Journal of Applied Research in Memory and Cognition 8 (2019) 255–272

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Review

A Unified Cognitive/Differential Approach to Human Intelligence: Implications for IQ Testing

Kristof Kovacs
ELTE Eotvos Lorand University, Hungary

Andrew R.A. Conway*
Claremont Graduate University, United States

Individual differences have been mostly ignored in cognitive/experimental psychology since the birth of the field, while the measurement of cognitive abilities has become a successful field of applied psychology. Because of its separation from mainstream research psychology, cognitive ability testing has focused on application, often without providing sound theoretical basis for the tests. More recently, the gap between cognitive/experimental psychology and differential/psychometric research has been closing. This stems from a rediscovery of variation in cognitive abilities in experimental psychology, owing largely to the concept of working memory. We present process overlap theory, a new theory of intelligence that is informed by cognitive psychology. The theory explains the positive correlations between diverse tests on the basis of overlapping cognitive processes and reinterprets the general factor of intelligence, *g*, as a formative construct. The consequences of this approach are discussed, including a focus on specific abilities rather than on global scores in cognitive test results.

General Audience Summary

There are many examples of human cognitive performance, from reading difficult texts to performing mathematical operations to solving complex problems. Success in these various activities is correlated: those who are better in one area are usually better in the rest, too. This is the most important finding in the field of human intelligence and it has led to the idea that, despite superficial dissimilarity, these cognitive activities all depend on the same general cognitive ability. However, in cognitive psychology and neuroscience there is ample evidence *against* this idea and for the fractionation of cognition into distinct faculties. But due to historical reasons the study of cognition mostly ignored individual differences, while the study of human intelligence was mostly uninformed by the general study of cognition and neuroscience. The question of general intelligence versus specific abilities is one of the oldest debates in psychology. In this paper we present process overlap theory (POT), which explains the correlations among performance measures from different tests without proposing a general cognitive ability. Instead, POT focuses on limitations of cognitive capacity, determined by processes involved in sustained attention, mental flexibility, planning, and the like. Limited processing capacity will affect performance in a number of areas, regardless of specific abilities. According to POT, general intelligence is a summary of different but correlated abilities rather than the reflection of a single, unitary ability. This approach

Author Note.

Kristof Kovacs, Department of Psychology, ELTE Eotvos Lorand University, Hungary.

Andrew R.A. Conway, Department of Psychology, Claremont Graduate University, United States.

We would like to thank Jean-paul Snijder, Han Hao, and Chris Schmank for assistance with the simulations and figures.

* Correspondence concerning this article should be addressed to Andrew R.A. Conway, Claremont Graduate University, United States. Contact: andrew.conway@cgu.edu.

has consequences for applied cognitive testing. For if the theory of general intelligence is correct then the optimal level of evaluating performance on cognitive ability tests is a global score. In contrast, if POT is correct then the focus should be on specific abilities that can provide a cognitive profile of strengths and weaknesses.

Keywords: Intelligence, IQ, Working memory, Executive function, Cognitive ability

While most cognitive psychologists and neuroscientists today would probably say that there is no such thing as general intelligence, psychometricians have become remarkably good at measuring it. We hope that by the end of this paper we will have convinced you that the previous sentence is not sarcastic. How is it possible, then, to be remarkably good at measuring something that does not exist? Before we are able to provide a (hopefully) reassuring answer, we need to present a story of two scientific disciplines, experimental/cognitive psychology and intelligence research/psychometrics, that diverged for most of the 20th century and only started to converge quite recently. This will be followed by a summary of our own unified approach—process overlap theory—that attempts to explain major findings about human intelligence on the basis of cognitive and neural processes. Finally, the implications of this approach are discussed for both research and application.

A Long History of Divergence and a Short History of Convergence

IQ tests have a bad reputation outside the field of human intelligence. This is quite remarkable for an instrument designed to help children. When Alfred Binet constructed the first test of intelligence his goal was to design an objective tool to decide which children would benefit from educational interventions. Subsequently, the history of IQ tests has become over-politicized as a result of the societal implications of the nature—nurture debate and the very existence of cognitive inequalities. Their reputation was further damaged with the involvement of IQ testing in practices such as the forced sterilization of the so-called "feeble-minded" or educational segregation—even though most of the claims about the historical role of early IQ testing in such practices are greatly exaggerated (Mackintosh, 2011). Nevertheless, the reputation of intelligence and IQ is bad among many intellectuals.

> Just mention IQ in polite company, and you'll be informed (sometimes rather sternly) that IQ tests don't measure anything real, and reflect only how good you are at doing IQ tests; that they ignore important things like 'multiple intelligences' and 'emotional intelligence;' and that those who are interested in intelligence testing must be elitists, or perhaps something more sinister. (Ritchie, 2015, p. 8)

Achievement testing is somewhat less controversial than ability testing, even though the correlation between IQ and educational measures such as the SAT in the US (Frey & Detterman, 2004) or GCSE in the UK (Deary, Strand, Smith, & Fernandes, 2007) is so high that, purely on the basis of correlations, these well known educational tests could qualify as tests of intelligence themselves.

Paradoxically, IQ tests are also one of psychology's most visible success stories. Cognitive ability is regularly assessed in clinical settings, as well as in industrial/organizational (I/O) psychology and school settings. Testing methods have become sophisticated, and there has been substantial development in testing technology, too. Yet, since psychometrics has evolved independently of mainstream cognitive/experimental psychology for most of the last century, test construction has not been based on psychological theory—or had not been up to relatively recently.

As a result, ever since Alfred Binet created the first version of the modern IQ test measuring individual differences as accurately as possible was more important than explaining them (Mackintosh, 2011). This does not mean that intelligence researchers have since been working completely detached from cognitive or mainstream psychology. On the contrary, there have been several attempts to merge intelligence research with cognitive psychology (Hunt, Lunneborg, & Lewis, 1975; Spearman, 1923; Sternberg, 1980). Yet, by and large, the fields of intelligence and cognitive psychology remained largely separated, with the former taking a correlational approach to measuring individual differences and the latter an experimental approach to explaining universal cognitive phenomena. Cronbach, in his 1957 presidential address to the American Psychological Association, discussed what he called the two disciplines of psychology, representing these two approaches, and concluded that "psychology continues to this day to be limited by the dedication of its investigators to one or the other method of inquiry rather than to scientific psychology as a whole" (Cronbach, 1957, p. 671).

Despite the achievements of numerous intelligence researchers who incorporated the findings of cognitive psychology to their own research, the two communities of researchers are still mostly distinct. Intelligence research is still mostly correlational/differential, with intelligence researchers often being involved in other differential areas, such as personality, rather than cognitive psychology. This is also apparent in little overlap in journals and conferences.

This is a remarkable situation because, strictly speaking, most of the field of intelligence, at least those areas that deal with the structure and causes of individual differences in abilities, could have been labelled "individual differences in cognition" rather than intelligence all along. As well, variation in cognition could have been a standard subsection of major cognitive-psychology conferences, just like autobiographical memory or decision-making. That this is not the case is due at least as much to attitudes as to scientific differences. In most of cognitive/experimental psychology variation is traditionally treated as noise. As Cronbach pointed out, “the correlational psychologist is in love with just those variables the experimenter left home to forget” while “individual variation is a source of embarrassment to the experimenter” (Cronbach, 1957, p. 674).

As arbitrary as it seems, this attitude difference seems to be fundamental. Take, for instance, one of cognitive psychology’s most well cited discoveries, Miller’s “magical number 7,” as the limit of short-term memory capacity:

> There is a clear and definite limit to the accuracy with which we can identify absolutely the magnitude of a unidimensional stimulus variable. I would propose to call this limit the span of absolute judgment, and I maintain that for unidimensional judgments this span is usually somewhere in the neighborhood of seven. (Miller, 1956, p. 90)

The focus on the universality of this finding is a matter of choice. As later research revealed, individual differences in such judgments are important for a number of real-life outcomes (Engle, 2002). Importantly, the magical number seven was in fact seven *plus or minus two*; individual differences were acknowledged, but without attributing much significance to that variation.

Now for the sake of the argument imagine a test battery of several memory-span tasks. The average raw score on this battery of tests would have a meaningful standard deviation, and it is only a matter of rescaling to transform this raw score to an IQ-type scale with a mean of 100 and standard deviation of 15. If the battery actually consisted of 14 tasks then even the magnitude of the raw score distribution might remotely resemble an IQ-scale (i.e., $14 \times 7 = 98$). It might sound absurd, but it is also certainly possible to interpret the results of intelligence assessment in a way similar to the magical number, that is, with an exclusive focus on the mean while dispersion is worded as a side note. Yet obviously no one would ever claim that the magical number of human intelligence is 100 (plus or minus 30), even though approximately 96% of people indeed do have an IQ of 100 plus or minus 30, just like they have a short-term memory span of 7 plus or minus 2. One could, of course, point out that IQ is a scaled score, while short term memory span is not. Which is indeed the case, but from that it also follows that IQ could be scaled to a mean of 7 instead of 100 and with a standard deviation of 1 instead of 15. It appears as if approaching variation as meaningful or as noise around a mean was indeed a mostly subjective matter.

Of course, the magnitude of variation is not the only issue to consider. There appears to be a general opinion in cognitive psychology, especially in areas informed by evolutionary biology or evolutionary psychology, than human universals are much more important phenomena than individual differences, both functionally and in terms of complexity (Brown, 1991; Tooby & Cosmides, 1992). Indeed, it appears that being able to create a mental representation of the world that includes colors is a more complex and fascinating phenomenon than individual differences in color discrimination. Similarly, the phenomenon that almost all humans are able to use language, to have a mental lexicon or read written texts, seems to be functionally much more important than the variation in reading comprehension or vocabulary. And given that someone with an IQ of 130 and someone with an IQ of 100 are both capable of perceiving colors and learning to read, the importance of abilities measured by intelligence tests seem to be rather restricted.

