatory IRT model may also include parameters that can represent variables in information-processing mechanisms proposed by cognitive theories (Embretson & McCollam, 2000). Thus, IRT has been utilized in not only psychometric testing, but also in the statistical fitting of experimental data on information processing models such as the diffusion model in decision-making (e.g., Molenaar et al., 2015).

Another trending technique that can be applied to modern intelligence research is the psychometric network analysis (Borsboom et al., 2021). The network approach can be applied to multivariate data without an assumption of a latent common cause. In latent factor models, the covariances among observed variables are assumed to be caused by a common source that is represented by the common variance. In network models, complex systems of the observed variables are represented as interconnected networks. This technique complements a factor analysis approach by focusing on the patterns of pairwise conditional dependencies among observed variables rather than the dimensional reduction of multivariate data. The observed positive manifold in traditional psychometric theories is therefore regarded as an emergent property resulting from interactions among the observed variables (e.g., performance on specific tests or task conditions).

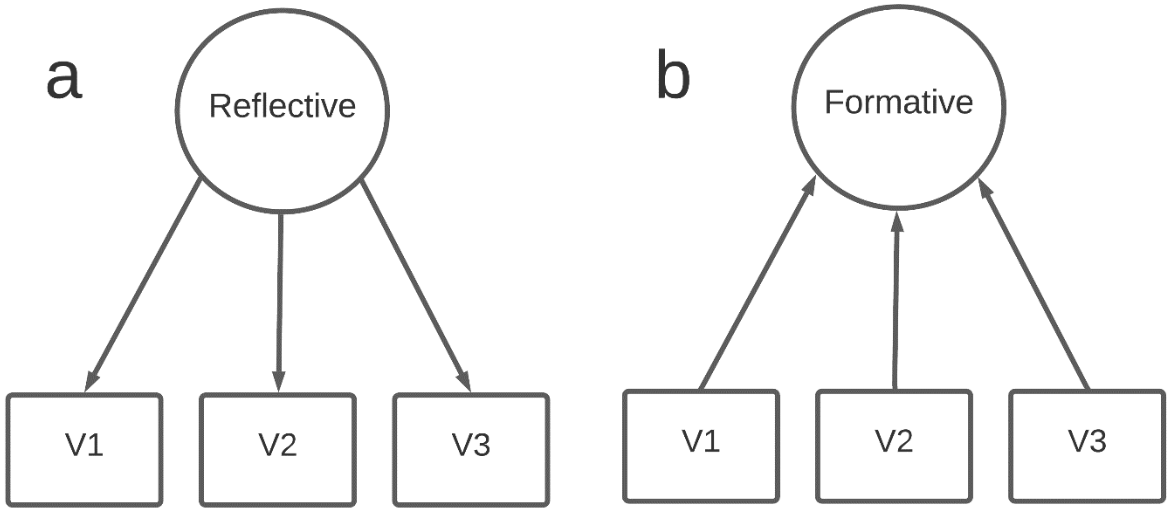
4 Process Overlap Theory: Bridging Psychometric and Cognitive Theories

Given the growing awareness of the problems in traditional theories of intelligence, contemporary frameworks to explain intelligence have been proposed based on new computational techniques and theoretical perspectives. Contemporary theories propose an alternative explanation of intelligence to address the problems in traditional theories, and more importantly, to provide quantifiable frameworks that comply with the observed phenomena at psychometric and cognitive levels. Specifically, some contemporary theories attempt to explain the positive manifold of intelligence without relying on the premise that the intercorrelations among all intelligence tests reflect a common cause of intelligence. The dynamic mutualism theory (van der Maas et al., 2006; Kan et al., 2019), for example, provides a network explanation of general intelligence from a developmental perspective, inspired by dynamic explanations of ecosystems such as consumer-resource systems (Savi et al., 2019). The idea of the dynamic mutualism model is that human cognition consists of many basic processes that are connected in a network with (mostly) positive interactions. These processes have mutually beneficial or facilitating relations throughout development, and their interactions are reinforced because of a collaborative application in different intellectual activities. This theory rejects the psychological *g* as a cognitive construct that causes the common covariances among intelligence tests. Instead, the correlations among the intelligence tests are caused by the intercorrelations of basic cognitive processes during the development of the cognitive system.

Another contemporary theory of intelligence, and the focus of this section, is the process overlap theory (Kovacs & Conway, 2016). Motivated by the sampling theory of intelligence (Thomson, 1916, 1951; see also Bartholomew et al., 2009) and cognitive psychology research on working memory, Kovacs & Conway (2016; 2019) introduced a unified theory of intelligence, the process overlap theory (POT). The theory attempts to bridge the gap between psychometric and cognitive approaches by offering a new perspective on the cognitive foundation of the positive manifold of intelligence and Spearman's *g*. POT regards *g* as a formative factor instead of a reflective one. (See Figure 5 for a review of the difference between the formative and reflective models; see also Kievit et al., 2019). It exhibits a computational framework of intelligence that does not rely on the latent common cause assumption of factor analysis, but rather applies a sampling mechanism of cognitive processes, in which domain-specific and domain-general processes are sampled in an overlapping manner across a battery of tests (Kovacs & Conway, 2016). As in the mutualism theory, POT rejects the notion of psychological *g* as a reflective cognitive construct. However, while the mutualism theory proposes an interactive network mechanism among cognitive processes during development that could lead to the emergence of a positive manifold and *g*, POT proposes an alternative sampling mechanism of cognitive processes during specific intellectual activities, which eventually leads to a positive manifold and *g* across a battery of intelligence tests (Kovacs & Conway, 2016).

Figure 5.

The Reflective and Formative Latent Variable Models



Note. In a reflective factor model (panel a), the reflective common factor is interpreted as the common cause of the indicators. However, in a formative factor model (panel b), the formative common factor, or more accurately, the statistical artifact, is interpreted as an index that summarizes the indicators. The different directions of the arrows in the two panels reflect the different directions of causality.

In the present paper, process overlap theory is described at four different levels of representation: (a) the conceptual model that is verbally described (POT-V), (b) the multi-dimensional item response model (POT-I) that presents a computational framework based on item response theory, in which lower-level cognitive processes are sampled and reflected in higher-level behavior performance, (c) the structural model (POT-S) that represents the intercorrelations of higher-level behavior performance based on the sampling algorithm in a traditional latent factor structure, and (d) the psychometric network model (POT-N) that represents the same intercorrelations, but in a psychometric network that does not rely on the assumption of latent common causes of task performance.

4.1 The Conceptual Model (POT-V)

According to the early sampling theories (Thomson, 1916; Thorndike, 1925), the general factor of intelligence (*g*) does not represent a common cause of intelligence but a sampling system of a large number of mental elements that each add a small amount of influence. Thus, *g* is only a statistical artifact that emerges from a pattern of positive correlations across tests. The correlations are caused by an overlap of the elements, or "mental bonds", sampled by different tests.

Process overlap theory adopted this sampling explanation of the positive manifold, but extended the initial sampling theories by proposing a nonadditive sampling mechanism of processes. Specifically, in POT, any battery of intelligence tests requires a number of domain-general processes and domain-specific processes. While the domain-general processes, such as the cognitive processes involved in working memory, are sampled by all kinds of tests, the domain-specific processes are sampled only in some of them. In this sampling mechanism, there is no single process that is required by all items and tests. Instead, some processes, the domain-general ones, are sampled more frequently than the domain-specific processes. Thus, the overall and the clustered intercorrelations observed in intelligence tests are caused by an overlap between the cognitive processes tapped by different tests. Furthermore, while the processes within the same domain are compensatory, the processes across domains are non-compensatory. Therefore, the limits posed by individual differences in domain-general processes will function as a bottleneck to overall performance (Kovacs & Conway, 2019).

4.2 The Multidimensional Item Response Model (POT-I)

This conceptual model of POT is represented in a multi-dimensional item response (mIRT) model, which is referred to as POT-I (see Equation 1).

$$P\left(U_{pi} = 1 \mid \Theta_{plm}, a_{il}, b_{il} \right) = \prod_{l=1}^{D} \frac{e^{\sum_{m=1}^{C} a_{il}(\Theta_{plm} - b_{il})}}{1 + e^{\sum_{m=1}^{C} a_{il}(\Theta_{plm} - b_{il})}} \quad (1)$$

According to POT-I, the probability (*P*) of a person (*p*) answering an individual test item (*i*) correctly is a function of their ability level (Θ) on the processes required by that item, as well as the discrimination and difficulty parameters for that item. More formally, Θ_{plm} represents the level of ability for the *p*-th individual on the *m*-th process in the *l*-th domain; a_{il} is the discrimination parameter for the *i*-th item in the *l*-th domain; b_{il} is the difficulty parameter for the *i*-th item in the *l*-th domain. *D* is the number of domains sampled by an item and *C* is the number of processes in a given domain sampled by an item.

In the POT-I algorithm, the cognitive processes are compensatory within domains but non-compensatory across domains. Specifically, the cognitive processes constitute the assumed latent traits required for a specific item. These cognitive processes are sampled and composed in two compensatory manners: additive or non-additive. For a response to a specific item, the ability levels (Θ) on the processes within the same domain are additive, but each domain of a test item functions as a separate dimension, and each of the dimensions has to be passed independently for the item to be solved correctly (reflected as the multiplication function in Equation 1). For example, if a visuospatial test requires processes in two domains, executive and spatial processes, then a weakness in a spatial process can be compensated by another spatial process, but such compensation cannot take place across domains. Thus, according to POT-I, for a specific item, the observed performance reflects a function of multiple domain-general abilities and multiple domain-specific abilities (the ability levels are expressed in Equation 1 as Θ_{plm}).

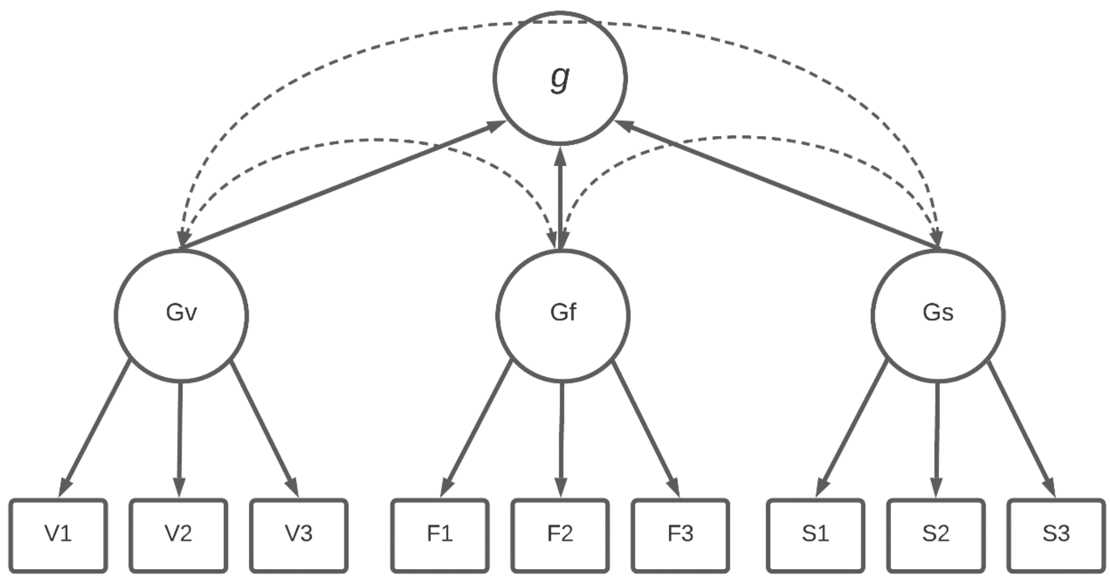
4.3 The Structural Model (POT-S)

According to POT, the multidimensional item response model described in POT-I could result in positive task-level intercorrelations among different cognitive tasks without any single cognitive process being sampled across all tasks. The patterns of these intercorrelations are expected to exhibit a positive manifold similar to the positive manifold observed in conventional psychometric theories. Traditional psychometric theories apply exploratory or confirmatory factor analyses to these correlations to find the common factors. These common factors are reflective and are interpreted as the broad ability(s) of intelligence that function as the common causes of the correlations. Similar to traditional psychometric theories, POT also predicts a higher-order factor structure of psychometric intelligence. However, POT interprets the commonly observed higher-order *g* factor in traditional psychometric models of intelligence as a formative factor that represents the common consequence of the positive manifold, instead of a reflective factor that represents the common cause (Kovacs & Conway, 2016, 2019). Thus, in POT-S the direction of causality between the broad ability factors and *g* should be opposite to that in the CHC theory. The opposite direction of causality indicates that the psychometric *g* is represented as a weighted combination of lower-level broad abilities as indicators. The psychological inference of this formative higher-order factor is dependent on the selection of indicators (intelligence tests). In other words, the individual difference in *g* is defined and determined by the cognitive measures included in a certain psychometric model of intelligence.

This formative model takes a more conservative perspective on the causal inference of *g* in a similar way as Boring (1923) stated, that intelligence is what the intelligence tests test. Thus, POT-S illustrates a factor structure of psychometric intelligence in which the subfactors, representing broad abilities measured by different tests, are reflective latent factors, while the higher-order g is a formative factor emerged from the subfactor correlations, which are caused by the sampling mechanisms proposed in POT-I. See Figure 6 for a conceptual illustration of POT-S, in which a higher-order formative g emerges from three broad abilities: Fluid Reasoning (Gf), Verbal Ability (Gv), and Spatial Ability (Gs).

Figure 6

A Conceptual Illustration of the Structural Model of Intelligence Based on POT-S



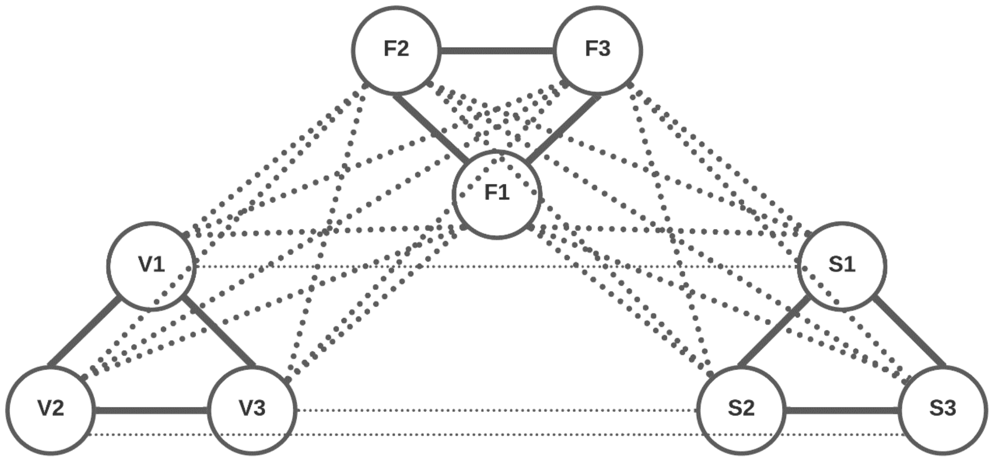
4.4 The Network Model (POT-N)

Due to their differences in the placement of the common cause assumption, the latent factor approach is not fully compatible with the theoretical perspective of POT. As previously mentioned, POT interprets *g* as a result of the sampling mechanism on cognitive processes. Importantly, none of the cognitive processes is exclusively involved across the entire battery of tests. Therefore, from a theoretical perspective, a network model (Kan et al., 2019) is more compatible with POT at the task level than a latent variable model, such that the *g* factor is not assumed to be the cause of the covariance among intelligence tests but is caused by an overlap of processes tapped by tests. Thus, the covariance structure among tests is represented in a network structure where the observed test performances are referred to as nodes and their connections are referred to as pairwise edges. Recent developments in the psychometric network approach (Borsboom et al., 2021; Epskamp et al., 2018; van der Maas et al., 2017) have made it possible to estimate conditional dependency among the tests as well as assess the stability and accuracy of the network structures, without assuming latent common causes.

Figure 7 presents a conceptual network structure based on POT-N. The conceptual model describes the same battery of cognitive measures as those in Figure 6 but with the nodes (the circles) representing the cognitive measures and the edges (the lines between the circles) representing partial correlations among them. Partial correlations reflect the remaining association between two nodes after controlling for all other information available (Epskamp & Fried, 2018). POT-N proposes that the broad abilities (Gf, Gv, and Gs in this example) are reflected in the small "clusters" of nodes, connected by strong edges among the test scores that are the same test type. The network model does not assume any latent factor to represent the covariance structure and therefore is more compatible with POT-N.

Figure 7

A Conceptual Illustration of the Network Model of Intelligence Based on POT-N



Note. The circles (nodes) represent the cognitive measures in the battery of tests. The lines (edges) can be regarded as pairwise partial correlations between two nodes.

However, it is worth mentioning that the psychometric network model presented in POT-N does not indicate that a network model can validate POT by itself (Conway et al., 2020). For a set of behavioral observations of intelligence (test scores), a network model, with nodes as subjects' performance on specific tests and edges as their relationships, is more compatible with POT than a reflective latent factor model, with manifest variables representing test performances and factor loadings representing their reliabilities of reflecting the latent construct (intelligence). Still, the network model does not represent the underlying cognitive mechanisms that POT describes, because the behavioral observations are also not direct representations of the sampled cognitive processes, and the between-test conditional relationships revealed by network models do not reveal the between-process conditional relationships.

5 Interim Summary

The past 100 years of research on intelligence have motivated various theories to explain the major phenomena observed in human intellectual behaviors. However, the focus of different theories has diverged during the progress of research to serve different purposes (Sternberg, 2019). As was mentioned, most of the traditional intelligence theories have been conceptualizing human intelligence by either making a "geographic map" of the human mind or by illustrating a "workflow system" of information processing. However, these traditional psychometric and cognitive theories have encountered theoretical or methodological problems. Process overlap theory, among other contemporary theories, attempts to explain intelligence and the corresponding major phenomena from a new perspective and with a computational framework that could, to some extent, unify the traditional psychometric and cognitive models (Kovacs & Conway, 2019). Specifically, POT proposes information-processing mechanisms from the cognitive perspective and maps the cognitive mechanisms into correlational structures of latent ability factors proposed by psychometric theories.

The current studies will focus on developing dynamics to investigate the cognitive and psychometric structure of intelligence from a process overlap theory's perspective. Specifically, this dissertation will start with a data simulation approach based on POT in Study 1. This simulation is developed to conceptually prove the sampling mechanism of POT at both the cognitive level and the psychometric level. The simulation approach also assesses if a formative *g* and a hierarchical structure of intelligence can emerge in the absence of a psychological common cause (Conway et al., 2021). The simulation applies a computational algorithm based on the central claims proposed in POT-V and POT-I, in which cognitive processes are sampled and overlapped in different intelligence tests, and the psychometric results of the simulated data, in the form of latent factor models (POT-S), are presented and investigated.

In Study 2, the simulation dynamics based on POT algorithms are examined by psychometric network models (POT-N). This study applies undirected network models to the simulated data and exhibits an alternative representation of the positive manifold among cognitive test scores. The network structure emphasizes direct mutual associations among cognitive measures over an unobservable common cause and is more compatible with POT. Furthermore, a latent network approach (Epskamp et al., 2017) is also applied to the simulated data, in which the benefits of latent factor models and psychometric network models are combined. The latent network model accounts for the concern of measurement error and retains the properties of latent variable models at the broad ability level, and represents the conditional dependence of different broad abilities in the form of an undirected network.

Study 3 further investigates the applicability of the network psychometrics on intelligence by applying the network models to existing empirical data. Specifically, an exploratory network analysis is conducted on a correlational dataset of cognitive abilities from Kane et al. (2004), in which the relationship between working memory and reasoning was investigated using structural equation modeling techniques. The network models illustrate the covariances among test scores of the cognitive tasks in a different manner compared to CFA and SEM models, which leads to alternative interpretations of the correlational associations among cognitive abilities represented by the tasks.

6 Study 1: A Simulation Framework Based on the POT-I Algorithm

This simulation attempts to translate the conceptual and IRT model of POT to sampling algorithms on randomly generated matrices, which represent subjects' individual differences in all the cognitive processes that are relevant to designated cognitive tests. This simulation dynamic is designed to show that, with the proper sampling mechanisms on random matrices and multidimensional item response model in POT-I, a hierarchical *g* structure of intelligence, which is commonly observed in psychometric theories (e.g., Cattell, 1963), can emerge without any general ability or specific intercorrelation of cognitive processes. In this study, subject scores on a battery of intelligence tests are calculated as a function of domain-general and domain-specific ability parameters. Test scores are simulated according to POT-V and POT-I, and are compared to a generic sampling model (GSM) in which all the cognitive processes are assumed to be equally weighted across the battery of tests. The GSM approach contrasts with POT such that GSM represents the lack of a distinction between general and specific cognitive processes.

For simplicity, the simulation specified three broad cognitive abilities: fluid, verbal and spatial. Each of the broad abilities is assumed to be measured by 3 tests. Each test consisted of 100 items. To obtain a psychometric model of intelligence with a higher-order *g* factor, a sample of 1000 subjects is simulated. Individual subject scores for the 9 tests are generated based on the POT algorithm and the GSM algorithm. Item scores (correct or incorrect; represented by 1 or 0, respectively) are determined by both domain-specific ability and domain-general ability in the two algorithms, which are represented by functions of simulated processes sampled by the item.

This simulation consists of four steps: (a) specification of the simulated cognitive processes and tests; (b) calculation of ability parameters (Θs) based on sampling algorithms; (c) calculation of item scores; and (d) task scoring and psychometric model fitting.

6.1 Step 1: Specification of Cognitive Processes and Tests

For each of the 1000 subjects, 60 cognitive processes are specified. The 60 cognitive processes are classified into four domains (15 per domain): fluid reasoning processes, verbal processes, spatial processes, and executive function processes. These are the four domains that are assumed to be involved in the 9 tests measuring 3 broad abilities (fluid, verbal, and spatial). Each subject is thought to have an “ability level” on each of the 60 cognitive processes. These cognitive process abilities are represented by numerical values from a standardized multivariate normal distribution, $X_p \sim \mathcal{N}_{200}$ ($\mu = 0$, $\Sigma = I$). This simulation step resulted in a 1000 × 60 matrix (number of subjects × number of processes) of orthogonal process ability values.

6.2 Step 2: Calculation of Ability Parameters (*Θ*s) Based on Two Algorithms

As previously mentioned, the POT model is based on the initial sampling theory by Thomson (1916). A simple elaboration of sampling theory is that the ability, *y* (in this simulation, *Θ*), required for an individual on a specific item of a test can be expressed as

$$y = \boldsymbol{a}^T \boldsymbol{x} = \sum_{k=1}^{n} a_k x_k \tag{2}$$

In Equation 2, *n* is the number of simulated cognitive processes in the corresponding domain, ***a*** is an *n*-dimensional row vector of random binominal values (0s and 1s) for process sampling, and ***x*** is an *n*-dimensional column vector that includes the *n* simulated cognitive processes in this domain for the individual. Therefore, the vector ***a*** expresses how potential processes of the domain are involved for the individual to respond to the item, with $a_k = 0$, meaning process x_i is not sampled, and $a_k = 1$, meaning process x_k is sampled in achieving the corresponding level of ability for the item. As follows, the ratio of 0s and 1s in ***a*** vector as the "sampling vector" determines the number of processes being sampled for a *Θ* value. Both the POT algorithm and the GSM algorithm follow this basic sampling dynamic in this simulation. Details for both algorithms are described as follows.

6.2.1 POT Algorithm

In the POT algorithm, each of the nine broad ability tests samples processes from two of the four different domains: one of the three types of domain-specific processes (fluid, verbal, and spatial processes for fluid, verbal, and spatial tests, respectively) plus the domain-general processes (executive function processes). While EF processes were allowed to be sampled in all three types of tests (fluid, verbal, and spatial), the other three types of processes could only be involved in their corresponding test. Hence, for each of the 100 item-level responses of an individual in a test, two aggregate process ability parameters (two Θs) were calculated: one domain-general and one domain-specific. The mathematical representation of the POT algorithm in Step 2 is described by the following equations (Equations 3 and 4):

$$\Theta g_{pti} = \sum_{n=1}^{15} b_{tin}\, G_{pn} \tag{3}$$

$$\Theta s_{pti} = \sum_{m=1}^{15} c_{tim}\, S_{pm} \tag{4}$$

In these two equations, Θg_{pti} and Θs_{pti} are the aggregate domain-general and aggregate domain-specific ability parameters for the *p*-th individual on the *i*-th item of the *t*-th test. The aggregate ability parameters are the sums of sampled processes from their corresponding domains (G as the domain-general EF processes and S as the domain-specific processes) based on the sampling vectors b_{ti} and c_{ti}.

An assumption of POT is that all cognitive tests sample domain-general processes and the probability of a domain-general process being sampled is higher in fluid reasoning tests than in other tests (Kovacs & Conway, 2016). In this simulation, for an item from a fluid reasoning test, the probability of an EF process being sampled is set to 0.28, while the probability of a fluid-specific process being sampled is 0.12, which on average sample about four EF processes and two fluid processes for an item on a fluid reasoning test. However, for an item from a domain-specific test (verbal or spatial), the probability of an EF process being sampled is set to 0.12, while the probability of a domain-specific process (verbal or spatial, respectively) being sampled is 0.28, which on average sample about two EF processes and four verbal/spatial processes for a domain-specific item. Either way, there are on average a total of 6 processes sampled when generating a simulated response to an item, but the EF processes are sampled more frequently in fluid reasoning tests than in domain-specific tests.

Finally, two 1000 × 9 × 100 (number of subjects × number of tests × number of items per test) three-dimensional arrays of aggregate ability parameters (Θg and Θs) are created as the result of this step for the 9 tests, representing the domain-general and domain-specific dimensions of traits required to respond correctly on the test items.

6.2.2 GSM Algorithm

For the generic sampling algorithm, no domains are specified when sampling the cognitive processes across the 9 tests, and thus only one aggregate ability parameter (Θ) is calculated for each of the 100 item-level responses for an individual on a test. Therefore, when generating simulated item-level responses for each test, the probabilities of the 60 cognitive processes being sampled for a corresponding process parameter (Θ) are all equal, as reflected in Equation 5:

$$\Theta_{pti} = \sum_{h=1}^{60} d_{tih}\, E_{ph} \tag{5}$$

In this equation, Θ_{pti} is the aggregate ability parameter for the p-th individual on the i-th item of the t-th test. Each aggregate ability parameter is the sum of sampled processes from the 60 processes of an individual based on the sampling vector d_{ti}. In this simulation, the probability of a cognitive process being sampled is set to 0.10. Thus, in concordance with the POT algorithm, an average of 6 processes are sampled when generating a simulated response to an item. One 1000 × 9 × 100 (number of subjects × number of tests × number of items per test) 3-dimensional array of aggregate ability parameters (Θ) is created as the result of this step for the GSM algorithm.

6.3 Step 3: Calculation of Item Scores

To generate item-level scores based on the POT algorithm and the GSM algorithm, different IRT functions are applied to the ability parameters generated from the two algorithms.