Yet since intelligence tests are widely used to make decisions about educability in schools, in cognitive treatment of the elderly, or even about the death penalty (Polloway, 2013) it is beneficial for psychology to have tools to measure individual differences that are informed by research on cognition. Fortunately, a long history of divergence was recently followed by a short history of convergence, mostly due to research on working memory. In the last 10–15 years the number of papers that discuss both working memory and intelligence has grown exponentially (Conway & Kovacs, 2013). But even that growth has antecedents in developmental psychology. In particular, in the 1970s developmental psychologists began to incorporate ideas from information processing models of cognition (Case, Kurland, & Goldberg, 1982; Pascal-Leone, 1970). This work had great impact on the field of working memory and set the stage for unified research on intelligence and working memory.

Working memory is a construct developed by cognitive psychologists to refer to the *intra-individual processes* that enable one to hold goal-relevant information in mind, often in the face of concurrent processing or distraction (Baddeley, 1992; Baddeley & Hitch, 1974). Working memory has become enormously influential in cognitive psychology, and so-called complex span tasks were developed on the basis of its theoretical assumptions. Complex span tasks require the parallel storage and processing of information, in contrast to earlier measures of short term storage, like forward digit span, that focus exclusively on storage.

One of the first complex span tasks was reading span, in which subjects have to read sentences and remember the last word of each sentence (Daneman & Carpenter, 1980). Variation in performance on this particular measure of working memory capacity predicts performance on reading comprehension tests. At the same time, developmental psychologists were using a task called counting span, in which children are instructed to count objects in a display and remember the sums for later recall (Case et al., 1982).

Over the years, many versions of complex span tasks have been developed (for a review of working memory span tasks, see Conway et al., 2005). Importantly, the construction of complex span tasks is embedded in mainstream research in cognitive psychology on intra-individual processes, yet such tasks not only demonstrated substantial variation, but proved to be able to predict the same outcomes that tests of intelligence predict. As a result, the most well-known batteries of intelligence started to incorporate tasks originally designed for working memory research in cognitive psychology and neuropsychology, such as the letter-number sequencing task (Gold, Carpenter, Randolph, Goldberg, & Weinberger, 1997) in the Wechsler Intelligence Scales or the nonword-repetition task (Gathercole, Willis, Baddeley, & Emslie, 1994) in the most recent version of the Woodcock-Johnson tests.

However, the convergence of differential and experimental cognitive psychology is not without challenges. In particular, there are practical and theoretical issues to face. The main practical issue stems from the fact that tasks in cognitive psychology are designed for within-subject studies and as such do not necessarily meet psychometric criteria of reliability that are required to study individual differences. Working memory tasks in general and complex span tasks in particular do demonstrate appropriate reliability (Engle & Kane, 2004). However, other important tasks in cognitive psychology are more problematic. For example, Stroop, Eriksen flanker, Posner cuing, stop-signal, and go/no-go tasks all exhibit poor test-retest reliability (Hedge, Powell, & Sumner, 2018). Neuropsychological tests are also problematic. For the Wisconsin Card Sorting test alternate-form reliabilities have been found between .25 and .63 with an average of .43 (Bowden et al., 1998). For the Tower of London an appropriate test-retest reliability (0.89) was found in one study (Korkman, Kirk, & Kemp, 1998), but other studies with the Tower of Hanoi found alarming results: a test-retest reliability of .72 with a test-retest interval of only 25 min (Gnys & Willis, 1991) and a test-retest reliability of .53 for an interval of 30 to 40 days (Bishop, Aamodt-Leeper, Creswell, McGurk, & Skuse, 2001).

Inadequate reliability indicates that the use of experimental tasks in individual differences research is problematic. It also highlights that a latent variable approach that gets rid of measurement error is recommended, whereas correlations between manifest variables are prone to type II error. From the perspective of the two disciplines the conclusion is that from a measurement perspective psychometric instruments greatly outperform most of the tools used in experimental/cognitive psychology. Although the cognitivist is often theoretically more advanced, the differentialist has better instruments to study variation.

The main theoretical problem stems from the conceptual distinction between within-individual and between-individual constructs. Psychologists studying human cognition in general usually aim to identify the processes or mechanisms required to perform a given cognitive activity. Neuropsychologists studying patient populations in which one or more of these processes is impaired aim at identifying how such impairment affects the same cognitive activity. The problem is that in healthy populations these processes might not contribute to individual differences in performance. This point is made clear by Borsboom, Mellenbergh, and van Heerden (2003), in the following fictional account of Einstein completing a single item on an IQ test:

> Einstein enters the testing situation, sits down, and takes a look at the test. He then perceives the item. This means that the bottom-up and top-down processes in his visual system generate a conscious perception of the task to be fulfilled; it happens to be a number series problem. Einstein has to complete the series *1, 1, 2, 3, 5, 8, … ?* Now he starts working on the problem; this takes place in working memory, but he also draws information from long-term memory (e.g., he probably applies the concept of addition, although he may also be trying to remember the name of a famous Italian mathematician of whom this series reminds him). Einstein goes through some hypotheses concerning the rules that may account for the pattern in the number series. Suddenly he has the insight that each number is the sum of the previous two (and simultaneously remembers that it was Fibonacci). Now he applies that rule and concludes that the next number must be 13. Einstein then goes through various motoric processes that result in the appearance of the number *13* on the piece of paper, which is coded as *1* by the person hired to do the typing. Einstein now has a 1 in his response pattern, indicating that he gave a correct response to the item. This account has used various psychological concepts, such as working memory, long- term memory, perception, consciousness, and insight. But where in this account of the processes leading to Einstein's item response did intelligence enter? The answer is nowhere. Intelligence is a concept that is intended to account for individual differences. (Borsboom et al., 2003, p. 213)

From the above quote it follows that from a differentialist perspective such universal processes are noise, strictly speaking, just like variation is noise for Cronbach's experimenter. Differential psychology's emphasis on variation in processes, coupled with a lack of importance of processes that do not contribute to variation in the outcome have resulted in arguments against the unification of the two disciplines (Borsboom, Kievit, Cervone, & Hood, 2009; Jensen, 2000). Borsboom et al. (2009) and Jensen present a similar thought experiment in which extraterrestrial creatures study cars and, because of the reasons discussed above, eventually develop separate lines of research for (a) how individual cars work, and (b) what makes them differ in performance.

We disagree. If aliens were indeed like psychologists and ended up having a Differential Car Science and an Experimental Car Science, then the former would suffer from this distinction—just like, arguably, the study of individual differences in intelligence suffers from being detached from mainstream psychology. For the differential car scientist would quickly find differences in measures of performance and would also find that these differences correlate. For instance, they would find that sports cars, which have better acceleration, also have larger engines, but also have more and louder speakers and often more vivid colors. Vividness of color, better speaker performance and larger engines would thus load on the same factor, which we could call "the *s* factor" for general sportiness.

So far, so good. But which of these things has a *causal role* in better acceleration? This is the point where being informed by the research of experimentalist aliens who study what causes cars to accelerate is enormously helpful. The subpopulation of alien scientists who specialize in the study of damaged cars is also useful, for they could easily inform the differentialist that brown cars without speakers still accelerate fine, but cars without engines do not. The point is that without an understanding of how cars work the study of how cars differ will be inherently limited, and will have no choice but to focus on prediction rather than explanation. But by a unification of the two lines of research the differentialist could study which of the components discovered by experimentalists show variation, and which of these variable mechanisms have a causal role in differences in performance.

Having said that, there is a final point that needs to be emphasized before we narrow our attention to the general factor of intelligence. As much as the unification of experimental and differential psychology is warranted in our opinion, one must be cautious about the commensurability of concepts in differential and in experimental psychology. For instance, domain-specificity bears different meanings in differential psychology and cognitive/experimental psychology (Kovacs, in press). In the former it relates to the finding that individual differences in tests with characteristic content (e.g., spatial or verbal) typically correlate more strongly with one another than with tests that have different content. In the latter it means that the mind can be fractionated into processors of specific content through double dissociation.

Factors representing different variables that covary also do not represent unitary constructs from a cognitive perspective—like speakers and engine size in the car example. Therefore, it is a fallacy to directly identify inter-individual structures, usually identified by factor analysis, as intra-individual constructs. In the field of intelligence, this manifests itself as the interpretation of the general factor of intelligence, an inter-individual construct, as general intelligence or general cognitive ability, an intra-individual construct—the topic of the next section. It might be the case that there is a direct agreement between within-individual and between-individual constructs. Such cases are referred to as ergodicity, but they are exceptions rather that the rule (Molenaar & Campbell, 2009).

The General Factor of Intelligence and/or General Intelligence

When a battery of diverse cognitive tests is administered to a large sample of people, a stable empirical pattern is observed, known as the positive manifold. The positive manifold refers to the pattern of all positive correlations that is observed among different tests of cognitive ability. In other words, if an individual performs above average on one kind of test (for example, vocabulary) they tend to perform above average on other kinds of tests as well (for example, mental rotation). Overall, 40–50% of the between-individual variation in IQ test scores is domain-general (Deary, Penke, & Johnson, 2010; Jensen, 1998), therefore factor analysis yields a strong general factor, *g*, that accounts for 40–50% of the variance.

Spearman, who discovered the positive manifold (Spearman, 1904), argued that different tests correlate because they measure the same general construct, and anything a test measures beyond this general construct is entirely test-specific, thus ruling out the existence of domain-specific abilities: "all branches of intellectual activity have in common one fundamental function (or group of functions), whereas the remaining or specific elements of the activity seem in every case to be wholly different from that in all the others" (Spearman, 1904, p. 284). Several researchers debated this claim and argued that intelligence is instead multi-dimensional (Horn & Cattell, 1967; Thurstone, 1938).

A large number of factor analytic studies conducted since have falsified Spearman's early claim that all variance beyond *g* is test-specific. Among the pattern of all-positive correlations there are clusters of correlations that are stronger than others, and these clusters of strong correlations are thought to reflect what are known as group factors, representing more specific cognitive abilities. For example, a vocabulary test, a reading comprehension test, and a test of general knowledge might reveal relatively strong positive correlations within the positive manifold. This cluster, then, is thought to reflect a group factor that we might refer to as verbal or crystallized ability. Yet group factors, or specific abilities so defined, are correlated, since a large part of the entire variance is across-domains. Therefore, higher-order models that accommodate both *g* and specific abilities are empirically sound, while *g*-only models and models with orthogonal specific abilities are not.

It is beyond the scope of this paper to evaluate or even review the literature on factor analytic studies describing the structure of human cognitive abilities. Instead, there are three summary points that we would like to make. The first is that factor analysis cannot directly reveal the "architecture of cognition" from a cognitive psychological perspective for both methodological and conceptual reasons. From a methodological perspective the problem is that performing factor analysis requires a number of decisions from the researcher and these decisions greatly influence the outcome:

> [E]very time anyone constructs a hierarchical model in a new battery, it comes out looking different from any one that the same person constructed in any other battery, not to mention different from the model some other researcher would construct in that same battery in that same sample. This is because the underlying factor-analytically based methods are inherently subjective and because the relative associations among specific cognitive tasks vary both with sample specifics and with the specific other cognitive tasks in any battery. (Johnson, 2018, p. 2)

Or, even more drastically, "inferences from the results of factor analysis and structural modeling should be primarily about the structure of test batteries rather than the structure of human mental abilities" (McFarland, 2017, p. 1168). The conceptual obstacle is that the constructs of factor analysis are entirely differential, that is, they describe individual differences and, as we have seen in the previous section, such constructs do not directly map onto intra-individual processes and mechanisms that could account for the functional architecture of cognition.

The second point, seemingly somewhat in contradiction with the above paragraph, is that evidence is converging towards the so-called CHC (Cattell-Horn-Carroll) model of cognitive abilities (see Figure 1). This model (McGrew, 2009) evolved from the theory of fluid/crystallized intelligence, where fluid intelligence is the ability to solve novel problems when someone cannot rely on already acquired skills or knowledge, while crystallized intelligence refers to the ability to use skills and knowledge (Cattell, 1971; Horn, 1994). The current CHC model has seven major factors now, as depicted in Figure 1. CHC thus accommodates both *g* and more specific cognitive abilities. The CHC model appears to be an adequate statistical model for describing covariance in existing cognitive test batteries.

The third point is that even though specific abilities of the CHC model seem to replicate in numerous data sets, the

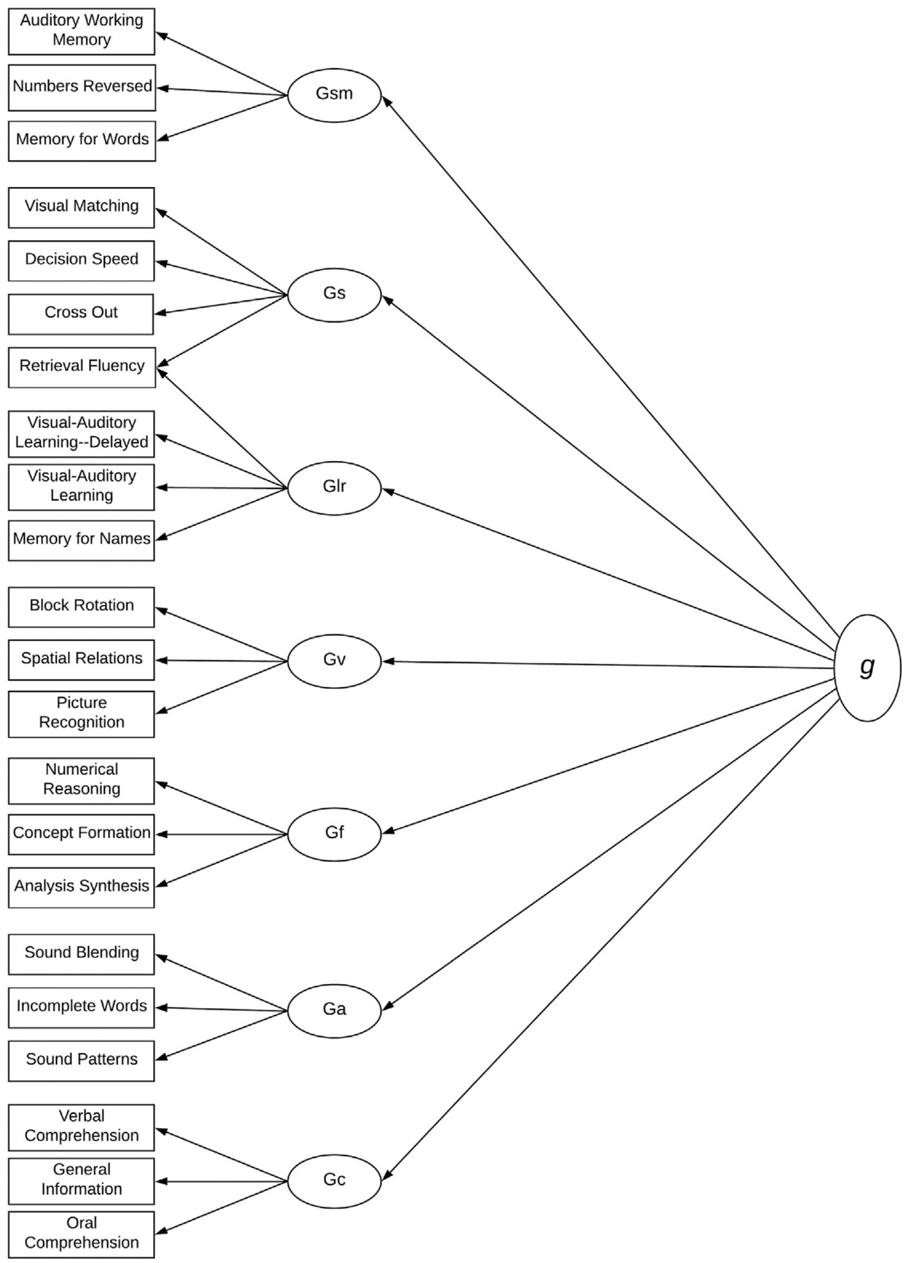

Figure 1. The factor structure of the Cattell–Horn–Carroll theory of intelligence. *Note*: Gsm = short-term memory; Gs = processing speed; Glr = long-term retrieval; Gv = visual-spatial thinking; Gf = fluid reasoning; Ga = auditory processing; Gc = crystallized intelligence; *g* = general intelligence.
Source: Adapted from Taub and McGrew (2014).

factor loadings—or correlations between specific abilities—are a function of ability itself. This phenomenon, called ability differentiation, refers to the finding that cross-domain correlations are higher in samples with lower average ability and so *g* explains more variance in such samples (Molenaar, Kő, Rózsa, & Mészáros, 2017).

Finally, from the perspective of the present paper, a crucial distinction needs to be made between a *g*-model and *g*-theory. A *g*-model is a latent variable model that can take many forms, but the common aspect is that there is a general factor explaining cross-domain variance. Note that the term "explaining" here is statistical; it means that correlations between variables, or in higher-order models like CHC, correlations between specific abilities, can be statistically accounted for by the variables' or abilities' correlation with *g*. Also note that the general factor is a necessary mathematical consequence of the positive manifold (Krijnen, 2004): It is always possible, though not necessary, to extract a general factor from a correlation matrix with only positive entries. That is, in a technical sense, *g* is a more sophisticated way of restating the positive manifold. Hence, "it is always important to remember that it is the positive manifold, not *g* as such, that needs explanation" (Mackintosh, 2011, p. 165).

In contrast, *g*-theory explains the positive manifold by interpreting the psychometric construct *g*—as in a *g*-model—as a psychological construct that is common to all tests; either an intra-individual construct, general intelligence or general cognitive ability, or a common parameter of information processing, like mental speed, that affects all cognitive activity. A *g*-theory, or the proposal of a unitary domain-general cognitive mechanism, is a sufficient but not necessary explanation of a *g*-model and, ultimately, the positive manifold. To illustrate this, consider the following two statements:

1. If John performs better on the vocabulary test than most people, it is likely that he will perform better on the mental rotation test as well. (the empirical basis of a *g*-model)
2. John used his general intelligence to correctly answer items on both the vocabulary test and the mental rotation test. (*g*-theory)

The second statement does not sit well with a number of findings from cognitive psychology and neuroscience (e.g., double dissociations, localization data, and patterns of sex differences) that point to the domain-specific fractionation of cognition and contradict the existence of a general cognitive ability. Importantly, as defined above, the *g*-model and the *g*-theory are logically differentiable since a *g*-model can fit well with the data even if the *g*-theory is false, and the positive manifold is the result of a causal mechanism other than a psychological equivalent of *g*, something that permeates all cognitive activity. Moreover, it is not only a logical possibility; in the next section we will present a mathematical model that shows that a *g*-model can fit simulated data even if *g*-theory is demonstrably false, namely, when data are generated without the effect of a single general causal mechanism.

Process Overlap Theory

Process overlap theory (POT) is a new approach to the study of intelligence that attempts to integrate psychometrics and cognitive psychology (Kovacs & Conway, 2016b; Kovacs & Conway, 2019). According to POT, intelligence is an inter-individual differences construct that can be explained in terms of intra-individual psychological processes. A consensus definition of intelligence remains elusive so, instead of a definition, POT adopts a view of intelligence that is consistent with the following statement, issued by a task force created by the Board of Scientific Affairs of the American Psychological Association:

> Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: a given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena. (Neisser et al., 1996, p. 77)

According to this view, intelligence is a complex construct; it allows for understanding, learning, and reasoning, all in multiple domains. This means that in order to conduct an optimal assessment of intelligence, the administration of a battery of diverse cognitive tests is required; a single score on a single test is rarely sufficient. As mentioned above, when a battery of diverse cognitive tests is administered to a large sample of subjects, the positive manifold is observed. Again, the positive manifold refers to the pattern of all positive correlations that are observed among different tests of cognitive ability.

The primary aim of POT, therefore, is to explain the positive manifold. POT provides an account of the positive manifold that differs from the traditional view in psychology. According to the traditional view, first proposed by Spearman (1904), the positive manifold can be explained by a general factor derived from a factor analysis (or specified in a latent variable model), and the general factor is interpreted as a general mental ability, or general intelligence. In other words, the positive manifold is explained by a *g*-model, which is interpreted in terms of a *g*-theory.