6.3.1 POT Algorithm

For the POT algorithm, a multidimensional IRT function is applied to the domain-general and domain-specific ability parameters (Θg and Θs). This function is a practical version of the conceptual POT-I model described in Equation (1). The mathematical representation is described by Equation (6).

$$P_{pti} = \frac{1}{1+e^{-(z(\Theta g_{pti})-bg_{ti})}} \cdot \frac{1}{1+e^{-(z(\Theta s_{pti})-bs_{ti})}} \qquad (6)$$

where P_{pti} is the probability of the *p*-th subject getting the *i*-th item in the *t*-th test correct; $z(\Theta g_{pti})$ and $z(\Theta s_{pti})$ are the standardized scores of Θg_{pti} and Θs_{pti}, respectively. In Equation (6), although processes are compensatory within a domain, Θg and Θs are not compensatory with each other. In this simulation, the discrimination parameters (*a* parameters) are set to 1 for parsimony across the 100 items and 9 tests. The difficulty parameters of the domain-general and the domain-specific processes (*b* parameters, b_g and b_s) for the 900 test items are randomly drawn from a standardized multivariate normal distribution, with each of the nine tests having average b_g and b_s of 0 and standard deviation of 1 and being independent of each other. Under Equation 6, the two three-dimensional arrays of Θg and Θs are transferred to one $1000 \times 9 \times 100$ (number of subjects × number of tests × number of items per test) three-dimensional array of probabilities. A simulated binary response (0 for incorrect response and 1 for correct response) on each item of each test for each individual is generated on the basis of the probabilities from the array. Thus, in general, a higher ability would generate a higher probability value to achieve the correct response to the item in a simulation.

6.3.2 *GSM Algorithm*

For the GSM algorithm, a unidimensional IRT function is applied to the sampled process values (Θe) because no domains are specified, and hence, only one ability parameter for each simulated response is calculated. The mathematical representation is described by Equation 7:

$$P_{pti} = \frac{1}{1 + e^{-\left(z\left(\Theta e_{pti}\right) - \frac{bg_{ti} + bs_{ti}}{2}\right)}} \tag{7}$$

Similar to the POT algorithm, P_{pti} is the probability of the p-th subject getting the i-th item in the t-th test correct, and $z\left(\Theta e_{pti}\right)$ is the standardized score of Θe_{pti}. All discrimination parameters (a parameters) are also set to 1 across the 100 items and 9 tests to maintain parsimony. The same sets of difficulty parameters of the domain-general and the domain-specific aspects (b parameters, bg and bs) are applied to the GSM algorithm, but unlike the POT algorithm, for each simulated response, the average of the corresponding bg and bs is taken as the b parameter for that item. According to this equation, the 3-dimensional array of Θe is transferred to a 1000 × 9 × 100 (number of subjects × number of tests × number of items per test) 3-dimensional array of probabilities. Similar to the POT algorithm, the 3-dimensional array of probabilities is applied to generate binary responses under the GSM algorithm.

6.4 Step 4: Task Scoring and Psychometric Model Fitting

For both algorithms, Step 3 generates a 1000 × 9 × 100 (number of subjects × number of tests × number of items per test) three-dimensional array of binary values (1 and 0), in which 1 represents a correct response and 0 represents an incorrect response. Thus, the simulated item-level responses are aggregated per individual, which results in a 1000 × 9 (number of subjects × number of tests) test-level dataset for each algorithm. In the two datasets, each subject has one score ranging from 0 to 100 for each of the nine tests, simulated under corresponding algorithms.

Two confirmatory factor models are applied to the two datasets to investigate theoretical latent factor structures that could best represent the data. First, a one-factor model is applied to both datasets, in which all 9 tests (as manifest variables) load onto the same *g* factor. Next, a three-factor model is applied, in which the 9 tests load onto three subfactors (fluid, verbal, and spatial) and the subfactors load onto a higher-order *g*. These two models are selected because the one-factor model is representative of Spearman's initial theory of general intelligence as well as Thomson's bonds model, whereas the higher-order three-factor model is representative of the current psychometric theories of intelligence, including the Cattell-Horn-Carroll model.

The fit of each model is evaluated with the following set of test statistics and fit indices: χ^2, Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), and Akaike Information Criterion (AIC). The criteria for an "acceptable" model fit were based on the following cut-off values, recommended by Schreiber et al. (2006) and Kline (2015): $p > .05$, $\chi^2/\text{df} \leq 3$, CFI $\geq .95$, RMSEA $\leq .06$, SRMR $\leq .08$. For AIC there is no cut-off value; lower values indicate better fit.

6.5 The Simulation Procedure

In the current study, the simulation process described above was conducted for 200 iterations. Factor loadings and fit indices from these 200 iterations were obtained and investigated. Parameters for these model parameters and indices were summarized and compared. All steps of the simulation were conducted in R (version 4.1.1). The confirmatory factor analyses were conducted using the package "lavaan" (Rosseel, 2012). R Scripts for the simulation were available at https://osf.io/jhzyp/.

The prediction of the results was based on the simulated cognitive algorithm of POT and GSM: for the POT algorithm, the one-factor model would convey poor model fit and the three-factor higher-order model would exhibit good fit. On the other hand, for the GSM algorithm, both the one-factor and the three-factor higher-order model would exhibit a good fit. Furthermore, under the POT algorithm, the higher-order model would indicate factor patterns that are consistent with real-world observations, such that the higher-order factor loading for the fluid factor will be stronger than the verbal and spatial factors. Under the GSM algorithm, factor loadings would be all balanced across tests and broad ability factors, due to the lack of distinction between domain-general and domain-specific processes.

6.6 Results, Study 1

For the simulated datasets based on either the POT algorithm or the GSM algorithm, the factor loadings, model chi-square statistics, and fit indices for the two confirmatory factor models were estimated for each of the 200 iterations. These parameters and indices were summarized and presented by the algorithm (POT or GSM) and the model structure (higher-order model or one-factor model). Model fit statistics (χ^2 statistics, CFI, RMSEA, and SRMR) and factor loadings of all 200 iterations in the current study were available as data files at: https://osf.io/jhzyp/.

Table 1 presented the descriptive statistics for the fit indices from the four sets of models. The results suggest that, under the POT algorithm, the higher-order model exhibited a good fit for the simulated data whereas the one-factor model exhibited a poor fit in general. For the higher-order model, χ^2 statistics were all acceptably small with an average of 26.40, ranging from 11.36 to 47.35. Whereas for the one-factor model, the average χ^2 statistics was 1637.46, ranging from 1392.91 to 1858.29. For reference, the critical value for a significant χ^2 (df = 24, for the higher-order model, α = .05) is 36.42, and the critical value for a significant χ^2 (df = 27, for the one-factor mode, $\alpha = .05$) is 40.11. Results of the fit indices also support that the higher-order model fits the simulated datasets under the POT algorithm better than the one-factor model.

Table 1. *Means (Standard Deviations) for the Fit Indices from the Four Sets of Models (by the Algorithm and the Model Structure)*

		χ^2	CFI	RMSEA	SRMR	AIC	BIC
POT	Higher-Order	26.21	1.00	.01	.02	58565.72	58668.78
		(6.57)	(< .01)	(.01)	(<.01)	(138.62)	(138.62)
	One-Factor	1637.46	.56	.24	.15	60170.97	60259.31
		(87.18)	(.02)	(.01)	(.01)	(169.36)	(169.36)
GSM	Higher-Order	26.77	.99	.01	.01	56411.03	56514.09
		(7.49)	(< .01)	(.01)	(<.01)	(127.53)	(127.53)
	One-Factor	27.69	1.00	.01	.01	56405.96	56494.30
		(7.65)	(< .01)	(.01)	(<.01)	(127.42)	(127.42)

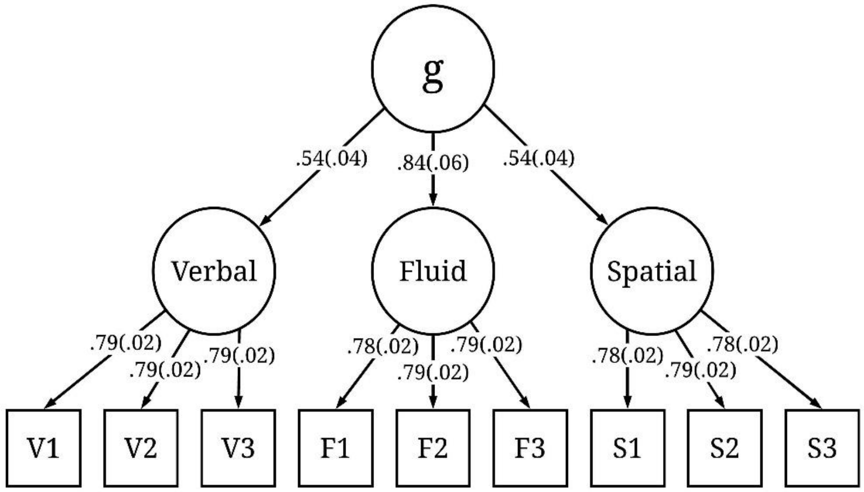
Note. χ^2 = Chi-Squared Statistics. CFI = comparative fit index. RMSEA = root mean square error of approximation. SRMR = standardized root mean square residual. AIC = Akaike information criterion. The degrees of freedom for the higher-order models were 24. The degrees of freedom for the one-factor models were 27.

On the other hand, under the GSM algorithm, both the higher-order model and the one-factor model had a great fit for the simulated data. As Table 1 presented, model fit statistics for the one-factor model indicated great fit: the average χ^2 statistics was 27.69, ranging from 9.80 to 50.39; all fit indices indicated almost perfect fit. For the higher-order model, the average χ^2 statistics was 26.77, ranging from 9.39 to 48.78; and the fit indices also indicated almost perfect fit. However, the results of the higher-order model also indicated a problem of overfitting. In some of the fitted models, the higher-order loadings were approximately equal to 1 due to the high intercorrelations among the subfactors. In fact, for most (163 out of 200) simulated datasets under the GSM algorithm, the model fitting results indicated that the higher-order model was an overfit to the simulated datasets. Given that the one-factor model have already fitted the data almost perfectly, it was clear that the higher-order structure was an over-specification of the GSM data. Thus, for the GSM algorithm, the one-factor model was the preferred model due to its more parsimonious nature. This was also supported by the comparison of the AIC and BIC indices from the two types of models, such that the one-factor models had slightly lower AICs and BICs.

Based on the results from these CFAs, Figure 8 presented means and standard deviations of the standardized factor loadings across the 200 iterations for the higher-order model of the POT algorithm. Figure 9 presented the same summary of the standardized factor loadings for the one-factor model of the GSM algorithm. The results were consistent with our prediction based on the two algorithms. For the POT algorithm, factor loadings for the manifest variables (representing the tasks) under their corresponding subfactors were balanced and stable across the 200 iterations. Furthermore, the higher-order factor loading for the fluid factor was stronger than the verbal and spatial factors. For the GSM algorithm, factor loadings for the 9 manifest variables were all balanced in the one-factor model and were stable across the 200 iterations.

Figure 8

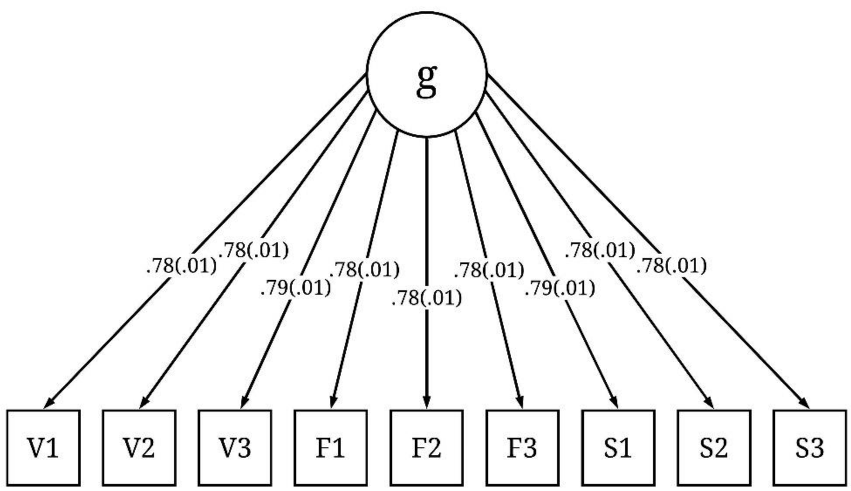
The Higher-Order Model of Intelligence Based on 200 Iterations of Simulated Test Scores Under POT Algorithm



Note. Values represent the means (and standard deviations) of standardized factor loadings.

Figure 9

The One-Factor Model of Intelligence Based on 200 Iterations of Simulated Test Scores Under GSM Algorithm



Note. Values represent the means (and standard deviations) of standardized factor loadings.

6.7 Discussion, Study 1

The simulation in Study 1 is a conceptual demonstration that the positive manifold of intelligence (the statistical/psychometric *g*) could emerge from the covariance of test scores in the absence of a general cognitive ability (a psychological *g*). It also provided supportive evidence that contemporary intelligence theories such as the process overlap theory could provide an alternative theoretical framework that unifies the psychometric theories and cognitive theories of intelligence.

Two simulation algorithms (POT and GSM), both based on cognitive theories, were designed and described in Study 1. The GSM reflected a generic sampling method that was based on the traditional bonds model (Thomson, 1916). The POT algorithm was based on the process overlap theory, which was inspired by the traditional bonds model but proposed several extra central claims of the cognitive mechanisms underlying the sampling of cognitive processes. Both algorithms were specifically designed to avoid any simulated cognitive processes being sampled across all items or tasks. The POT algorithm contrasted with the GSM algorithm such that the POT algorithm specified a distinction between general and specific cognitive processes and how they were sampled in different cognitive tasks based on cognitive theories.

Simulated test scores were generated from these two algorithms and were fitted onto two CFA models. Across multiple iterations, the test scores generated from these two algorithms consistently reflected a positive manifold among 9 simulated tasks (3 fluid tasks, 3 verbal tasks, and 3 spatial tasks) without a “general” cognitive process as the common source of variance. The CFA results of these simulated test scores were consistent with corresponding psychometric theories. As was predicted by traditional bonds models, a one-factor model (conceptually similar to Spearman’s original model) fitted the simulated data based on the GSM algorithm well, even in the absence of a common cognitive process across the tests. The POT algorithm took a step further by specifying different types of cognitive processes in the simulation algorithm. Based on the types of these hypothetical cognitive tests, the simulated cognitive processes were sampled in different manners per the corresponding cognitive properties of the tests. This specification as a simulated cognitive mechanism was reflected in the POT results. For the simulated data from the POT algorithm, a one-factor model was no longer appropriate and resulted in an unacceptable fit, whereas a higher-order three-factor model that was similar to the CHC model provided a great fit. Moreover, the POT algorithm also replicated a common psychometric finding that the fluid ability subfactor has a stronger loading on the higher-order *g* factor compared to other broad-ability subfactors in the higher-order model (Carroll, 1993).