The problem with this approach, which we outlined above, is that a factor derived from a factor analysis is not equivalent to a cognitive process. To illustrate the problem, again consider the two statements presented earlier:

1. If John performs better on the vocabulary test than most people, it is likely that he will perform better on the mental rotation test as well.
2. John used his general intelligence to correctly answer items on both the vocabulary test and the mental rotation test.

The first statement follows from the positive manifold but the second statement does not. According to POT, the second statement is invalid because there is no cognitive process that is equivalent to "general intelligence." In other words, there is no intra-individual psychological process, or ability, or parameter that is common to all cognitive tests. Furthermore, cognitive tests are not process pure, meaning that each individual test requires multiple cognitive processes, or abilities, for accurate performance.

In a psychometric sense, *g* is real, and it is predictive of many important life outcomes, such as academic achievement and job performance (Gottfredson, 1997). However, according to POT, in a psychological sense, *g* is not real. It does not represent a psychological attribute. This is precisely why our opening statement in the current paper was entirely serious, and not sarcastic. It is indeed possible for there to be no such thing as general intelligence, yet it is also possible to accurately measure a thing called *g*.

This approach to explaining the positive manifold is not entirely new. Godfrey Thomson, as a direct challenge to Spearman, proposed a sampling approach to the positive manifold (Thomson, 1916). Thomson argued that there are a large number of mental "bonds" (which we now interpret as cognitive processes) and a sample of bonds is required to complete any single test of intelligence. Furthermore, different kinds of IQ tests require different samples of bonds. The correlations between IQ tests are caused by an overlap between the bonds tapped by the tests. For example, suppose there are 10 bonds and Test A samples bonds 1–5, Test B samples bonds 1–3, 6, and 7, and Test

C samples bonds 1, 2, and 7–9. The correlation between test A and B will be stronger than the correlation between test A and C, yet all the correlations will be positive.

We will not go into further details about Thomson's theory because it is the idea of sampling that is important from the perspective of this discourse. Sampling, as a general approach, can be contrasted with the factor analytic approach to *g*; these approaches provide different explanations of the positive manifold. There are different factorial models but as long as they incorporate a general factor they all propose a causal mechanism reflected by all tests to some extent and this explains why the tests are correlated. Sampling, on the other hand, proposes a large number of resources that are tapped by different tests. Under this framework tests are correlated because there are common processes involved in test performance. However, there is no single, unitary cause of the correlations, and there need not even be processes that are involved in performance on *all* tests in order for the positive manifold to emerge.

POT is a modern sampling model, largely motivated by research on working memory. Cognitive models of working memory inform the study of intelligence because scores on working memory tasks are strongly correlated with scores on IQ tests (Kane, Hambrick, & Conway, 2005). This means that working memory tasks and IQ tests measure something similar—either the same ability or the same set of abilities. Cognitive models of working memory suggest the latter. This is because working memory is a multi-component system, consisting of domain-general processes involved in cognitive control and domain-specific processes involved in maintenance of information. Working memory capacity, as measured by cognitive task performance, is therefore determined by multiple processes. Indeed, measures of working memory capacity, such as complex span tests, require parallel storage and processing and tap domain-general cognitive control processes as well as domain-specific storage processes. Complex span tests are therefore different than so-called simple span tests, such as digit span, in which subjects simply have to recall a list of items.

Indeed, in contrast to simple span tests, variance in complex span tests is primarily domain-general (Kane et al., 2004). Therefore, similar to IQ tests, a general factor of working memory capacity can be extracted, and this factor correlates strongly with fluid intelligence: two meta-analyses of latent variable studies investigating the relationship between working memory capacity and fluid intelligence estimate that the correlation is somewhere between $r = .72$ to $r = .81$ (Kane et al., 2005; Oberauer, Schulze, Wilhelm, & Süss, 2005).

Furthermore, the processes that working memory tasks measure beyond storage most likely reflect individual differences in the executive attention (a.k.a. cognitive control) component of working memory (Engle & Kane, 2004; Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2002). According to the executive attention theory of individual differences in working memory capacity (Engle & Kane, 2004; Kane et al., 2001), working memory and fluid intelligence correlate strongly because both constructs rely to a great extent on executive functions, such as updating, inhibition, and task-switching. Indeed, several recent latent variable studies have demonstrated strong correlations between executive attention and fluid intelligence (Engelhardt et al., 2016; Shipstead, Lindsey, Marshall, & Engle, 2014; Unsworth, Fukuda, Awh, & Vogel, 2014).

POT is also influenced by research on goal neglect (Duncan, Emslie, Williams, Johnson, & Freer, 1996; Duncan et al., 2008) as well as work on the neural underpinnings of fluid intelligence (Burgess, Gray, Conway, & Braver, 2011; Kane, 2005; Tschentscher, Mitchell, Duncan, Unit, & Sciences, 2017; Woolgar et al., 2010). The latter studies point to a fronto-parietal network that correlates with performance on tests of fluid intelligence but is also involved in diverse cognitive tasks (Duncan, 2010; Duncan & Owen, 2000).

A main premise of POT is that a battery of IQ tests requires a number of domain-general processes, such as those involved in working memory and executive attention, as well as a number of domain-specific processes. Importantly, domain-general processes are thought to be required by the majority (but not all) of test items, whereas domain-specific processes are thought to be required less frequently, depending on the nature of the test (e.g., verbal vs. spatial). Therefore, domain-general processes associated with working memory and executive attention will constrain performance on most items on most IQ tests, whereas domain-specific processes will impact a narrower range of tests. Such a pattern of overlapping processes explains the positive manifold and thus the general factor as well as the domain-specific clusters of correlated tests that result in group factors.

Since POT is a sampling model, it is necessarily similar to Thomson's sampling model (Thomson, 1916), but is also different in crucial ways (Kovacs & Conway, 2016a). The most important and novel aspect of POT, and its main divergence from Thomson, is that it proposes that the processes involved in IQ test performance are non-additive. This is because individual differences in executive processes pose general limits on total performance, acting as a bottleneck, and masking individual differences in more domain-specific processes. That is, insufficient executive processing is likely to be the cause of failure on various test items measuring different content regardless of the specific abilities that are also measured. According to POT, executive performance cannot be compensated by domain-specific performance. In contrast, in Thomson's original account, performance on any test is simply the sum of the hypothetical bonds involved. Therefore, Thomson's model is fully compensatory and the correlation between two tests is simply a linear function of the ratio of overlapping processes. The unique aspect of POT is that it is compensatory within domains but non-compensatory across domains. This feature of POT is formalized in a multidimensional item-response model, which we discuss in more detail below.

Besides providing an account of the positive manifold, POT also explains a number of important phenomena observed in the study of human intelligence. The first such phenomenon is ability differentiation, which refers to the finding that cross-domain correlations are higher in samples with lower average ability and so *g* explains more variance in such samples. The second is that the more complex a task the higher its correlation with *g*. Finally, through proposing that the positive manifold is caused by the overlapping activation of the executive attention processes that are involved in both working memory and fluid reasoning, the theory accounts for the central role of fluid reasoning in the structure of human abilities and for the finding that the fluid reasoning factor (Gf) seems to be statistically identical or near-identical to *g* (Gustafsson, 1984).

POT is therefore able to explain why *g* is both population and task-dependent (i.e., it explains the most variance in populations with lower ability and cognitively demanding tasks). POT focuses on the limitations of executive attention processes in explaining *g* and proposes an interaction between the executive demands of the task and the executive functioning of the individual. This is expressed in a formal mathematical model that specifies the probability of arriving at a correct answer on a given mental test item as the function of the level of domain-specific as well as domain-general cognitive processes (Kovacs & Conway, 2016a).

POT-I: A Multidimensional Item Response Model

POT is mathematically formalized as a multidimensional item-response model, which we refer to as POT-I, and is expressed here in Eq. (1):

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

According to POT-I, the probability of an individual 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, which are both domain-general and domain-specific. More formally, Θ_{plm} is the process score (ability) for the p^{th} individual on the m^{th} process in the i^{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 tapped by an item and C is the number of processes in a given domain tapped by an item. Finally, POT-I is compensatory within domains and non-compensatory across domains. This is achieved in Eq. (1) by taking the sum of process scores within domains (compensatory) and multiplying the sum process scores across domains (non-compensatory).

Thus, according to POT-I, overall test performance reflects multiple domain-general abilities and multiple domain-specific abilities (expressed in Eq. (1) as process scores). This differs from the standard view, motivated by factor analysis, that test performance reflects a single domain-general ability and a single domain-specific ability. As stated earlier, a further claim of POT is that some tests are more dependent on domain-general processes while other tests are more dependent on domain-specific processes. This is precisely the reason why POT is able to account for the hierarchical structure of intelligence; it allows for distinct specific ability factors that vary with respect to their relationship to a higher-order general factor.

In a recent simulation study, fictional IQ test scores were calculated based on POT-I (Kovacs, Conway, Snijder, & Hao, 2018). In order to demonstrate a higher-order *g* factor, we assumed the presence of 3 group factors representing specific cognitive abilities: verbal ability, visual-spatial ability, and fluid ability. Test scores were simulated for 3 tests of each group factor, for a total of 9 tests. Each test consisted of 100 items. The sample size for each simulation was set at $N = 400$, and 1000 iterations of the simulation conducted. Each item on each test required both domain-general and domain-specific cognitive processes, consistent with POT-I. However, in tests of fluid ability, the probability of a domain-general process being required was nearly double the probability of a domain-specific process being required. For tests of verbal and visual-spatial ability, the opposite was true; the probability of a domain-specific-process being required was nearly double the probability of a domain-general process being required.

For each individual subject, a set of 200 cognitive process scores was generated. These cognitive processes can be regarded as the potential processes that could be involved in the performance of the 9 tasks. We assumed 4 different kinds of cognitive processes (50 of each kind): Executive function processes, fluid reasoning processes, verbal processes, and visual-spatial processes. While executive function processes could be involved in all three types of tasks (fluid tasks, verbal tasks, and visuospatial tasks), the other three types of processes could only be involved in their corresponding type of tasks. The values for the process scores were drawn from a multivariate standard normal distribution, which resulted in a 400×200 matrix (N $\times$ processes) of orthogonal process scores.

The IQ-test scores for the 9 tasks were analyzed using structural equation modeling and the result is presented in Figure 2. We observed a higher-order *g* factor, as well as specific group factors. This model fit the data well. We were therefore able to demonstrate the presence of a higher-order *g* factor in the absence of a general ability parameter. Furthermore, the fluid ability factor was more strongly correlated with the *g* factor than were the other factors, which is consistent with most latent variable models of intelligence.

The simulation demonstrates that a *g* factor can emerge even if the data are generated without the influence of a general mechanism involved in all cognitive performance. Moreover, a standard factor model, with a causal *g* accounting for the correlation between specific abilities, fit the data well; the structural model *appears as if* there were a causal mechanism at play, even though that is not the case. Therefore, the results of this simulation mean that being able to fit a *g*-model to any given data set does not prove that the data reflect the causal influence of a general cognitive ability.

POT-S: A Structural Model

POT can also be expressed as a structural latent variable model (POT-S; Kovacs & Conway, 2016a). As mentioned above, POT interprets *g* as equivalent to the positive manifold, and so POT-S includes a general factor. However, according to POT, *g* as a construct is different from the general factors that typically appear in factor models of cognitive ability. In most psychometric models of intelligence, *g* is specified as a reflective latent variable. That is, in such models *g* is a construct that is reflected

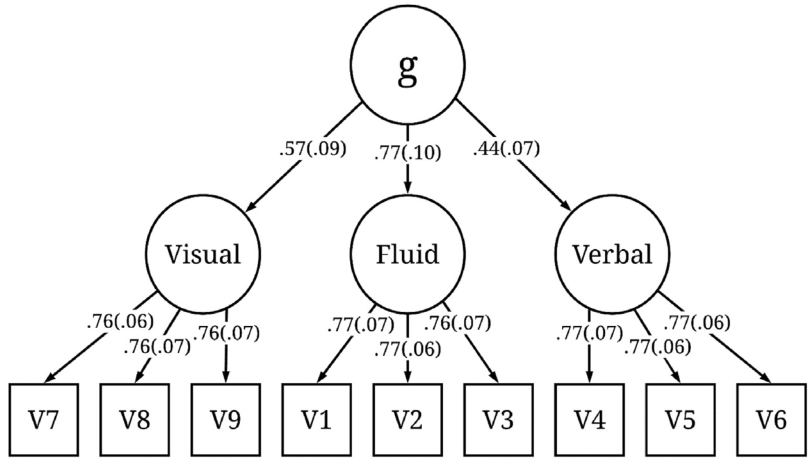

Figure 2. A hierarchical latent variable model of intelligence generated in a simulation study based on POT-I (Kovacs et al., 2018). The values are the mean standardized factor loadings (and standard deviations) across 1000 iterations of the simulation.

by its indicators, the cognitive ability tests, and is also the common cause of covariance between specific cognitive abilities.

Unfortunately, being able to fit a reflective latent variable model does not, in itself, mean that the reflectively modeled variable corresponds to a real entity. We have seen that it is possible to fit a reflective model, such as the one in Figure 2, to data generated without a causal mechanism involved in all tests. In other words, a reflective model can fit even if, in fact, there is nothing to reflect.

However, positing a reflective latent variable one implicitly subscribes to a number of strong theoretical positions, even if these positions are not explicitly articulated. Most importantly, depicting a variable as reflective implies a realist ontology with regard to the variable (Borsboom et al., 2003). In other words, reflective g is real, in both a psychometric and psychological sense. Among other things, this means that it would exist even if no one ever attempted to measure it. Importantly, the variables that reflect the latent construct are replaceable as long as each new variable also correlates with the latent construct.

In a reflective latent variable model, all tests x_i reflect an underlying construct η (the common cause) plus measurement error:

$$x_i = \lambda_i \eta + \varepsilon_i \tag{2}$$

Here x_i is the ith indicator of the latent variable η, λ_i is the loading of x_i on η, and ε_i is the measurement error of the ith indicator. Thus, in reflective models the tests are the dependent variables.

Yet if there is no general ability reflected by test scores (that is, if there is no psychological equivalent of the g factor), then this kind of model is inappropriate from a substantive perspective (Borsboom et al., 2003). According to POT, g is the common consequence rather than the common cause of the positive manifold. Hence g cannot be interpreted as a psychological construct or indeed as a "thing" of any kind, but as the consequence of a set of overlapping cognitive processes sampled by a battery of tests. Therefore, POT considers g to be a formative construct (see Figure 3).

Formative variables drastically differ from reflective ones. They are the result of measurement, without which they would not exist. This also means that the indicators of a formative construct have great importance since, unlike in a reflective model, they are not freely replaceable without altering the nature of the construct. A formative model can be formalized as follows:

$$\eta = \sum_{i=1}^{n} \gamma_i x_i + \zeta \tag{3}$$

Here γ_i represents the effect of indicator x_i on the latent variable η, and ζ represents all remaining causes of the construct that are not represented by the indicators. Technically, this is a multiple regression equation where the formative construct is the dependent variable.

Thus, according to POT-S, the direction of causality is the opposite as in traditional models of g: g does not cause individual differences in cognitive performance; rather, it is the outcome of individual differences in cognitive performance. A similar approach to intelligence is the mutualism model (van der Maas et al., 2006), based on which it has been proposed that "intelligence is what the intelligence test measures. Seriously" (van der Maas, Kan, & Borsboom, 2014). This is because a *formative g* can be thought of as a kind of global index of mental functioning.

The difference between formative and reflective models is easiest to comprehend through examples. An example where a reflective latent model is perfectly appropriate is the measurement of body temperature (Cramer et al., 2012). Specifically, body temperature can be measured by a number of different methods (e.g., mercury-based, phase-change, liquid crystal, electronic), in a number of different places (e.g., under the tongue, in the rectum, in the ear, in the armpit). Different methods can yield substantially different results, as anyone who has ever measured the temperature of a child with more than one method knows all too well. But there is no doubt that differences between these results are due to the inaccuracy of instruments, and there clearly is a real temperature reflected by the measurements. Living children would have a body temperature even if no one ever measured it. Naturally, the instruments are replaceable, too; any thermometer can be replaced by another since all of them reflect the same construct.

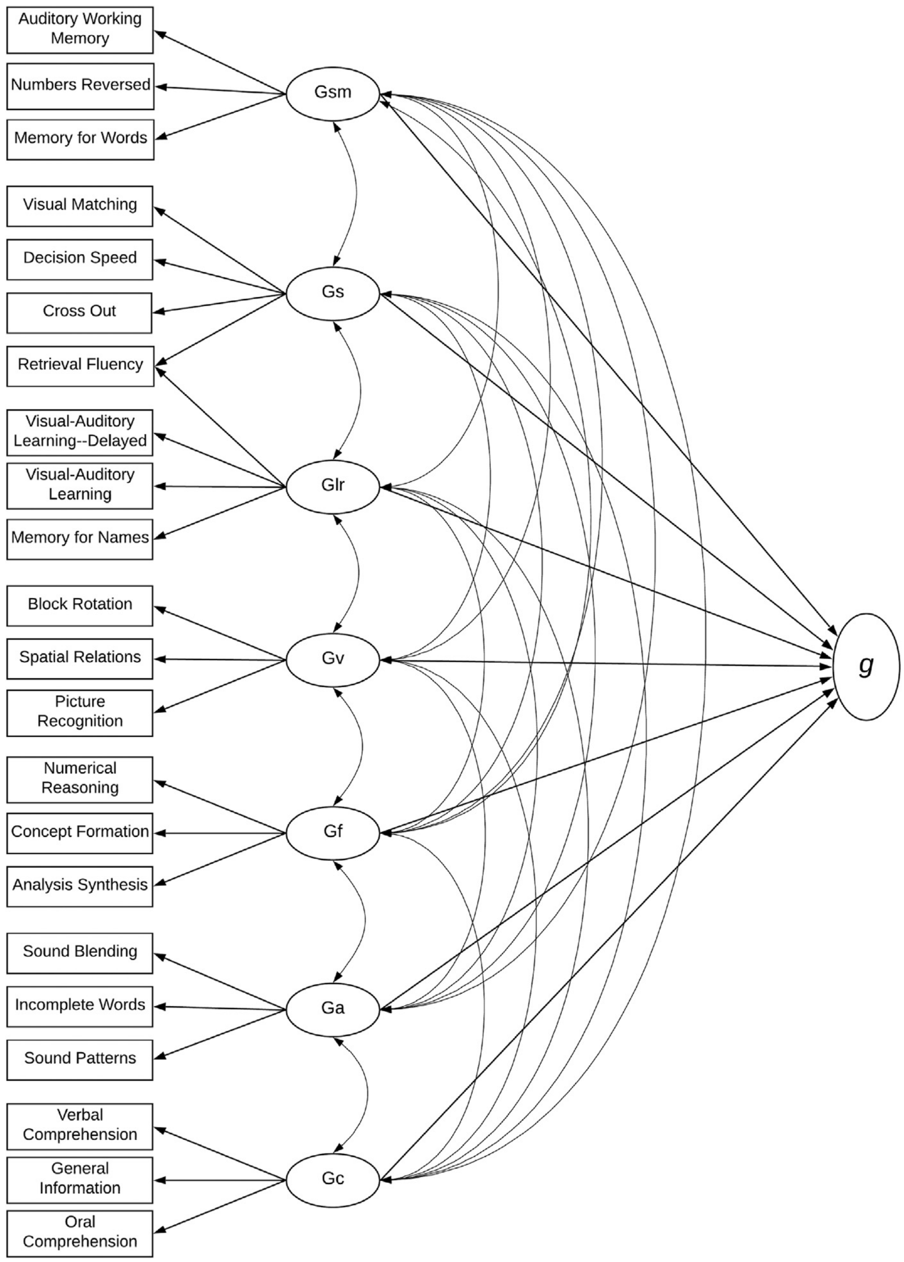

Figure 3. The factor structure of the Cattell–Horn–Carroll theory of intelligence, but with a formative *g* factor, to be consistent with POT. *Note*: Gsm = short-term memory; Gs = processing speed; Glr = long-term retrieval; Gv = visual-spatial thinking; Gf = fluid reasoning; Ga = auditory processing; Gc = crystallized intelligence; *g* = general intelligence. Single-headed arrows depict causality. Double-headed arrows depict correlation.
Source: Adapted from Taub and McGrew (2014).