Thus, the simulation results supported POT as a unified theory that attempts to bridge the gap between traditional cognitive theories and psychometric approaches. The POT algorithm proposed a conceptual illustration of the cognitive mechanisms in process overlap theory, which is an alternative interpretation of *g*, and the modeling results from this algorithm were consistent with findings in contemporary psychometric theories such as the CHC model. Although these results do not guarantee POT as the explanation of the inexistence of psychological *g*, they do support the idea that a psychometrical/statistical *g* could be derived from a systematic sampling mechanism as was proposed by POT without a common source of variance to different types of cognitive tests.

7 Study 2: A Psychometric Network of the POT Simulation

In Study 1, the simulated test scores based on the POT algorithm and GSM algorithm were examined by latent factor models that are similar to conventional psychometric models. The latent factor models of the simulated data were applied to illustrate that a common cause is not a necessary condition for a common factor model to fit the data. However, according to POT-S, the *g* factor is a formative construct and does not assume a psychological general intelligence. Therefore, POT proposes a network structure of intelligence (POT-N) to represent the positive manifold among cognitive tests. For latent factor models, the covariance among observed variables is assumed to be caused by a common source that is represented by a reflective latent variable, g. However, for network models, complex systems of the observed variables are represented as interconnected networks. The psychometric network models regard the manifest variables as nodes in interconnected networks and their pairwise interactive associations as edges (Epskamp & Fried, 2018). Conceptually, these partial correlation edges among nodes are similar to undirected alternatives to the directed regression paths in regression models, except that there is no distinction between a "predictor" or an "outcome". Thus, the partial correlations could be used to describe massive multicollinear relationships among cognitive testing data and transform the positive manifold into a graphic network. Hence, the psychometric network models are more compatible with process overlap theory than are traditional latent factor models. Study 2 applies a psychometric network analysis on simulated data based on the POT algorithm from Study 1 to provide an alternative representation (POT-N) of the intercorrelations among the simulated test scores.

Furthermore, another goal of Study 2 is to extend the standard psychometric network analysis approach by combining latent factor models and network models. Specifically, the simulated data are also investigated using a latent network model (Epskamp et al., 2017). Conceptually, the current POT-S is a hybrid model in which the higher-order *g* is regarded as a formative construct, but the broad ability factors are still regarded as being reflective. Compared to POT-S, POT-N is represented in a standard psychometric network, which rejects not only the notion of general intelligence but also the notion of all broad abilities, by describing the common variance among all tests as an interconnected network. A better POT-N should exhibit psychometric networks at the broad ability level, but retain the latent factor structures at the lower level of specific task measures. In a latent network model, confirmatory hypotheses regarding different types of tasks are specified in a similar way to the hypotheses in a measurement model, and relationships among these latent factors are estimated in a similar way to those in a standard psychometric network model. Thus, the nodes in latent network analysis are latent factors based on prior theories and account for the measurement error in specific tasks, while the edges are estimated conditional dependence relationships among these latent factors. In this study, we refer to these nodes as "latent nodes" to reflect their conceptual similarity with the latent factors in factor models. A latent network model of POT therefore bridges POT-S and POT-N by accounting for measurement error in cognitive tasks while assuming no psychological *g*.

7.1 Method, Study 2

The simulation process for the current study replicated the process in Study 1 but with a few updates. First, in the current study, only the POT algorithm was included in the simulation. Second, due to model convergence issues, a lower discrimination parameter of a = .70 was introduced to this simulation. Third, an iteration of the simulation process was not conducted. The simulation in the current study resulted in one 1000 × 9 matrix simulating test-level scores of 9 tests for 1000 subjects.

The psychometric network analyses were conducted with the "psychonetrics" package (v 0.9; Epskamp, 2021) in R. R Scripts for the simulation and the network analyses are available at https://osf.io/syxed/. Network models were estimated by modeling the variance-covariance matrix of the data as Gaussian graphical models (Epskamp et al., 2017). For the standard network model with simulated scores of the 9 tests as nodes, a model optimization procedure was conducted, in which the initial model was pruned by a step-down search process with a significance level of .01 in a recursive manner, such that edges that were not significant at α

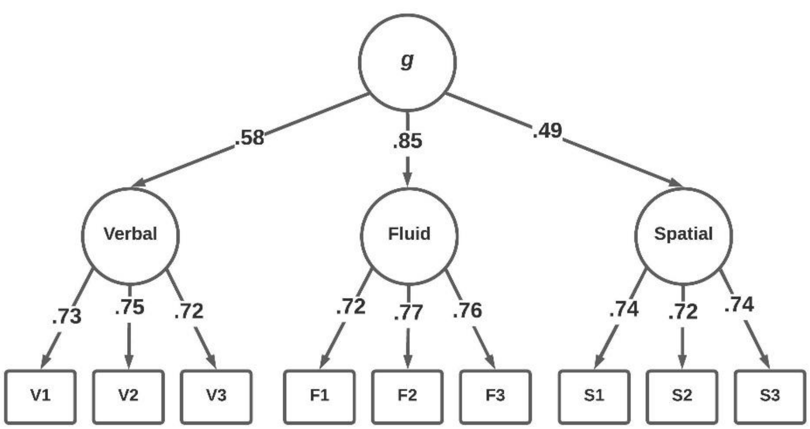
= .01 were automatically and recursively removed. The pruned model was then optimized by a step-up search process with a significance level of .01, such that the edges that were removed in the previous steps were added back, based on modification indices, until BIC no longer increased. For the latent network model, a total of 3 latent nodes were recognized based on the simulation specifications: Fluid, Verbal, and Spatial. Each latent node consisted of three corresponding tasks. Due to the small set of latent nodes, no model optimization process for the latent network was conducted. The latent network model was presented in its initial form of a graphical model. For the current study, the higher-order factor model was also estimated for the simulated data for conceptual comparison of latent factor model, network model, and latent network model.

7.2 Results, Study 2

The higher-order latent factor model exhibited excellent fit to the simulated data: $\chi^2(24)$ = 20.24, p = .683; CFI = 1.00, TLI = 1.00; RMSEA < .01, SRMR = .02; AIC = 55450.35, BIC = 55553.41 (see Figure 10). The results of the simulation replicated the simulation in Study 1 (see Figure 8).

Figure 10

The Higher-Order Latent Factor Model of the Simulated Data Based on POT



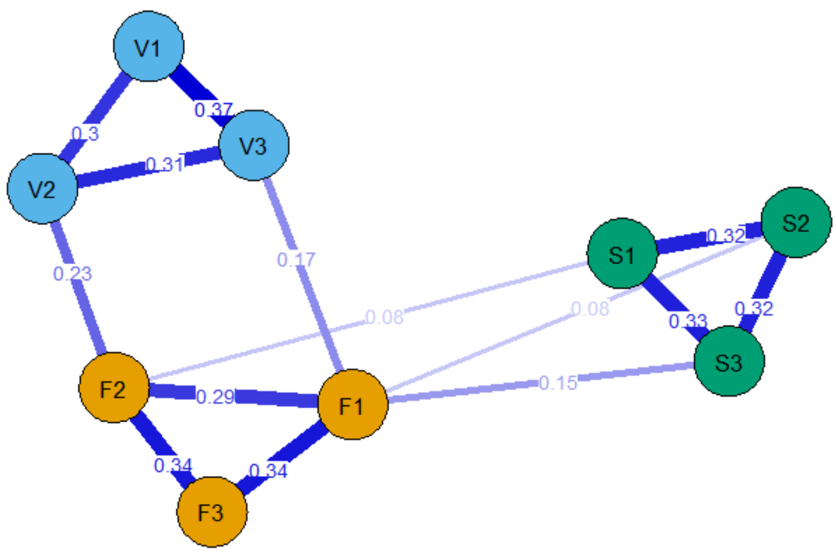
Note. F1, F2, and F3 were fluid tasks; V1, V2, and V3 were verbal tasks; S1, S2, and S3 were spatial tasks. All factor loadings were statistically significant.

As mentioned, two types of psychometric network models were constructed to explore the specific associations among the tasks. First, a standard network model analysis was applied to the full simulated dataset with nine manifest variables as nodes for the network. Second, a latent network model analysis was applied to the same dataset but with three latent variables (Fluid, Verbal, and Spatial) specified as nodes for the network, whereas the nine manifest variables loaded onto the three latent variables correspondingly. As described in the method section, the optimization procedure was conducted for the standard network analysis. The network model was pruned by a step-down search process and then optimized by a step-up search process. Due to the small set of latent nodes, no model optimization process for the latent network was conducted. The current latent network model retained and presented all three edges among the three latent nodes regardless of statistical significance. Both final models exhibited acceptable fit according to the fit statistics (with the exception of significant χ^2 statistics): for the standard network model, $\chi^2(22) = 98.93$, $p < .001$, CFI = .97, TLI = .95, RMSEA = .06, AIC = 55804.93, BIC = 55961.98; for the latent network model, $\chi^2(24) = 179.19$, $p < .001$, CFI = .96, TLI = .93, RMSEA = .08, AIC = 22952.53, BIC = 23099.76.

The weighted, undirected optimized standard network model of the nine tests simulated from the POT algorithm is presented in Figure 11. Nodes in this model, representing the manifest variables, were colored to reflect the theoretical clusters based on types of the corresponding tasks (Fluid, Verbal, and Spatial). Edges were weighted so that the width represents the magnitude of conditional dependent association between each pair of nodes. Wider edges indicate higher partial correlations between the two nodes. The estimated edges, representing estimated conditional dependencies between pairs of the nodes, are presented numerically, where larger values indicate stronger partial correlations. Compared to the higher-order latent factor model, this standard network model described the same correlations among the nine simulated variables without assuming any latent factor. However, the three broad abilities were still represented by the three small clusters of nodes. Furthermore, compared to the verbal and spatial tests, the fluid tests played a more central role in the network model: the fluid nodes connected to both spatial and verbal nodes, while edges between the spatial and verbal nodes were not significant and therefore excluded. This pattern was a network representation of the larger factor loading of the fluid factor to the higher-order *g* in the latent factor model (Figure 10).

Figure 11

The Standard Network Model of the Simulated Data Based on POT

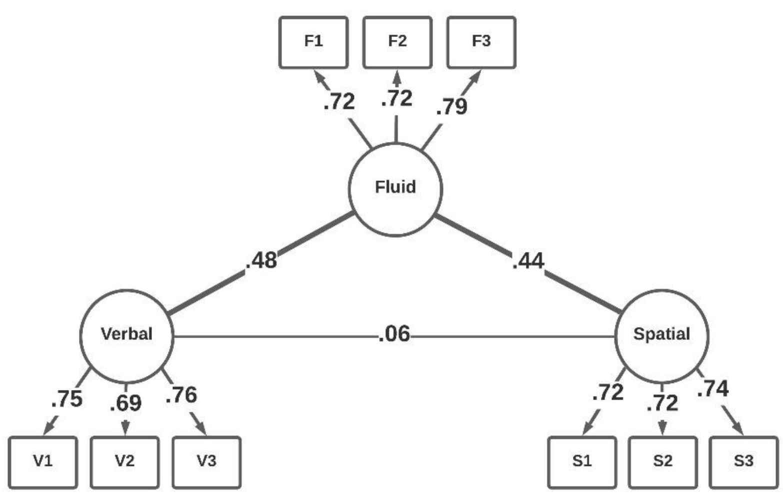


Note. The network was a weighted, undirected Gaussian graphical model that was optimized using step-down search and step-up search. F1, F2, and F3 were fluid tasks; V1, V2, and V3 were verbal tasks; S1, S2, and S3 were spatial tasks.

The latent network model is presented in Figure 12. Unlike the former model, nodes in this model represented latent variables estimated from the nine manifest variables (presented as boxes in the figure). Similar to the former model, the conditional dependent associations (partial correlations) between these latent nodes were presented as weighted edges. Wider edges indicated higher partial correlations between the two nodes. Furthermore, factor loadings of the observed variables were presented as directed paths (arrows from latent nodes to manifest variables). The estimated edges and factor loadings were all presented in numbers. In this latent network, the edges between Verbal and Fluid factors and Spatial and Fluid factors were statistically significant, both with $p < .001$. The edge between Verbal and Spatial was not statistically significant, $p = .14$. The latent network model, again, described the same correlations among the nine variables and combines features of the latent factor model and the network model. In the current latent network model, the broad ability nodes were positively associated with each other without attributing to a latent common cause. Furthermore, compared to either the Verbal or the Spatial node, the Fluid node was more central to the latent network.

Figure 12

The Latent Network Model of the Simulated Data



Note. The latent network was a weighted undirected Gaussian graphical model without optimization. F1, F2, and F3 were fluid tasks; V1, V2, and V3 were verbal tasks; S1, S2, and S3 were spatial tasks. All factor loadings were statistically significant. The edge between Verbal and Spatial factors was not statistically significant ($p = .14$).

7.3 Discussion, Study 2

The psychometric network models presented in Study 2 offered an alternative perspective to investigate the correlational structure of individual differences in cognitive abilities. The network models estimated conditional dependent associations (partial correlations) among test scores, in which the underlying interconnectedness of observed data could be directly represented without assuming unobserved latent factors (Schmank et al., 2019). In the network models of intelligence, no common cause of covariance was emphasized to represent the positive manifold of cognitive ability, and thus, the network approach was more compatible with process overlap theory (POT-N). According to POT, the correlations among cognitive testing data are caused by the overlaps among the sampling of cognitive processes involved in different tests.

As was presented in the current standard network model, these mutual associations due to process overlap could be directly estimated in the network model by the edges between nodes, with stronger edges indicating more overlaps of sampled cognitive processes between two tests. Thus, the shared variance among a set of variables was represented as small clusters of nodes with strong edges in the network model (Figure 11) instead of latent factors and factor loadings in the CFA model (Figure 10). In the POT algorithm of this simulation, domain-general processes (defined as the executive function processes) were sampled across all 9 tests, and the three types of domain-specific processes (fluid, verbal, and spatial processes) were sampled only by their corresponding types of tests. This sampling algorithm was represented by the current standard network model, as the 9 nodes were generally connected as a graphic network among the 9 nodes, and three broad abilities were represented by three clusters of corresponding nodes.