Though technically not the result of a formative measurement model, Global Competitiveness Index (GCI) is a good illustration of a construct that only allows for a formative interpretation. GCI is a weighed sum score computed from different variables summarized in Table 1. GCI does not cause economies to be more competitive than others and it would not exist without the actual economic measures from which it is constructed. According to *g*-theory, *g* is like temperature and intelligence tests are like thermometers. In this approach *g* has a realist ontology, it would exist without measurement and the tests that measure *g* are replaceable: ceteris paribus, any two tests with an equal *g* loading are equally appropriate. According to POT, *g* is like GCI: a consequence of independent but correlated cognitive abilities.

Since POT is, in principle, agnostic about any particular modeling of first- or second-order cognitive abilities, and CHC is probably the most accepted model of the structure of abilities, we decided to illustrate this interpretation of *g* on the model already depicted in Figure 1. Figure 3 demonstrates a modified version of the model: the seven CHC abilities are considered to be reflective latent variables, while *g* is an emergent formative construct.

Table 1
"The Twelve Pillars of Competitiveness" that Constitute the Global Competitiveness Index (Artadi & Sala-i-Martin, 2004)

	#	Pillars
Basic requirements	1.	Institutions
	2.	Infrastructures
	3.	Macroeconomic stability
	4.	Personal security
	5a.	Basic human capital
Efficiency enhancers	5b.	Advanced human capital
	6.	Good market efficiency
	7.	Labor market efficiency
	8.	Financial market efficiency
	9.	Technological readiness
	10.	Openness/market size
Innovation and sophistication factors	11.	Business sophistication
	12.	Innovation

POT-N: A Network Model

A more radical approach to modeling the structure of intelligence is to abandon latent variable models altogether and conduct psychometric network analyses instead. In network models it is possible to examine the covariance structure of observed variables and the estimated partial correlations among observed variables, without assuming latent common causes. In network science and network modeling, observations are referred to as *nodes* and the connections between pairs of nodes are referred to as *edges*. Recently, Kan et al. (2019) provided a description of the differences between traditional latent variable modeling and psychometric network analysis and proposed that psychometric network analysis lends itself well to theories of intelligence like POT (Kovacs & Conway, 2016a) and their own model of intelligence known as mutualism (van der Maas et al., 2006).

To illustrate the network modeling approach, we conducted a psychometric network analysis on the data from our previously discussed simulation study. As a reminder, in the simulation study we assumed three broad ability factors (spatial, verbal, and fluid), each measured by three tests (the latent variable model of the simulation data was presented in Figure 2). The resulting network model is presented in Figure 4. There are three important features of this model: (1) the three broad abilities are revealed, represented by three tight clusters of nodes; (2) there are no latent variables and therefore no *g*; (3) fluid ability is more central to the network than either spatial ability or verbal ability. In our view, these features make the network model more compatible with cognitive psychology and neuroscience interpretations of intelligence. For example, the centrality of fluid ability corresponds to the centrality of domain-general processes in cognition and the importance of the fronto-parietal network in neuroscience.

Implications for Intelligence Research

This new conceptualization of human cognitive abilities has implications both for research and application. With regard to research, it follows from a formative concept of *g* that "the most fruitful path [for researchers] would be to focus on those lower order variables that do allow for a realist, causal interpretation" (Kan, van der Maas, & Kievit, 2016, p. 220). That is, research on the genetic basis or neural correlates of intelligence should focus on reflective latent variables rather than on an emergent property. Naturally, neural correlates of *g* can be identified in imaging studies and genome-wide association studies are also able to identify correlates of *g*. There is nothing in principle that would prevent finding such correlates, since neural or genetic correlates of clearly formative variables such as socioeconomic status can also be identified. But since a formative *g* does not exert a causal influence on cognition in general and IQ test scores in particular, a scientific understanding of intelligence can only be achieved through understanding the effect of reflective variables.

This is also the case because a formative construct is largely dependent on the actual instruments that are indicators of the construct. For instance, neuroimaging studies have found that *g*-factors obtained from different test batteries show different neural correlates even though the *g*-factors extracted from those batteries are statistically identical (Colom, Jung, & Haier, 2006). At the same time, fluid intelligence has a consistent pattern of neural correlates in the dorsolateral and partly in the posterior parietal cortex (Jung & Haier, 2007; Kane, 2005). This is exactly what one would expect if *g* is formative and Gf is reflective. Moreover, fluid intelligence is one of the few constructs that seem to demonstrate ergodicity: the neural correlates of task difficulty and the neural correlates of differences in performance are nearly identical (Kievit, 2014). Finally, research on the evolution of intelligence is informed by a formative approach to *g* and a reflective one to CHC abilities; it follows that *g* does not cross species. It matters, for instance, if *g* in one species has a verbal component but in another it does not, hence dog *g*, for instance, is not identical to human *g*. At the same time, it is logically possible that the factor Gf represents individual differences in the construct fluid intelligence which, since both reflective and ergodic, represents the ability to solve novel problems. Such an ability, then, can be meaningfully interpreted across species (Kovacs & Conway, 2017), provided that the differences between the theoretical status of intra-individual and inter-individual processes is not ignored (Penke et al., 2011).

Implications for IQ Testing

The difference between a reflective and a formative interpretation of *g* has consequences for the practice of IQ-testing, too. First, in a formative approach indicators are not automatically interchangeable, so in a formative approach the content of IQ-tests has greater relevance. For instance, it has been argued that a focus on *g* in intelligence research has prevented the study of a number of important constructs, such as economic decision-making (Stankov, 2017). Future research might therefore broaden current conceptions on intelligence and include measures of additional constructs to existing batteries. Relatedly, a cultural approach to intelligence, which emphasizes the role of cultural values in which abilities are regarded as intelligence (Sternberg & Grigorenko, 2004) is more in accordance with a formative approach.

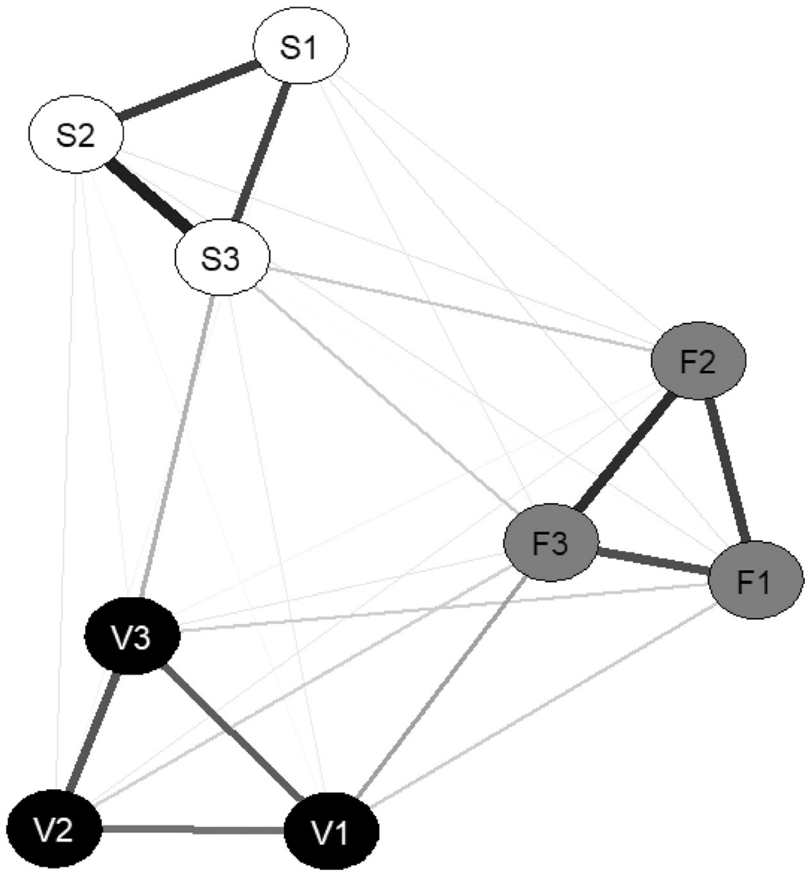

Figure 4. A network model of intelligence based on simulated test scores from 9 tests. The circles represent measures of spatial ability (S), fluid ability (F), and verbal ability (V).

Second, in *g*-theory variation in lower-order abilities has much lower importance than variation in *g*, and since in *g*-theory IQ is usually interpreted as a proxy for *g*, global IQ scores are more important than specific abilities. In a formative approach, on the other hand, the first-order factors that allow for a realist interpretation—for instance, CHC abilities—have greater significance. Therefore, a detailed assessment of specific abilities is more informative than a global index of mental functioning. This can be illustrated with the example of global competitiveness index: even though it is useful for global comparisons of different economies, a thorough understanding of any given economy will not be provided by the GCI; an evaluation of all the pillars that compose GCI is much more useful. Similarly, a global IQ score or a score on the general factor represent summary statistics that can be used to predict various phenomena, ranging from everyday cognitive performance (e.g., academic achievement and job performance) to non-cognitive life outcomes (e.g., socioeconomic status or longevity). As long as *g* factors obtained from different batteries are statistically equivalent, which seems to be the case (Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004) if the batteries are large and diverse enough (Major, Johnson, & Bouchard, 2011), for the purpose of pure statistical prediction the problem on instrument-dependence in formative models can be overcome.

There is a debate about the adequate level of interpretation of IQ-test results in general (Hale, Fiorello, Kavanagh, Hoeppner, & Gaither, 2001; Nelson, Canivez, & Watkins, 2013), and for clinical (Fiorello et al., 2010; Watkins, Glutting, & Lei, 2007) and school psychology (Bowman, Markham, & Roberts, 2001; Fiorello, Hale, McGrath, Ryan, & Quinn, 2001; Kranzler, 2001; Vanderwood, McGrew, Flanagan, & Keith, 2001) and for I/O psychology (Schmidt, 2002) in particular. It is certainly beyond the aims or possibilities of this paper to review the arguments in each of these debates. It is sufficient to point out that generally there are three possible levels of the evaluation of performance in cognitive test batteries that consist of a number of subtests: (a) the level of subtests, (b) the specific factors, or (c) global IQ scores or scores on a general factor.