Another important feature observed in the network model was the central role of fluid ability. In this network model, after optimization based on the statistical significance of edges and the goodness of model fit, there were edges retained between fluid and verbal nodes as well as fluid and spatial nodes, but no edges were retained between verbal and spatial nodes. This central role of fluid tests in the current network was also compatible with the common finding in psychometric models that the tests that are designed to directly measure fluid cognitive functions tend to have higher loadings on g than other types of tests (Gustafsson, 1988; Blair, 2006). In this study, the feature in the network model directly illustrated one statistical specification of the sampling mechanism in the POT algorithm, such that the EF processes were sampled with a higher probability ($P = 0.28$) in fluid tests than in verbal and spatial tests ($P = 0.12$). Therefore, this specification represented more sampling overlaps between the cognitive processes sampled by fluid and verbal tests as well as fluid and spatial tests, compared to those between verbal and spatial tests.

A standard psychometric network that estimates the partial correlations between pairs of nodes is primarily exploratory and data-driven. However, in the real world, depending on the tests, the local independent associations estimated from a network model could be biased by specific measurement errors. Therefore, the latent network model of intelligence went a step further and combined the benefits of both latent factor models and psychometric network models. That is, the latent nodes of broad abilities (Fluid, Verbal, and Spatial in Figure 12) retained the benefits of latent factor models and account for measurement error in cognitive tasks; and the network structure among these latent nodes rejects the notion of general intelligence as a common cause of the positive manifold and reflects the connectedness of the broad ability factors using partial correlations. In the current latent network model, the three latent nodes were conceptually and statistically comparable to the latent factors in the latent factor model, such that the nine simulated tests were loaded onto their corresponding latent nodes that represented the three broad abilities. Thus, the three latent nodes are statistically error-free estimates of the three broad abilities. Unlike the higher-order *g* in the latent factor model, a partial correlation network was estimated based on the intercorrelations among the three latent nodes. Similar to the edges in a standard network model, the partial correlation edges among latent nodes were similar to undirected alternatives to the directed regression paths among latent factors in structural equation models. The results of the latent network model also replicated the central role of fluid ability in the standard network model. In the model presented in Figure 12, the edge between fluid and verbal latent nodes was as strong as that between fluid and spatial latent nodes, but the edge between verbal and spatial latent nodes was not significant. This was a more direct network representation of the higher g loading of fluid intelligence that was commonly observed in previous psychometric studies. Again, in the current study, the centrality of fluid ability in the network was obtained by the sampling specification of domain-general processes based on POT, which led to a higher overlap of sampled cognitive processes between fluid tests and other types of tests.

8 Study 3: The Domain-Generality of Working Memory and Fluid Intelligence: Rethinking the Relationship via Network Analyses of Kane et al. (2004)

Study 2 applied psychometric network analysis on a simulated dataset based on POT and offered an alternative perspective to investigate the correlational structure of individual differences in cognitive abilities. Furthermore, it applied a latent network approach that combines the benefits of both latent factor models and psychometric networks. In study 3, we apply this network approach to an empirical dataset with multiple measures of cognitive abilities, including working memory capacity (WMC) and fluid intelligence. Study 3, therefore, attempts to “revisit” the relationship between WMC and fluid intelligence using the exploratory network approach.

Individual differences research on working memory has largely focused on understanding the relationship between WMC and fluid intelligence by estimating and investigating the amount of shared variance between the two constructs using latent factor models. For example, Kane et al. (2004) investigated the relationship between short-term memory (STM), working memory and general fluid intelligence using structural equation models (SEM) on 25 tasks including verbal and spatial simple span tasks, verbal and spatial complex span tasks, and three types of reasoning tasks (fluid, verbal, and spatial). The confirmatory factor models in the study tested whether the shared variance between simple and complex span, verbal and visuospatial memory span tasks could be explained by a single domain-general ability factor. The study also contrasted the predictive validity of complex span tasks to that of simple span tasks on general fluid intelligence.

Kane et al. (2004) found that the verbal and spatial complex span tasks were more strongly correlated with each other (sharing 70–85% of variance) than were the verbal and spatial simple span tasks (sharing 40% of variance), suggesting that complex span tasks drew more strongly from a domain-general ability. Also, WMC measured by the complex span tasks and short-term memory (STM) measured by the simple span tasks are closely related but separable in latent factor analyses. Additionally, Kane et al. (2004) found that, using a bi-factor approach, a general factor of executive attention from all memory span tasks was more predictive of general fluid intelligence than were the domain-specific factors (Kane et al., 2004). These are important findings because they suggest that the shared variance between memory span tasks and fluid intelligence occurs largely due to a domain-general ability (e.g., executive attention).

It is important to note that Kane et al. (2004) approached their analysis using confirmatory, theory-driven latent modeling techniques. However, while the confirmatory latent factor models are adequate to test and estimate relationships among observed and unobserved variables based on specific theories, they usually rely on both conceptual and statistical assumptions, and could miss theory-inconsistent information. In Study 3, we attempt to show that the network modeling approach, which relies on fewer assumptions and is more exploratory, can provide important new insights regarding the nature of cognitive abilities underlying the relationship between working memory and fluid intelligence. This technique complements a factor analysis approach by focusing on the patterns of pairwise conditional dependencies (edges) among observed variables (nodes) rather than the dimensional reduction of multivariate data. Therefore, network modeling is ideal for modeling the correlational structure of cognitive measures that consist of overlapping processes. For example, the verbal complex span tasks overlapped with other verbal tasks on domain-specific (verbal) processes such as the verbal simple span and the verbal reasoning tasks, and they also overlapped with other span tasks on domain-general (executive attention) processes. This complex nature of overlapping processes is difficult to be fully accounted for in a confirmatory latent factor model but can be represented well in a network model.

8.1 Method, Study 3

Study 3 was a re-analysis of data from Kane et al. (2004). In the original study, a total of 25 cognitive tasks was administered to 236 subjects. The task battery included 12 memory span tasks and 13 reasoning tasks. The 12 memory span tasks consisted of 6 simple span tasks and 6 complex span tasks, half of which used verbal materials and half visuospatial materials. The 13 reasoning tasks consisted of 5 verbal reasoning tasks, 5 spatial reasoning tasks, and 3 matrix reasoning tasks. For a summary of the 25 tasks, see Table 1. For further details about the administration of the tasks, see Kane et al. (2004).

Table 2. *Description of Cognitive Tasks from Kane et al. (2004)*

Latent Construct	Cognitive Task
Verbal Short-Term Memory	*Word span*: Participants recalled sequences of one and two syllable nouns that were presented for 1s each with a 500 ms ISI. *Letter Span*: Participants recalled sequences of letters. *Digit Span*: Participants recalled sequences of digits.
Spatial Short-Term Memory	*Arrow Span*: participants recalled sequences of long/short arrows in eight possible directions. *Matrix Span*: Participants recalled sequences of the positions of red squares in a 4x4 matrix. *Ball Span*: Participants recalled sequences of the movement of balls moving across a screen.
Verbal Working Memory	*Operation span*: Participants recalled sequences of words while also verifying the answers to math problems on a screen. *Reading span*: Participants recalled sequences of letters on a screen while also judging whether sentences were sensical or not. *Counting span*: Participants recalled sequences of digits while also counting the number of colored circles on a screen.
Spatial Working Memory	*Rotation span*: Participants recalled sequences of long/short arrows on the screen while also deciding whether rotated letters on the screen were mirror-reversed or not. *Symmetry span*: Participants recalled the positions of red squares on a 4x4 matrix while also judging whether an abstract figure was symmetrical or not across its vertical axis. *Navigation span*: Participants recalled the movement of balls across a screen while also mentally navigating a moving target's position in a maze.
Verbal Reasoning	*ETS inference test*: Participants read passages and selected from an answer set of sentences which sentence could be correctly inferred from the information given. *AFOQT Analogies test*: Participants read incomplete analogies and selected the best word or phrase that completed the analogy.

	AFOQT Reading Comprehension: Participants read 2-6 sentence passages and completed the final sentence of each passage from one of five choices. *Remote Associates Test*: Participants generated a word that was semantically related to a set of three distantly related words. *ETS nonsense syllables*: Participants read two nonsensical premises that were followed by a conclusion and determined whether the conclusion could be correctly derived.
Spatial Reasoning	*DAT Space Relations Test*: Participants were presented with the illustration of a piece of paper that could be folded along the edges to make a three-dimensional shape. Participants selected the correct folded three-dimensional shaped paper from five answer choices. *AFOQT Rotated Blocks Test*: Participants were presented with an irregular-shaped block and selected the correct rotated block from five answer choices that best matched the target block. *ETS Surface Development Test*: Participants were shown pairs of unfolded paper with edges and corresponding folded shapes. The edges for both were numbered/lettered. Participants matched the edges between the unfolded paper and the folded shape. *ETS Form Board Test*: Participants were shown geometric shapes that could potentially be put together in some combination to form a geometrical figure. Participants selected the correct pieces to form the figure. *ETS Paper Folding Test*: Participants were shown a square piece of paper being folded and then having one or two holes being punched through it. Participants selected from five answer choices, which choice would represent the paper if unfolded.
Fluid (Matrix) Reasoning	*RAPM, Set II*: Each item presented a pattern of eight black and white figures arranged in a 3x3 matrix with one figure missing. Participants fill in the most appropriate figure from a set of 5 options. *WASI Matrix Reasoning*: Participants were presented with a pattern of colored figures and were arranged in a matrix. There was a missing piece in the matrix. Participants selected the best figure from an answer set that best completed the matrix. *BETA III Matrix Reasoning*: Participants were presented with three novel, black and white figures arranged in a 2x2 matrix. Participants selected the one figure from a set of five possible choices that would best complete the matrix.

In Study 3, similar to Study 2, psychometric network analyses and latent network analyses were conducted on the data. The network analyses were first conducted on the 12 memory span tasks. Conditional dependencies among the span tasks were investigated using both the standard task-level network analyses and the latent network analyses. For the latent network analyses of memory span tasks, a total of 4 "latent nodes" were recognized among the 12 tasks: Spatial Complex Span (Ws), Verbal Complex Span (Wv), Spatial Simple Span (Ss), and Verbal Simple Span (Sv). The network analyses were then conducted on the full dataset with 12 memory span tasks and 13 reasoning tasks. Conditional dependencies among the 25 manifest variables were investigated using both the regular task-level network analyses and the latent network analyses. Three "latent nodes" were recognized among the 13 reasoning tasks: Matrix Reasoning (Rm), Spatial Reasoning (Rs), and Verbal Reasoning (Rv). These latent factors/nodes were specified because we were interested in exploring the network structure among the complex cognitive tests while retaining the benefits of latent factor models to account for measurement errors of tasks. Thus, we specified the latent factors based on the types of tasks instead of the theoretical cognitive constructs. All psychometric network analyses in the current study were conducted with the "psychonetrics" package (v 0.9; Epskamp, 2021) in R using the step-down and step-up model optimization procedures described in Study 2. Data and R Scripts were available at https://osf.io/d4e98/.

8.2 Results, Study 3

Descriptive statistics for the memory tasks and reasoning tests are presented in Table 2. For all simple and complex span tasks, task scores were mean composites of item scores calculated using the partial-credit unit scoring procedure (Conway et al., 2005). For all reasoning tasks, task scores were mean composites of binary item scores, except for the ETS Surface Development Test, where item scores were the proportions of correct responses in each set.

Table 3. *Descriptive Statistics for the Memory Span Tasks and Reasoning Tests (N = 235)*

Task	Label	*M*	*SD*	*Skewness*	*Kurtosis*
Word Span	wor	0.78	0.11	-0.46	0.22
Letter Span	let	0.78	0.10	-0.29	-0.31
Digit Span	dig	0.80	0.10	-0.39	0.26
Ball Span	bal	0.58	0.12	-0.12	0.82
Array Span	arr	0.65	0.16	-0.76	1.26
Matrix Span	mat	0.51	0.15	-0.14	-0.15
Operation Span	ope	0.65	0.17	-0.47	0.05
Counting Span	cou	0.69	0.15	-0.26	-0.39
Reading Span	rea	0.67	0.16	-0.32	-0.13
Navigation Span	nav	0.48	0.18	0.10	-0.44
Symmetry Span	sym	0.47	0.20	0.26	-0.58
Rotation Span	rot	0.61	0.18	-0.29	-0.45
ETS Inference	etsin	0.61	0.22	-0.35	-0.58
ETS Paper Folding	papfo	0.60	0.24	-0.25	-0.63
AFOQT Analogies	afqan	0.56	0.22	-0.12	-0.54
DAT Space Relations	datsr	0.57	0.21	-0.23	-0.52
Remote Associates	remass	0.46	0.17	-0.21	-0.18
AFOQT Rotated Blocks	afqrb	0.37	0.23	0.48	-0.37
ETS Nonsense Syllogisms	etssy	0.54	0.17	0.26	-0.22
ETS Surface Development	etssd	0.60	0.25	-0.10	-1.02
AFOQT Reading Comprehension	afqrc	0.51	0.26	0.16	-1.03
ETS Form Board	etsfb	0.39	0.24	0.37	-0.72
WASI	wasim	0.69	0.14	-0.53	0.75
RAPM	raven	0.55	0.15	-0.58	0.66
BETA III	beta3	0.86	0.13	-1.80	5.05

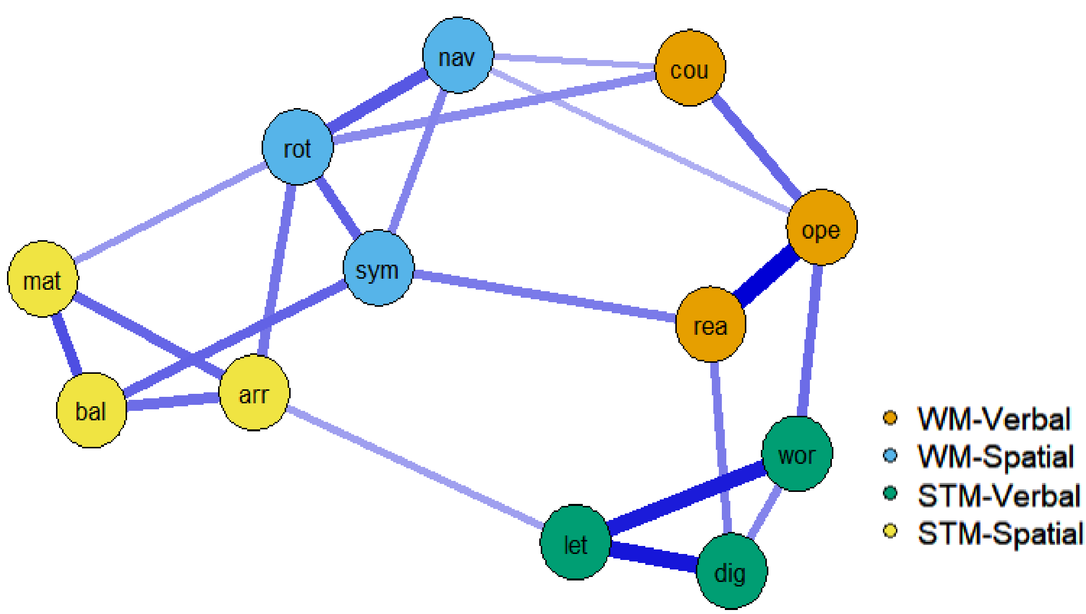
8.2.1 Network Models of Memory Span Tasks

The standard network model analysis was applied to the memory span data with 12 manifest variables as nodes for the network: 3 verbal simple span tasks (Word, Letter, and Digit), 3 spatial simple span tasks (Arrow, Matrix, and Ball), 3 verbal complex span tasks (Operation, Reading, and Counting), and 3 spatial complex span tasks (Rotation, Symmetry, and Navigation). Model optimization procedure (step-down and step-up selection) was conducted on the network. The optimized model exhibited acceptable fit according to the fit statistics (except significant χ^2 statistics). For the standard network model after optimization, $\chi^2(45) = 64.58$, p

=.029; CFI = .99, TLI = .98; RMSEA = .04; AIC = -4763.46, BIC = -4607.77. Figure 13 presents the weighted, undirected optimized regular network model of the 12 memory span tasks in Kane et al. (2004). Edges were weighted so that the width represents the magnitude of conditional dependent association between each pair of nodes. Overall, the network model illustrated a positive manifold of memory span tasks in a network manner without relying on any common cause of the variance. The small clusters were conceptually reflecting the grouping of the tasks based on their corresponding types (simple or complex, verbal or spatial). Despite the strong edges (partial correlations) that constituted these clusters by task type, there were also strong edges between verbal and spatial complex span tasks. Also, the connectedness was strong between verbal complex and simple span tasks, and so was that between spatial complex and simple span tasks. However, there were fewer directly connected nodes between verbal and spatial simple span tasks.

Figure 13

The Standard Network Model of the Memory Span Data in Kane et al. (2004)

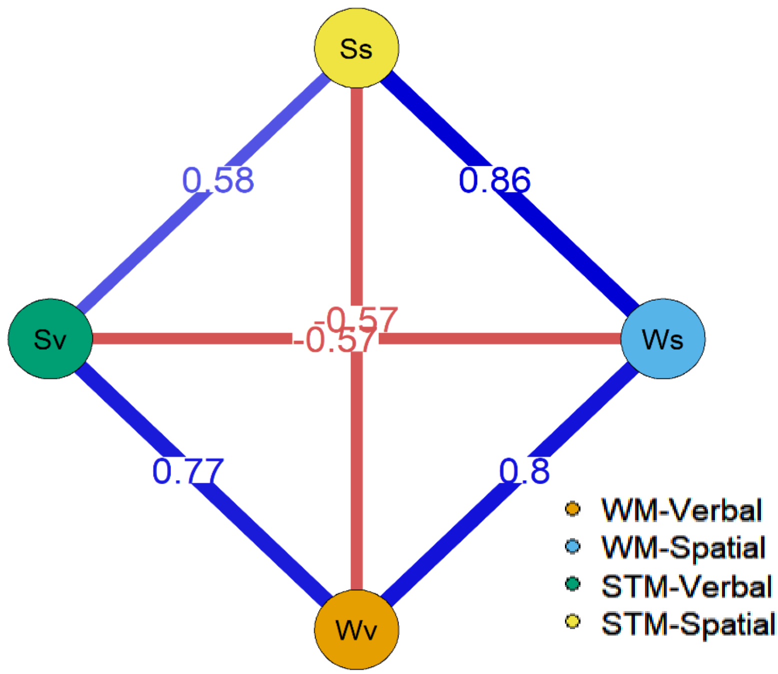


Note. Nodes in this model represent the manifest variables in the data and are colored to reflect the theoretical cluster based on the types of corresponding tasks. Edges in the model are weighted so that the width represents the magnitude of conditional dependent association between each pair of nodes. Blue edges indicate positive associations.

The latent network model analysis was applied to the same data with 12 tasks but with 4 latent variables: Verbal Short-term Memory (Sv), Spatial Short-term Memory (Ss), Verbal Working Memory (Wv), and Spatial Working Memory (Ws). The four latent variables were specified as nodes for the network, whereas the 12 manifest variables loaded onto the 4 latent variables correspondingly. Model optimization procedure (step-down and step-up selection) was conducted in the same manner but only to the latent network. For the latent network model without optimization, $\chi^2(48) = 77.88$, $p = .004$; CFI = .98, TLI = .97; RMSEA = .05; AIC = -4744.13, BIC = -4598.83. Figure 14 presented the weighted, undirected latent network model of the same data at the latent factor level. All standardized factor loadings of the 12 manifest variables were significant, and thus no manifest variables were included for the simplicity of the visualization. Unlike the former model, nodes in this model represent latent variables. Edges between the nodes represented the conditional dependent associations (partial correlations) between the latent variables. In the latent network model of memory span data, the partial correlation between the two latent nodes of working memory measures (Ws and Wv) was strong and positive. Furthermore, the partial correlation between the spatial latent nodes (Ws and Ss) and that between the verbal latent nodes (Wv and Sv) were also strong. Compared to the previous two edges, the partial correlation between the two simple span nodes (Ss and Sv) was relatively smaller. There were also two negative edges in the current latent network model between Ss and Wv, and Sv and Ws. These negative edges were not predicted but were interpreted in the discussion.

Figure 14

The Latent Network Model (Only Latent Nodes) of the Memory Span Data in Kane et al. (2004)



Note. Nodes in this model represent latent variables estimated from the observed variables. The latent nodes are colored in the same colors as the observed variables in Figure 13. Edges in the model are weighted so that the width represents the magnitude of conditional dependent association between each pair of nodes. Blue edges indicate positive associations and red edges indicate negative associations. The estimated partial correlations were also presented in numbers.

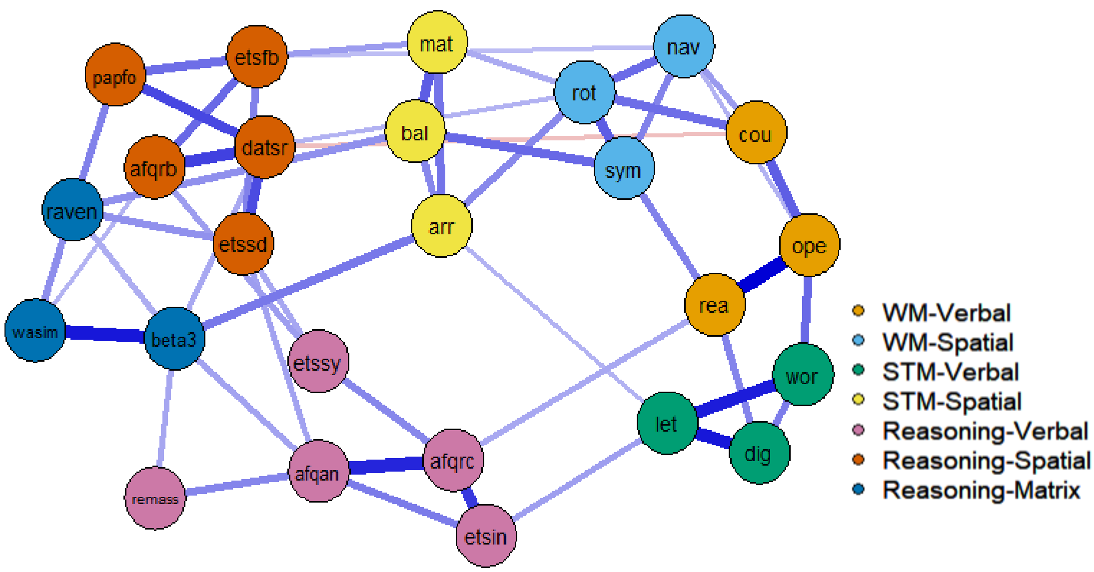
8.2.2 The Networks of Memory Span Tasks and Reasoning Tests

The same network analyses were conducted on all the 25 cognitive tasks administered in the original Kane et al. (2004) study. The regular network model analysis was applied to the memory span data with 25 manifest variables as nodes for the network. Other than the 12 memory span tasks, the 13 reasoning tests were also included: 3 matrix reasoning tests, 5 verbal reasoning tests, and 5 spatial reasoning tests. For the latent network model, a total of seven latent factors were specified: the verbal simple span factor with all 3 verbal simple span tasks (Sv), the spatial simple span factor with all 3 spatial simple span tasks (Ss), the verbal working memory factor with 3 verbal complex span tasks (Wv), the spatial working memory factor with 3 spatial complex span tasks (Ws), the verbal reasoning factor with 5 verbal reasoning tests (Rv), the spatial reasoning factor with 5 spatial reasoning tests (Rs), and the fluid reasoning factor with 3 matrix reasoning tests (Rf). Again, the latent factors were specified based on the types of cognitive tasks. This was due to the interest in exploring the network structure among these complex cognitive tests while retaining the benefits of latent factor models to account for measurement errors of tasks. The same model optimization procedure as was specified in previous methods was conducted on both the regular network and the latent network. Figure 15 presented the weighted, undirected optimized network model of memory span tasks and reasoning tests in Kane et al. (2004). Figure 16 presents the weighted, undirected optimized latent network model of the same data at the latent factor level. Both models exhibited good fit according to the fit statistics. For the regular network model, $\chi^2(df)$ = 286.14 (248), p = .048; CFI = .99, TLI = .99; RMSEA = .03; AIC = -7855.23, BIC = -7502.36. For the latent network model, $\chi^2(df)$ = 363.94 (265), p < .001; CFI = .97, TLI = .97; RMSEA = .04; AIC = -7786.39, BIC = -7492.32.

The standard network model illustrated a positive manifold of all cognitive tasks and small clusters of nodes by their task types. Furthermore, the nodes that were considered to reflect more domain-general individual differences, namely those of the complex span tasks and the matrix reasoning tests, played relatively central roles within memory span tasks and reasoning tests, respectively. However, these two types of nodes were not directly connected in the standard network model. The latent network of memory span tasks and reasoning tests revealed similar patterns but relied on confirmatory assumptions of latent broad abilities by task types. All latent nodes were generally positively connected. The reasoning nodes (Rm, Rv, and Rs) are only connected to the memory nodes through the STM nodes (Sv and Ss) but not the WM nodes (Wv and Ws). Also, the latent network indicated that, after controlling for all other information, spatial and verbal reasoning (Rs and Rv) were positively associated, as were the complex span nodes (Ws and Wv), but the simple span nodes were not associated.

Figure 15

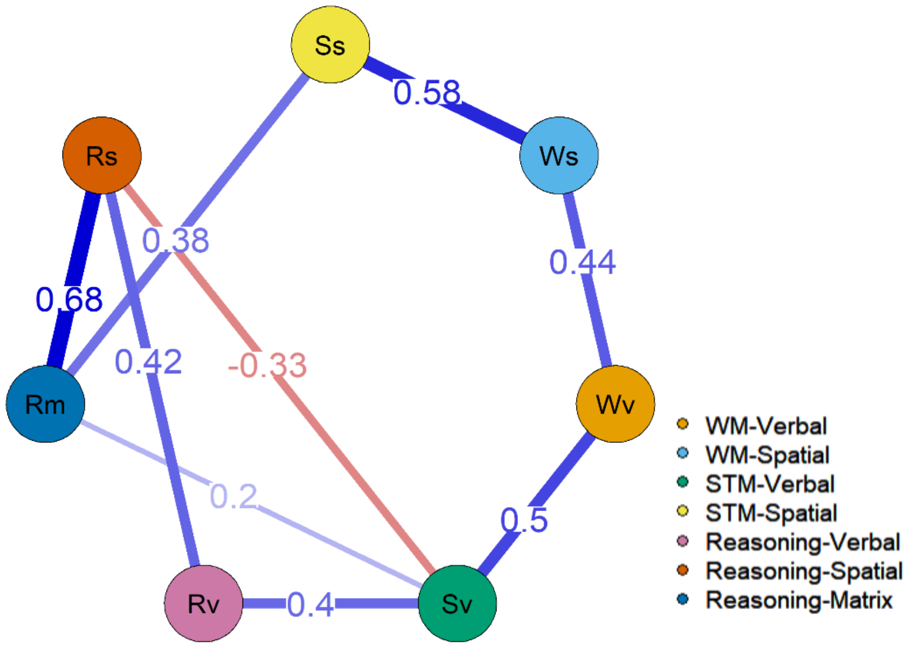
The Standard Network Model of Memory Span Tasks and Reasoning Tests in Kane et al. (2004)



Note. Nodes in this model represent the manifest variables in the data and are colored to reflect the theoretical cluster based on the types of corresponding tasks. Edges in the model are weighted so that the width represents the magnitude of conditional dependent association between each pair of nodes. Blue edges indicate positive associations and red edges indicate negative associations.