The level of subtests is arguably too specific, hence scores contain too much task-specific variance as well as, inevitably, measurement error, which means that such scores are less reliable than more global scores. The problem with a global score such as IQ is that, if a formative approach is correct, they do not reflect a real ability. Naturally, an IQ score can still be useful for a number of reasons, as already discussed, as long as it is interpreted as index of different abilities rather than the reflection of some general intelligence. But the optimal level of interpretation is the level of specific abilities (e.g., fluid, spatial, verbal) that are global enough not to be task-specific but still allow for a reflective interpretation.

Overall, from a formative perspective in general, or from the perspective of POT in particular, it follows that a good measure of IQ is a large enough battery that is also able to provide a profile-type assessment that highlights individual strengths and weaknesses. Such a profile is especially more informative than an overall IQ if there is substantial discrepancy between the score on specific abilities. Generally, in educational settings, especially with a focus on identifying individual strengths in students, a profile-based assessment seems more appropriate and also more in line with the theoretical approach of POT. This is especially the case when testing with the purpose of identifying giftedness since because of the phenomenon of ability differentiation (outlined in the introduction) correlations between different tests are lowest at highest levels of ability. Therefore an overall IQ score might mask individual strength in a specific area. Or, if for some reason a single measure is sufficient, the best candidate to predict overall IQ is fluid intelligence because of its central role in the structure of abilities and its near-identity with *g*.

In a large number of studies, clinical or otherwise, researchers control for IQ, but IQ is measured different ways in different studies. This practice is only without problems as long as IQ scores obtained on different batteries are interchangeable. If, however, different IQ batteries do not measure the exact same thing (as indicated in a formative approach) but rather possibly different things (as in a formative approach) then different "IQs" might reflect different processes. The solution is to use batteries that are large and diverse enough so that they tap the same factor (Johnson et al., 2004)—or, under POT's framework, the same overlapping processes.

The use of overall IQ scores can become particularly problematic when they are used for personnel selection. In the I/O literature, it well established that *g* is the single best predictor of job performance (Schmidt & Hunter, 2004). However, I/O psychologists tend to interpret *g* as a "general mental ability," and often refer to it as "the ability to learn" (Schmidt, 2002). According to POT, this interpretation of *g* is incorrect because there is no such thing as general mental ability. When it comes to personnel selection, the incorrect interpretation of *g* can have legal consequences. For example, in the United States, employers may be required to provide evidence that whatever is measured by the IQ test is also required to do the job. If the employer relies on an overall IQ score, derived from a large battery of tests, then it may be difficult for them to link the score to job performance. It is possible, for example, for a person to have a below average IQ score but above average verbal ability. If the overall IQ score was used for personnel selection but only verbal ability is needed for the job then the employer could face legal action.

A better approach to personnel selection is to use measures of more specific cognitive abilities that can easily be linked to job performance, or to use tests of fluid intelligence, under the assumption that novel problem solving is common to most jobs. The over-arching goal should be to maximize predictive validity and minimize adverse impact. Adverse impact refers to the negative effect that a selection procedure has on a protected class of citizens. For example, in the United States, there is a group difference in overall IQ between whites and blacks. Therefore, if IQ is used for personnel selection then there is adverse impact, discrimination, against blacks. This is why is it essential to interpret scores correctly and to use tests that minimize adverse impact. This issue was made famous in the case of Griggs vs. Duke Power Co. (1971). Duke Power Co. had been using high school completion and an intelligence test to make hiring decisions. The court ruled that those criteria resulted in discrimination against African-Americans. In their decision they argued,

> On the record before us, neither the high school completion requirement nor the general intelligence test is shown to bear a demonstrable relationship to successful performance of the jobs for which it was used. Both were adopted, as the Court of Appeals noted, without meaningful study of their relationship to job performance ability.

A final consequence of a formative approach to *g* relates to training and improvement of cognitive abilities. Without taking sides on the actual efficiency of existing interventions and the general trainability of cognitive abilities, it is worth noting that if *g* is formative then the focus of such attempts should once again be the lower-order specific abilities that have a realist ontology—because, well, they exist. Again, the example of global competitiveness index illustrates the point. For if a government aims to improve the competitiveness of the country's economy it will not try to directly improve GCI itself, but rather it will target one or more of the 12 pillars that constitute the index, such as infrastructure, labor market efficiency, or innovation. Similarly, the target of cognitive interventions should be reflective factors like the special abilities of the CHC model (see Figure 3), not IQ per se.

Conclusions

We hope to have convinced you that the first sentence of this paper is meaningful; while most cognitive psychologists and neuroscientists today agree that there is no such thing as general intelligence, psychometricians have become remarkably good at measuring it. In contemporary cognitive psychology the dominant view on cognition is domain-specificity, which is based on ample empirical evidence on dissociations through brain injury, genetic disorders, and the localization of function through neuroimaging. Therefore equating a *g*-model with *g*-theory, that is, interpreting the general factor of intelligence as a general cognitive ability, is unlikely to be the basis of a unified cognitive/differential approach.

At the same time, instruments that are designed for differential psychology to measure individual differences generally greatly outperform trademark tasks in cognitive psychology that are designed for within-subject experimental studies that purport to discover or explain causal relations between cognitive processes. Psychometrics indeed has remarkably sophisticated methods to quantify variation in human cognitive abilities and as a formative variable *g* can be measured accurately and can be the basis of prediction.

We also hope to have presented food for thought about explaining the positive manifold in intelligence without a general cognitive ability as well as interpreting the general factor as a formative construct. This paper could only start to unravel all the consequences of such a shift of interpretation. Hopefully the unified approach presented here can serve as a meaningful basis for further research on cognitive abilities as well as for applied measurement that is in agreement with theoretical advances in human cognitive abilities.

Author Contributions

K.K. and A.C. both contributed to the writing of the manuscript with K.K. taking on the role of lead author. K.K. wrote the first draft and both K.K. and A.C. wrote revisions. A.C. conducted the statistical analyses.

Conflict of interest

The authors declare no conflict of interest.

References

Artadi, V., & Sala-i-Martin, X. (2004). *The global competitiveness index*. pp. 51–80. *World economic forum global competitiveness report* (Vol. 2005)

Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. http://dx.doi.org/10.1126/science.1736359

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. http://dx.doi.org/10.1016/S0079-7421(08)60452-1

Bishop, D. V. M., Aamodt-Leeper, G., Creswell, C., McGurk, R., & Skuse, D. H. (2001). Individual differences in cognitive planning on the Tower of Hanoi task: Neuropsychological maturity or measurement error? *Journal of Child Psychology and Psychiatry*, *42*, 551–556. http://dx.doi.org/10.1111/1469-7610.00749

Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In *Dynamic process methodology in the social and developmental sciences*. pp. 67–97. New York, NY: Springer US. http://dx.doi.org/10.1007/978-0-387-95922-1_4

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–219.

Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The reliability and internal validity of the Wisconsin Card Sorting Test. *Neuropsychological Rehabilitation*, *8*, 243–254. http://dx.doi.org/10.1080/713755573

Bowman, D. B., Markham, P. M., & Roberts, R. D. (2001).. *Expanding the frontier of human cognitive abilities: So much more than (plain) g! Learning and individual differences* (Vol. 13) http://dx.doi.org/10.1016/S1041-6080(02)00076-6

Brown, D. E. (Donald E.). (1991). *Human universals*. McGraw-Hill.

Burgess, G. C., Gray, J. R., Conway, A. R. A., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationship between fluid intelligence and working memory span. *Journal of Experimental Psychology. General*, *140*, 674–692. http://dx.doi.org/10.1037/a0024695

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386–404. http://dx.doi.org/10.1016/0022-0965(82)90054-6

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Colom, R., Jung, R. E., & Haier, R. J. (2006). Distributed brain sites for the g-factor of intelligence. *NeuroImage*, *31*, 1359–1365. http://dx.doi.org/10.1016/j.neuroimage.2006.01.006

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*, 769–786. http://dx.doi.org/10.3758/BF03196772

Conway, A. R. A., & Kovacs, K. (2013). Individual Differences in intelligence and working memory. *Psychology of Learning and Motivation*, *58*, 233–270. http://dx.doi.org/10.1016/B978-0-12-407237-4.00007-4

Cramer, A. O. J., Van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., . . ., & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, *26*, 414–431.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684. http://dx.doi.org/10.1037/h0043943

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466. http://dx.doi.org/10.1016/S0022-5371(80)90312-6

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews. Neuroscience*, *11*, 201–211. http://dx.doi.org/10.1038/nrn2793

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*, 13–21. http://dx.doi.org/10.1016/J.INTELL.2006.02.001

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*, 172–179. http://dx.doi.org/10.1016/j.tics.2010.01.004

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30*, 257–303. http://dx.doi.org/10.1006/cogp.1996.0008

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475–483. http://dx.doi.org/10.1016/S0166-2236(00)01633-7

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., . . ., & Nimmo-Smith, I. (2008). Goal neglect and Spearman's g: Competing parts of a complex task. *Journal of Experimental Psychology. General*, *137*, 131–148. http://dx.doi.org/10.1037/0096-3445.137.1.131

Engelhardt, L. E., Mann, F. D., Briley, D. A., Church, J. A., Harden, K. P., & Tucker-Drob, E. M. (2016). Strong genetic overlap between executive functions and intelligence. *Journal of Experimental Psychology. General*, *145*, 1141–1159. http://dx.doi.org/10.1037/xge0000195

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *The Psychology of Learning and Motivation*, *44*, 145–199.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology. General*, *128*, 309–331. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10513398

Fiorello, C. A., Hale, J. B., Holdnack, J. A., Kavanagh, J. A., Terrell, J., & Long, L. (2010). Interpreting intelligence test results for children with disabilities: Is global intelligence relevant? *Applied Neuropsychology*, *14*, 2–12. http://dx.doi.org/10.1080/09084280701280338

Fiorello, C. A., Hale, J. B., McGrath, M., Ryan, K., & Quinn, S. (2001). IQ interpretation for children with flat and variable test profiles. *Learning and Individual Differences*, *13*, 115–125. http://dx.doi.org/10.1016/S1041-6080(02)00075-4

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? *Psychological Science*, *15*, 373–378. http://dx.doi.org/10.1111/j.0956-7976.2004.00687.x

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*, 103–127. http://dx.doi.org/10.1080/09658219408258940

Gnys, J. A., & Willis, W. G. (1991). Validation of executive function tasks with young children. *Developmental Neuropsychology*, *7*, 487–501. http://dx.doi.org/10.1080/87565649109540507

Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1997). Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia. *Archives of General Psychiatry*, *54*, 159–165. http://dx.doi.org/10.1001/archpsyc.1997.01830140071013

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*, 13–23. http://dx.doi.org/10.1016/S0160-2896(97)90011-8

Griggs v. Duke Power Co., 401 U.S. 424 (1971).

Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179–203. http://dx.doi.org/10.1016/0160-2896(84)90008-4

Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Hoeppner, J.-A. B., & Gaither, R. A. (2001). WISC-III predictors of academic achievement for children with learning disabilities: Are global and factor scores comparable? *School Psychology Quarterly*, *16*, 31–55. http://dx.doi.org/10.1521/scpq.16.1.31.19158

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*, 1166–1186. http://dx.doi.org/10.3758/s13428-017-0935-1

Horn, J. L. (1994). Theory of fluid and crystallized intelligence. In R. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 443–451). New York: MacMillan Reference Library.

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

Hunt, E., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology*, *7*, 194–227. http://dx.doi.org/10.1016/0010-0285(75)90010-9

Jensen, A. R. (1998).. *The g factor: The Science of Mental Ability. East* (Vol. 17). Retrieved from http://www.jstor.org/stable/10.2307/4236447

Jensen, A. R. (2000). "The g factor" is about variance in mental abilities, not a cognitive theory of mental structure. Reply to Anderson on Jensen on Intelligence-g-Factor. *PSYCOLOQUY*, *11*.

Johnson, W. (2018). A tempest in a ladle: The debate about the roles of general and specific abilities in predicting important outcomes. *Journal of Intelligence*, *6*, 24. http://dx.doi.org/10.3390/jintelligence6020024

Johnson, W., Bouchard, T. J., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence*, *32*, 95–107. http://dx.doi.org/10.1016/S0160-2896(03)00062-X

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *The Behavioral and Brain Sciences*, *30*, 135–154. http://dx.doi.org/10.1017/S0140525X07001185, discussion 154–187

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

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

Kane, M. J. (2005). Full frontal fluidity. In O. Wilhelm, & R. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 141–165). Thousand Oaks, CA, US: Sage Publications.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*, 169–183. http://dx.doi.org/10.1037/0096-3445.130.2.169

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*, 637–671. http://dx.doi.org/10.3758/BF03196323

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71. http://dx.doi.org/10.1037/0033-2909.131.1.66, author reply 72–75

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*, 189–217. http://dx.doi.org/10.1037/0096-3445.133.2.189

Kievit, R. A. (2014). *Turtles all the way down? Psychometric approaches to the reduction problem*. University of Amsterdam.

Korkman, M., Kirk, U., & Kemp, S. (1998). . pp. 156–157. *A developmental NEuroPSYchological assessment (NEPSY)* (Vol. 64) San Antonio, TX: Psychological Corporation.

Kovacs, K. (in press). Process overlap theory: How the interplay between general and specific cognitive processes accounts for the positive manifold in intelligence. In D. J. McFarland (Ed.), *General and specific mental abilities*. Cambridge Scholars Publishing.

Kovacs, K., & Conway, A. R. A. (2016a). Has g gone to POT? *Psychological Inquiry*, *27*, 241–253. http://dx.doi.org/10.1080/1047840X.2016.1202744

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

Kovacs, K., & Conway, A. R. A. (2017). The evolution of fluid intelligence meets formative g. *Behavioral and Brain Sciences*, *40*, e208. http://dx.doi.org/10.1017/S0140525X16001680

Kovacs, K., & Conway, A. R. A. (2019). What is IQ? Life beyond "general intelligence". *Current Directions in Psychological Science*, http://dx.doi.org/10.1177/0963721419827275

Kovacs, K., Conway, A. R. A., Snijder, J., & Hao, H. (2018). General intelligence explained (away). In *Poster presented at: The 59th Annual Meeting of the Psychonomic Society*.

Kranzler, J. H. (2001). Commentary on "Is g a viable construct for school psychology?". *Learning and Individual Differences*, *13*, 189–195. http://dx.doi.org/10.1016/S1041-6080(02)00078-X

Krijnen, W. P. (2004). Positive loadings and factor correlations from positive covariance matrices. *Psychometrika*, *69*, 655–660. http://dx.doi.org/10.1007/BF02289861

Mackintosh, N. J. (2011). *IQ and human intelligence*. Oxford: Oxford University Press.

Major, J. T., Johnson, W., & Bouchard, T. J. (2011). The dependability of the general factor of intelligence: Why small, single-factor models do not adequately represent g. *Intelligence*, *39*, 418–433. http://dx.doi.org/10.1016/j.intell.2011.07.002

McFarland, D. J. (2017). Evaluation of multidimensional models of WAIS-IV subtest performance. *Clinical Neuropsychologist*, *31*, 1127–1140. http://dx.doi.org/10.1080/13854046.2017.1320426

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

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97. http://dx.doi.org/10.1037/h0043158

Molenaar, D., Kő, N., Rózsa, S., & Mészáros, A. (2017). Differentiation of cognitive abilities in the WAIS-IV at the item level. *Intelligence*, *65*, 48–59. http://dx.doi.org/10.1016/J.INTELL.2017.10.004

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*, 112–117. http://dx.doi.org/10.1111/j.1467-8721.2009.01619.x

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., . . ., & Sternberg, R. J. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77.

Nelson, J. M., Canivez, G. L., & Watkins, M. W. (2013). Structural and incremental validity of the Wechsler Adult Intelligence Scale-Fourth Edition with a clinical sample. *Psychological Assessment*, *25*, 618–630. http://dx.doi.org/10.1037/a0032086

Oberauer, K., Schulze, R., Wilhelm, O., & Süss, H.-M. (2005). Working memory and intelligence—Their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 61–65. http://dx.doi.org/10.1037/0033-2909.131.1.61, author reply 72–75

Pascal-Leone, J. (1970). A mathematical model for the transition rule in Piaget's development stages. *Acta Psychologica*, *32*, 301–345.

Penke, L., Borsboom, D., Johnson, W., Kievit, R. A., Ploeger, A., & Wicherts, J. M. (2011). Evolutionary psychology and intelligence research cannot be integrated the way Kanazawa (2010) suggested. *American Psychologist*, *66*, 916–917. http://dx.doi.org/10.1037/a0024626

Polloway, E. A. (2013). *The death penalty and intellectual disability*. American Association on Intellectual and Developmental Disabilities.

Ritchie, S. J. (2015). *Intelligence: All that matters*. London: Teach Yourself.

Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, *15*, 187–210. http://dx.doi.org/10.4324/9781410608802-12

Schmid, F. L., & Hunter, J. (2004). General Mental Ability in the World of Work: Occupational Attainment and Job Performance. *Journal of Personality and Social Psychology*, *86*, 162–173.

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116–141.

Spearman, C. (1904). "General Intelligence," objectively determined and measured. *The American Journal of Psychology*, *15*, 201–292. http://dx.doi.org/10.2307/1412107

Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: Macmillan.

Stankov, L. (2017). Overemphasized "g". *Journal of Intelligence*, *5*, 33. http://dx.doi.org/10.3390/jintelligence5040033

Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences*, *3*, 573–584. http://dx.doi.org/10.1017/S0140525X00006932

Sternberg, R. J., & Grigorenko, E. L. (2004). Intelligence and culture: How culture shapes what intelligence means, and the implications for a science of well-being. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *359*, 1427–1434. http://dx.doi.org/10.1098/rstb.2004.1514

Taub, G. E., & McGrew, K. S. (2014). The Woodcock–Johnson Tests of Cognitive Abilities III's Cognitive Performance Model. *Journal of Psychoeducational Assessment*, *32*, 187–201. http://dx.doi.org/10.1177/0734282913504808

Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology, 1904–1920*, *8*, 271–281. http://dx.doi.org/10.1111/j.2044-8295.1916.tb00133.x

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

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19–136). New York, NY: Oxford University Press.

Tschentscher, N., Mitchell, D., Duncan, J., Unit, B. S., & Sciences, B. (2017). *Fluid intelligence predicts novel rule implementation in a distributed frontoparietal control network*. http://dx.doi.org/10.1523/JNEUROSCI.2478-16.2017

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26.

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*, 842–861. http://dx.doi.org/10.1037/0033-295X.113.4.842

van der Maas, H. L. J., Kan, K.-J., & Borsboom, D. (2014). Intelligence is what the intelligence test measures. Seriously. *Journal of Intelligence*, *2*, 12–15. http://dx.doi.org/10.3390/jintelligence2010012

Vanderwood, M. L., McGrew, K. S., Flanagan, D. P., & Keith, T. Z. (2001). The contribution of general and specific cognitive abilities to reading achievement. *Learning and Individual Differences*, *13*, 159–188. http://dx.doi.org/10.1016/S1041-6080(02)00077-8

Watkins, M. W., Glutting, J. J., & Lei, P. W. (2007). Validity of the full-scale IQ when there is significant variability among WISC-III and WISC-IV factor scores. *Applied Neuropsychology*, *14*, 13–20. http://dx.doi.org/10.1080/09084280701280353

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., . . ., & Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14899–14902. http://dx.doi.org/10.1073/pnas.1007928107

Received 31 March 2019;
received in revised form 15 May 2019;
accepted 26 May 2019
Available online 24 June 2019