Figure 16

The Latent Network Model of Memory Span Tasks and Reasoning Tests in Kane et al. (2004)



Note. Nodes in this model represent latent variables estimated from the observed variables. The latent nodes are colored in the same colors as the observed variables in Figure 15. Edges in the model are weighted so that the width represents the magnitude of conditional dependent association between each pair of nodes. Blue edges indicate positive associations and red edges indicate negative associations. The estimated partial correlations were also presented in numbers.

8.3 Discussion, Study 3

Study 3 applied the standard and latent network analysis introduced in Study 2 to an empirical dataset with multiple measures of human cognitive abilities. Specifically, this study reinvestigated the correlational structure of test scores on 12 memory span tasks and 13 reasoning tests from Kane et al. (2004) using psychometric network analyses. Compared to latent factor models, the psychometric network analyses relied on fewer common-cause premises to illustrate the intercorrelations (positive manifold) among the working memory and reasoning measures in a more exploratory manner using partial correlation networks.

The standard network analysis of WM indicated several important features of the 12 memory span tasks. First, the positive manifold of WM was reflected in the standard network model. The 12 nodes were positively connected by weighted edges in the network, reflecting conditional dependent associations between the span task scores. Second, the small clusters in the network generally reflected the corresponding types of tasks and were comparable to a correlated multi-factor structure in latent factor modeling: verbal WM, spatial WM, verbal STM, and spatial STM. Third, compared to the edges between nodes for verbal complex span and spatial complex span tasks, there were fewer edges retained among verbal and spatial simple span tasks. These patterns were compatible with existing cognitive theories such that the domain-general cognitive mechanisms are essential to the active maintenance and rapid retrieval of information in working memory tasks (Conway & Kovacs, 2020). There were also strong connections between spatial tasks and verbal tasks in the network, indicating the influence of domain specificity (e.g., verbal/spatial storage) on individual differences in WM.

Similar to Study 2, we also extended the standard network analysis and applied latent network analysis to the memory span data in the current study. An important pattern in the latent network of WM was that the edge between spatial and verbal WM nodes (Ws and Wv) was stronger than the edge between spatial and verbal STM modes (Ss and Sv). This replicated the original finding in Kane et al. (2004) that the verbal and spatial complex span tasks were more strongly correlated with each other than were the verbal and spatial simple span tasks, suggesting that complex span tasks drew more strongly from domain-general cognitive processes than simple span tasks.

Another interesting result in the latent network of WM was the negative edge between Sv and Ws, and that between Ss and Wv. These negative edges were unexpected but were interpretable. They were likely due to the “common effects” among the four types of tasks (Pearl, 2000; Epskamp & Fried, 2018). Specifically, they reflected the common impacts of domain-general abilities (executive functions) and domain-specific abilities (e.g., verbal or spatial short-term storage) on the performance of different types of memory span tasks.

For example, there was a negative edge between verbal STM tasks (Sv) and spatial WM tasks (Ws) in the current latent network model. This negative partial correlation represented that, after controlling for the influence of all other variables in the network (specifically, Ss and Wv), the model estimated a negative correlation between Sv and Ws. According to the common effect explanation (Epskamp & Fried, 2018), this negative association could be interpreted as the common effect of verbal storage and domain-general ability on verbal WM performance. For verbal complex span tasks, it is expected that the performance is positively influenced by subjects’ verbal storage as well as their domain-general executive functions. Thus, if a subject with relatively low verbal storage (as primarily reflected by Sv) scored high on verbal working memory tests (Wv), the subject could be expected to be relatively more “capable” on domain-general abilities. Therefore, this subject could be expected to have relatively lower scores at verbal simple span tasks but high scores at spatial working memory tests. Thus, after controlling for subjects’ performance on Wv, Sv (highly influenced by verbal storage) and Ws (highly influenced by domain-general abilities) could be negatively correlated in the network. Therefore, considering the inclusion of the verbal WM node in the network, the negative partial correlation edge between Sv and Ws reflected the common effect of domain-general ability and (verbal) domain-specific ability to verbal memory span tasks. In general, these negative edges in the current latent network model may be a reflection of the complex and overlapping nature of different types of span tasks.

Similar to the standard network model of memory span tasks, the standard network model of memory span tasks and reasoning tests illustrated a positive manifold of all cognitive tasks and the small clusters of nodes by their task types. In the standard network model, the memory span nodes and reasoning nodes were positively associated in general. Furthermore, in the current network, the nodes of the complex span tasks played relatively central roles among the memory span tasks, and the nodes of the matrix reasoning tests played relatively central roles among the reasoning tests. However, although complex span tasks (as WMC measures) have been demonstrated to be strongly correlated with matrix reasoning tests (as fluid intelligence measures), and both types of tasks were considered to tap into similar cognitive functions, they were not directly connected in the current model. In the current standard network model, there were few edges between WM nodes and reasoning nodes retained after optimization, indicating that those edges were either not significant or were low in magnitude, so they were neglected during the optimization procedure. In other words, the associations between complex span tasks and matrix reasoning tests were primarily “mediated” by the simple span tasks and domain-specific reasoning tests.

Therefore, the results of the standard network model were partially consistent with the original findings from Kane et al. (2004). Consistent with Kane et al. (2004), the current network model indicated that the STM measures were indeed strong predictors of domain-specific reasoning and less-strong predictors of Gf. However, the current standard network model did not clearly support that WM was a strong predictor of Gf and a weaker predictor of domain-specific reasoning, as was indicated by the SEM models in Kane et al. (2004). Although the current network model was data-driven, these exploratory results indicated that the potential influence of domain-specific cognitive processes between WM and reasoning may have been overlooked in traditional latent factor models.

The latent network model of WM and reasoning relied on a relatively confirmatory structure that specified the latent broad abilities by task types and recognized a total of seven latent nodes based on task types. Similar network patterns to the standard network model were observed. The results indicated that the latent nodes of STM, WM, and reasoning were positively connected in general, illustrating the positive manifold of the cognitive tests in a network statistics framework. Furthermore, the STM nodes, but not the WM nodes, were connected directly to the reasoning nodes. This result replicated the standard network model of WM and reasoning and emphasized that working memory and reasoning may have more domain-specific overlap than conventional (confirmatory) latent factor models have suggested. It was also worth noting that in the current latent network model after optimization, there was no edge between Ss and Sv, while there was an edge between Rs and Rv, as well as an edge between Ws and Wv. The edges were partial correlations between nodes when all other information was controlled for, thus they could be interpreted as either interaction between the nodes or the shared variance from a common cause that could not be partial out by controlling for other variables. One potential explanation of these unique edges of Rs – Rv and Ws – Wv is that these edges indicate domain-general overlaps of cognitive processes in reasoning tests and complex span tasks, respectively, that could not be accounted for by the other variables in the current network model. Another potential explanation is that the more complex tasks, such as reasoning tests and complex span tasks, may involve mutual interactive effects of domain-general cognitive processes and domain-specific processes. Such interaction effects are dependent on task types and the complexity of the task (the degree to which corresponding cognitive processes are sampled in a task), leading to the unique shared variance that could not be partial out in the network.

Overall, the current study reanalyzed the working memory and reasoning data in Kane et al. (2004) using standard and latent network models. Compared to the original study using CFA and SEM models, the psychometric network analyses were exploratory and data-driven. However, compared to the original findings from Kane et al., the results of the network analyses emphasized more on the roles of domain-specific abilities. These results of network models on Kane et al. (2004) data suggested that working memory and reasoning could have more domain-specific overlap than conventional factor models have presented, indicating alternative interpretations of the relationships between domain-general and domain-specific abilities in working memory and reasoning based on POT.

9 General Discussion

This dissertation investigated a contemporary theoretical framework of intelligence, the process overlap theory (POT; Kovacs & Conway, 2016; 2019), that attempts to bridge the gap between psychometric and cognitive theories of human intelligence. POT proposed a novel psychometric structure and cognitive architecture to explain individual differences in higher-order cognition. The dissertation consisted of three studies that systematically illustrated and investigated the POT framework using a combination of latent factor models, item response models, and psychometric network models of both simulated and empirical cognitive testing data.

Study 1 illustrated a simulation based on the sampling algorithm of POT (as was proposed in POT-V and POT-I) to demonstrate that, under the POT algorithm, the positive manifold can emerge at the psychometric level in the absence of a general cognitive ability at the cognitive level. This simulation was developed to conceptually prove the sampling mechanism of POT at both the cognitive level and the psychometric level. In the POT algorithm, simulated cognitive processes were sampled and overlapped in different intelligence tests under regulations based on cognitive mechanisms. The psychometric results of the simulated data, in the form of latent factor models (POT-S), replicated the higher-order factor model observed in empirical psychometric theories such as the CHC theory. A standard higher-order “general intelligence” model fitted the simulated testing data well, even though there was no general cognitive ability involved in generating the data.

Study 2 examined the simulation dynamics based on the POT algorithm in the form of network models (POT-N). This study aimed to investigate the correlational structure of the simulated cognitive testing scores based on POT using both a standard network model (Borsboom et al., 2021) and a latent network model (Epskamp et al., 2017). The standard network model is fully exploratory and data-driven, while the latent network model retains benefits from both exploratory and confirmatory approaches. In a standard network model, no latent common causes are assumed, instead, the covariance of observed variables is examined by the estimated partial correlations between pairs of variables. In a latent network model, confirmatory latent factors can be specified for manifest variables and the covariance of these factors is still explored by partial correlations between pairs of factors. In study 2, the standard network model presented a network of the 9 simulated test scores that revealed the three broad abilities. The latent network model replicated similar findings from the standard network model, but under confirmatory assumptions regarding the broad abilities as latent factors (latent nodes). In both models, measures of fluid ability played a more central role than measures of either spatial or verbal ability, such that the edges between fluid nodes and spatial/verbal nodes were stronger than those between spatial and verbal nodes. This emphasis on fluid ability in the two network models reflected the sampling mechanism proposed by POT, in which domain-general cognitive processes were sampled more often than domain-specific processes across different types of tests. Overall, the latent network model was more compatible with POT-S than the standard network model, in which the latent nodes are equivalent to the latent broad ability factors. The formative *g* in POT-S is, in fact, conceptually comparable to the network structure of the latent nodes: both of them are determined by the indicators rather than being the cause of the variance among indicators.

Study 3 investigated the applicability of network modeling, which is currently the most compatible psychometric approach to POT, on an existing dataset of cognitive ability tests. Using a standard network model and a latent network model, Study 3 re-estimated the covariances among 12 working memory tasks and 13 reasoning tests in the absence of the assumption of a domain-general cause. The network models of Kane et al. (2004) shed light on alternative interpretations of the relationships between domain-general and domain-specific abilities in working memory and reasoning based on POT.

Overall, the results of these three studies demonstrated that, based on the POT algorithm, the positive manifold of intelligence can emerge at the psychometric level in the absence of a general mental ability at the cognitive level. This dissertation provides critical supportive evidence for POT and illustrates an alternative theoretical and statistical framework for contemporary research of human cognition that combines psychometric and cognitive theories of intelligence.

There are also potential limitations of the current studies that could inspire corresponding future research. An important limitation of the current POT simulation framework is that in the current sampling mechanism, the "cognitive processes" are only random numbers that are conceptually equivalent. The cognitive distinctions are specified only by the sampling mechanism such as the sampling probabilities. This could be updated by introducing different types of cognitive processes in the simulation that could account for more behavioral and neural science observations, such as simulated cognitive processes/mechanisms that could reflect the statistical parameters and cognitive properties of processing speed. One potential direction is to integrate the drift-diffusion model in the current POT-I algorithm. The drift-diffusion model, combined with the current POT-I simulation framework, may represent a more realistic decision-choice mechanism that is related to response time parameters in cognitive tasks. Another important limitation of the current simulation framework is that the simulated tests are relatively simple compared to empirical observations in cognitive and psychometric research. In the current simulation framework, there are a total of 4 types of cognitive processes and each of the tasks only samples no more than 2 types. This is highly unlikely in the real world. More types of simulated cognitive processes and complex tasks could be introduced in the current simulation method. For the network modeling aspect, the current dissertation only applied the psychometric network models to one empirical dataset of cognitive testing. Given the exploratory nature of network psychometrics, it is important to apply this contemporary psychometric approach to other samples of cognitive tests and subjects before reaching a generalizable interpretation of the domain-general and domain-specific emphasis based on POT. Network psychometrics is also developing rapidly, and recent techniques encourage more confirmatory analyses using network models such as multi-group invariance comparison, time-series modeling, or even meta-network analysis (Epskamp, 2020; Isvoranu et al., 2021). Network models using these techniques could be more comparable to conventional psychometric models of behavioral data as well as neuroscience data. The applications of network psychometrics could inspire unification between psychometric and cognitive research and improve the measurement, explanation, and prediction of individual differences in human intellectual behaviors at different levels of analysis.

10 References

Anderson, M. (1992). *Intelligence and development: A cognitive theory*. Blackwell Publishing.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.005

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain*, *135*(4), 1154-1164. https://doi.org/10.1093/brain/aws021

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review*, *116*(3), 567-579. https://doi.org/10.1037/a0016262

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, *29*(2), 109-125. doi:10.1017/S0140525X06009034

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic*, *35*(6), 35-37.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., ... & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, *1*(1), 1-18. https://doi.org/10.1038/s43586-021-00055-w

Brody, E. B., & Brody, N. (1976). *Intelligence: Nature, determinants, and consequences*. New York: Academic Press.

Brown, A., & Croudace, T. J. (2015). Scoring and estimating score precision using multidimensional IRT models. In Reise, S. P., & Revicki, D. A. (Eds.), *Handbook of item response theory modeling: Applications to typical performance assessment* (pp. 325-351). Routledge.

Routledge.Burns, N. R., Nettelbeck, T., & White, M. (1998). Testing the interpretation of inspection time as a measure of speed of sensory processing. *Personality and Individual Differences*, *24*(1), 25-39. https://doi.org/10.1016/S0191-8869(97)00142-6

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies* (No. 1). Cambridge University Press.

Carroll, J. B. (1997). Psychometrics, intelligence, and public perception. *Intelligence*, *24*(1), 25-52. https://doi.org/10.1016/S0160-2896(97)90012-X

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*(3), 153–193. https://doi.org/10.1037/h0059973

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. https://doi.org/10.1037/h0046743

Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, *15*(3), 139-164. https://doi.org/10.1111/j.1745-3984.1978.tb00065.x

Cowan, N. (2005). Understanding Intelligence: A Summary and an Adjustable-Attention Hypothesis. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 469–488). Sage Publications, Inc. https://doi.org/10.4135/9781452233529.n25

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, *24*(4), 1158-1170. https://doi.org/10.3758/s13423-016-1191-6

Conway, A. R. A, Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786. https://doi.org/10.3758/BF03196772

Conway, A. R. A, Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547-552. https://doi.org/10.1016/j.tics.2003.10.005

Conway, A. R. A, & Kovacs, K. (2015). New and emerging models of human intelligence. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(5), 419-426. https://doi.org/10.1002/wcs.1356

Conway, A. R. A, & Kovacs, K. (2020). Working Memory and Intelligence. In R. Sternberg (Ed.), *The Cambridge handbook of intelligence* (Cambridge Handbooks in Psychology, pp. 504-527). Cambridge: Cambridge University Press. Doi:10.1017/9781108770422.022

Conway, A. R. A, Kovacs, K., Hao, H., Goring, S. A., & Schmank, C. (2020). The Struggle is Real: Challenges and Solutions in Theory Building. *Psychological Inquiry*, *31*(4), 302-309. DOI: 10.1080/1047840X.2020.1853468

Conway, A. R. A., Kovacs, K., Hao, H., Rosales, K. P., & Snijder, J.-P. (2021). Individual Differences in Attention and Intelligence: A United Cognitive/Psychometric Approach. *Journal of Intelligence*, *9*(3), 34. https://doi.org/10.3390/jintelligence9030034

Deary, I. J. (2003). Reaction time and psychometric intelligence: Jensen's contributions. In *The Scientific Study of General Intelligence* (pp. 53-75). Pergamon. https://doi.org/10.1016/B978-008043793-4/50040-4

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Ecker, U. K., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language*, *74*, 77-90. https://doi.org/10.1016/j.jml.2013.09.003

Embretson, S., & McCollam, K. (2000). Psychometric Approaches to Understanding and Measuring Intelligence. In R. Sternberg (Ed.), *Handbook of intelligence* (pp. 423-444). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511807947.020

Engle, R. W., & Kane, M. J. (2004). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory*, Vol. 44, pp. 145–199). Elsevier Science.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195-212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617. https://doi.org/10.1037/met0000167

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, *82*(4), 904-927. https://doi.org/10.1007/s11336-017-9557-x

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd ed.). John Wiley & Sons Inc.

Flanagan, D. P., Alfonso, V. C., & Reynolds, M. R. (2013). Broad and narrow CHC abilities measured and not measured by the Wechsler scales: Moving beyond within-battery factor analysis. *Journal of Psychoeducational Assessment*, *31*(2), 202-223. https://doi.org/10.1177/0734282913478047

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186-204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172-179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Frischkorn, G. T., & Schubert, A. L. (2018). Cognitive models in intelligence research: Advantages and recommendations for their application. *Journal of Intelligence*, *6*(3), 34. https://doi.org/10.3390/jintelligence6030034

Frischkorn, G. T., & von Bastian, C. C. (2021). In Search of the Executive Cognitive Processes Proposed by Process-Overlap Theory. *Journal of Intelligence*, *9*(3), 43. https://doi.org/10.3390/jintelligence9030043

Frischkorn, G. T., Schubert, A. L., & Hagemann, D. (2019). Processing speed, working memory, and executive functions: Independent or inter-related predictors of general intelligence. *Intelligence*, *75*, 95-110. https://doi.org/10.1016/j.intell.2019.05.003

Gardner, M. K. (2011). Theories of intelligence. *The Oxford handbook of school psychology*, 79-100. DOI: 10.1093/oxfordhb/9780195369809.013.0035

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence*, *29*(6), 523-535. https://doi.org/10.1016/S0160-2896(01)00078-2

Gustafsson, J. E. (2014). Hierarchical models of individual differences in cognitive abilities. In *Advances in the psychology of human intelligence* (pp. 35-71). Psychology Press.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, *75*(3), 242–259. https://doi.org/10.1037/h0025662

Horn, J. L., & Blankson, N. (2005). Foundations for Better Understanding of Cognitive Abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 41–68). The Guilford Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*(5), 253–270. https://doi.org/10.1037/h0023816

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta psychologica*, *26*, 107-129. https://doi.org/10.1016/0001-6918(67)90011-X

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*(6), 417–441. https://doi.org/10.1037/h0071325

Hunt, E. (2010). *Human intelligence*. Cambridge University Press.

Jensen, A. R. (1987). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 101–175). Ablex Publishing.

Jensen, A. R. (1998). The suppressed relationship between IQ and the reaction time slope parameter of the Hick function. *Intelligence*, *26*(1), 43-52. https://doi.org/10.1016/S0160-2896(99)80051-8

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

Jensen, A. R., & Munro, E. (1979). Reaction time, movement time, and intelligence. *Intelligence*, *3*(2), 121-126. https://doi.org/10.1016/0160-2896(79)90010-2

Kan, K. J., van der Maas, H. L., & Kievit, R. A. (2016). Process overlap theory: Strengths, limitations, and challenges. *Psychological Inquiry*, *27*(3), 220-228. https://doi.org/10.1080/1047840X.2016.1182000

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47. https://doi.org/10.1037/0096-3445.132.1.47

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kaufman, J. C., Kaufman, S. B., & Plucker, J. A. (2013) Contemporary Theories of Intelligence. In *The Oxford Handbook of Cognitive Psychology*.

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Wicherts, J. M., Scholte, H. S., & Borsboom, D. (2011). Mind the gap: a psychometric approach to the reduction problem. *Psychological Inquiry*, *22*(2), 67-87. DOI: 10.1080/1047840X.2011.550181

Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry*, *27*(3), 151-177. https://doi.org/10.1080/1047840X.2016.1153946

Kovacs, K., & Conway, A. R. A. (2019). A unified cognitive/differential approach to human intelligence: Implications for IQ testing. *Journal of Applied Research in Memory and Cognition*, *8*(3), 255-272. https://doi.org/10.1016/j.jarmac.2019.05.003

Kovacs, K., Conway, A. R. A., Snijder, J., & Hao, H. (2018). General intelligence explained (Away): Poster presented at: The 59th annual meeting of the psychonomic society. *New Orleans*, *2018*, 15-18.

Kovacs, K., Molenaar, D., & Conway, A. R. A. (2019). The domain specificity of working memory is a matter of ability. *Journal of Memory and Language*, *109*, 104048. https://doi.org/10.1016/j.jml.2019.104048

Kovacs, K., Conway, A. R. A., Hao, H., & Snijder, J. (2022). *General Intelligence Explained (Away)* [Unpublished manuscript]. Institute of Psychology, ELTE Eotvos Lorand University

Lally, M., & Nettelbeck, T. (1980). Inspection time, intelligence, and response strategy. *American Journal of Mental Deficiency*, *84*(6), 553–560.

Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*, *14*(3), 309-325. https://doi.org/10.1016/0160-2896(90)90021-K

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281. https://doi.org/10.1038/36846

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*(1), 201-226. https://doi.org/10.1146/annurev.psych.51.1.201

McGrew, K. S. (2005). The Cattell-Horn-Carroll Theory of Cognitive Abilities: Past, Present, and Future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136–181). The Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *1*(37), 1-10. https://doi.org/10.1016/j.intell.2008.08.004

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Molenaar, D., Tuerlinckx, F., & van der Maas, H. L. (2015). Fitting diffusion item response theory models for responses and response times using the R package diffIRT. *Journal of Statistical Software*, *66*(4), 1-34. https://doi.org/10.18637/jss.v066.i04

Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., ... & Urbina, S. (1996). Intelligence: knowns and unknowns. *American Psychologist*, *51*(2), 77. https://doi.org/10.1037/0003-066X.51.2.77

Nettelbeck, T. (2001). Correlation between inspection time and psychometric abilities: A personal interpretation. *Intelligence*, *29*(6), 459-474. https://doi.org/10.1016/S0160-2896(01)00072-1

Nettelbeck, T. (2003). Inspection time and g. *The Scientific Study of General Intelligence*, 77-91.

Nettelbeck, T., Zwalf, O., & Stough, C. (2020). Basic Processes of Intelligence. In R. Sternberg (Ed.), *The Cambridge handbook of intelligence* (Cambridge Handbooks in Psychology, pp. 471-503). Cambridge: Cambridge University Press. DOI:10.1017/9781108770422.021

Oberauer, K. (2005). The measurement of working memory capacity. In *Handbook of understanding and measuring intelligence* (pp. 393-408). SAGE Publications Inc..

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45-100. https://doi.org/10.1016/S0079-7421(09)51002-X

Oberauer, K. (2019). Working memory and attention–A conceptual analysis and review. *Journal of Cognition*, *2*(1). http://doi.org/10.5334/joc.58

Oberauer K. (2019). Working Memory Capacity Limits Memory for Bindings. *Journal of Cognition*, *2*(1), 40. https://doi.org/10.5334/joc.86

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. https://doi.org/10.1037/bul0000153

Ortiz, S. O. (2015). CHC theory of intelligence. In *Handbook of intelligence* (pp. 209-227). Springer, New York, NY.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46-59. https://doi.org/10.1002/hbm.20131

Pellegrino, J. W., & Glaser, R. (1979). Cognitive correlates and components in the analysis of individual differences. *Intelligence*, *3*(3), 187-215. https://doi.org/10.1016/0160-2896(79)90017-5

Rey-Mermet, A., Gade, M., Souza, A. S., von Bastian, C. C., & Oberauer, K. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology: General*, *148*(8), 1335–1372. https://doi.org/10.1037/xge0000593

Roid, G. H., & Pomplun, M. (2012). *The Stanford-Binet intelligence scales*. The Guilford Press.

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*(1), 5-21. https://doi.org/10.1016/j.neuroscience.2005.12.061

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403. https://doi.org/10.1037/0033-295x.103.3.403

Savi, A. O., Marsman, M., van der Maas, H. L., & Maris, G. K. (2019). The wiring of intelligence. *Perspectives on Psychological Science*, *14*(6), 1034-1061. https://doi.org/10.1177/1745691619866447

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136*(3), 414–429. https://doi.org/10.1037/0096-3445.136.3.414

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 99–144). The Guilford Press.

Schrank, F. A. (2011). Woodcock-Johnson III Tests of Cognitive Abilities. In A. S. Davis (Ed.), *Handbook of pediatric neuropsychology* (pp. 415–434). Springer Publishing Company.

Schrank, F. A., & Wendling, B. J. (2018). The Woodcock-Johnson IV: Tests of cognitive abilities, tests of oral language, tests of achievement. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 383–451). The Guilford Press.

Schrank, F. A., Decker, S. L., & Garruto, J. M. (2016). *Essentials of WJ IV cognitive abilities assessment*. John Wiley & Sons.

Schubert, A. L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence*, *51*, 28-46. https://doi.org/10.1016/j.intell.2015.05.002

Spearman, C. (1904). " General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201-292.

Spearman, C. (1927). *The abilities of man.* London: Macmillan.

Sternberg, R. J. (1985). Human intelligence: The model is the message. *Science*, *230*(4730), 1111-1118. https://doi.org/10.1126/science.230.4730.1111

Sternberg, R. J. (2012). Intelligence. WIREs Cognitive Science, 3 (5), 501–511.

Sternberg, R. J. (2013). Contemporary theories of intelligence. In W. M. Reynolds, G. E. Miller, & I. B. Weiner (Eds.), *Handbook of psychology: Educational psychology* (pp. 23–44). John Wiley & Sons, Inc..

Sternberg, R. (2020). A History of Research on Intelligence: Part 2: Psychological Theory, Research, and Practice in the Nineteenth and Twentieth Centuries. In R. Sternberg (Ed.), *The Cambridge handbook of intelligence* (Cambridge Handbooks in Psychology, pp. 31-46). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108770422.004

Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology*, *8*(3), 271.

Thomson, G. (1951). *The factorial analysis of human ability*. University of London Press.

Thorndike, E. L. (1925). *The measurement of intelligence*. New York: Columbia University, Teachers College.

Thurstone, L. L. (1935). *The vectors of mind: Multiple-factor analysis for the isolation of primary traits.* University of Chicago Press. https://doi.org/10.1037/10018-000

Thurstone, L.L. (1938). *Primary mental abilities.* University of Chicago Press: Chicago.

Thurstone, L. L. (1947). *Multiple-factor analysis: A development and expansion of the vectors of mind*. University of Chicago Press.

Thurstone, L. L., & Thurstone, T. G. (1941). Factorial studies of intelligence. *Psychometric Monographs, 2*, 94.

Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*(4), 842–861. https://doi.org/10.1037/0033-295X.113.4.842

Van Der Maas, H. L., Kan, K. J., Marsman, M., & Stevenson, C. E. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence, 5*(2), 16.

Kan, K. J., van der Maas, H. L., & Levine, S. Z. (2019). Extending psychometric network analysis: Empirical evidence against g in favor of mutualism?. *Intelligence, 73*, 52-62. https://doi.org/10.1016/j.intell.2018.12.004

Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen & Co.

von Bastian, C. C., Blais, C., Brewer, G. A., Gyurkovics, M., Hedge, C., Kałamała, P., … Wiemers, E. A. (2020, July 27). Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations. https://doi.org/10.31234/osf.io/x3b9k

Walrath, R., Willis, J., Dumont, R., & Kaufman, A. (2020). Factor-Analytic Models of Intelligence. In R. Sternberg (Ed.), *The Cambridge handbook of intelligence* (Cambridge

Handbooks in Psychology, pp. 75-98). Cambridge: Cambridge University Press.

https://doi.org/10.1017/9781108770422.006