Measurement of mental attention: Assessing a cognitive component underlying performance on standardized intelligence tests

Steven J. Howard[1,2], Janice Johnson[1] & Juan Pascual-Leone[1]

Abstract

Despite the widespread use of standardized IQ tests to measure human intelligence, problems with such measures have led some to suggest that better indices may derive from measurement of cognitive processes underlying performance on IQ tests (e.g., working memory capacity). However, measures from both approaches may exhibit performance biases in favour of majority groups, due to the influence of prior learning and experience. Mental attentional (*M*-) capacity is proposed to be a causal factor underlying developmental growth in working memory. Measures of *M*-capacity index important cognitive variance underlying performance on standardized intelligence tests. These measures appear to be reasonably culture-fair and invariant across content domains. The current study tested theoretical predictions regarding the content-invariance of *M*-measures and the development of *M*-capacity for groups of children differing in performance on standardized IQ tests. Ninety-one participants differentiated on the basis of academic stream (intellectually gifted vs. mainstream) and age (grade 4 vs. grade 8) received measures of *M*-capacity in the verbal and visuo-spatial domains. Children identified as gifted scored about one stage higher on both measures. Results suggest that measures of *M*-capacity may be useful adjuncts to standardized intelligence measures.

Key words: mental attention, working memory, intelligence, IQ, giftedness

---

[1] *Correspondence concerning this article should be addressed to:* Steven Howard, PhD, Faculty of Education, University of Wollongong, Wollongong, New South Wales, 2522, Australia; email: stevenh@uow.edu.au

[2] Interdisciplinary Educational Research Institute and Faculty of Education, University of Wollongong, Australia

Development of the IQ test to measure human intelligence has been lauded as one of the greatest achievements in the history of psychology (Nisbett et al., 2012). Advocates of IQ testing point to evidence that IQ scores in childhood are predictive of length of schooling (Neisser et al., 1996), academic success (Brody, 1997; Deary, Strand, Smith, & Fernandes, 2007; Gottfredson, 2004; Neisser et al., 1996; Nisbett et al., 2012), socio-economic and vocational success (Firkowska-Mankiewicz, 2011; Gottfredson, 2004; Neisser et al., 1996; Schmidt & Hunter, 1998, 2004; Strenze, 2007), and even cognitive declines in late adulthood (Bourne, Fox, Deary, & Whalley, 2007). Intellectually precocious children, as identified by exceptionally high scores on standardized intelligence tests, display heightened performance in areas such as mathematics (Hoard, Geary, Byrd-Craven, & Nugent, 2007), speed and efficiency of cognitive processing (Jausovec, 1998; Johnson, Im-Bolter, & Pascual-Leone, 2003; Saccuzzo, Johnson, & Guertin, 1994), and resistance to interfering stimuli (Johnson et al., 2003). On the strength of these findings, IQ measures have been widely adopted for selection, placement, and decision-making in educational, vocational, clinical, and research settings (Richardson, 2002; Weinberg, 1989).

From this perspective, intelligence is viewed as a cognitive trait that can be reliably measured by IQ tests to yield scores that are related (perhaps causally) to superior cognitive performances and achievements across the lifespan. It has been argued that this cognitive trait is highly stable and largely resistant to meaningful long-term change (Herrnstein & Murray, 1994; Murray, 1996; Rushton, 1995). Numerous studies have demonstrated, however, both the short-term and long-term malleability of intelligence (as measured by IQ tests), as in the case of increased IQ scores after adoption into a more affluent family (Capron & Duyme, 1989; Duyme, Dumaret, & Tomkiewicz, 1999; van Ijzendoorn, Juffer, & Poelhuis, 2005), initial IQ gains and occasional later regression after cognitive training (Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Klingberg, Forssberg, & Westerberg, 2002; Mackey, Hill, Stone, & Bunge, 2011; Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005; Wasik, Ramey, Bryant, & Sparling, 1990), change in IQ as a result of various non-cognitive interventions (e.g., nutritional changes, curing infection, increasing motivation; Duckworth, Quinn, Lynam, Loeber, & Stouthamer-Loeber, 2011; Johnson, Swank, Howie, Baldwin, & Owen, 1996; Nokes & Bundy, 1994; Schoenthaler, Amos, Eysenck, Peritz, & Yudkin, 1991), and the rise and decline of IQ scores with continued or delayed/disrupted schooling, respectively (Baltes & Reinert, 1969; Bedard & Dhuey, 2006; Brinch & Galloway, 2012; Ceci, 1991; Ceci & Gilstrap, 2000).

There also is evidence of a reciprocal relationship between IQ and academic experience: IQ is a strong predictor of academic success (Brody, 1997; Deary et al., 2007; Gottfredson, 2004; Neisser et al., 1996; Nisbett et al., 2012), but length of schooling is a similarly strong predictor of later IQ scores (Ceci & Williams, 1997; Wahlsten, 1997). Although those espousing the stability of intelligence argue that these changes in IQ are evident "in modest amounts, inconsistently, and usually temporarily" (the 'fading' objection; Murray, 1996, p. S145), these results nevertheless question what actually is measured by IQ tests. If we maintain that IQ tests do in fact index intelligence, then we must also allow that an individual's 'true' level of intelligence can fluidly wax and wane over

the course of years, months, weeks, or even days. Conversely, if these changes in IQ scores do not reflect genuine changes in levels of intelligence, then this presents a challenge to the assertion that IQ tests are a valid and reliable measure of an individual's intelligence.

In line with the latter perspective, Richardson (2002) concluded that "the assertion that IQ measures human intelligence in any general sense, or that the source of variance in IQ scores is primarily cognitive in nature, remains unsubstantiated after decades of investigation" (p. 306). Rather than IQ tests measuring an inherent ability to learn, problem solve, reason, and/or plan (components of intelligence suggested by Gottfredson, 1997, 2004), Richardson (2002) contends that IQ scores represent a complex combination of learned cognitive and linguistic structures, cultural context, and interpersonal factors that outweigh any variance attributable to the cognitive processes contributing to the purported general intelligence ('g') factor.

In support of this claim, research has documented IQ gains as a result of cognitive and academic experience (e.g., schooling, cognitive training; Baltes & Reinert, 1969; Bedard & Dhuey, 2006; Brinch & Galloway, 2012; Campbell et al., 2001; Ceci, 1991; Ceci & Gilstrap, 2000; Jaeggi et al., 2008; Klingberg et al., 2002; Mackey et al., 2011; Rueda et al., 2005; Wasik et al., 1990). Further, it has been argued that this experience is culturally bound and unevenly distributed throughout the population (Pascual-Leone & Goodman, 1979; Pascual-Leone & Ijaz, 1991; Richardson, 2002), leading to IQ scores that consistently are biased toward the ethnic majority (e.g., Edwards & Fuller, 2005; Hart & Risley, 1995; Mercer, 1988; Moore, 1986; Neisser et al., 1996; Nisbett et al., 2012; Rushton & Jensen, 2010) and those from a higher socioeconomic status (e.g., Duyme et al., 1999; Neisser et al., 1996; Nisbett et al., 2012; van Ijzendoorn et al., 2005). This evidence suggests that the primary source of variability in IQ scores may not be strictly cognitive in nature, as commonly purported, but rather a product of learned socio-cognitive and affective factors. In fact, ongoing efforts to isolate the cognitive variance contributing to IQ scores (either statistically or using measures of specific components of intelligence, such as fluid intelligence) suggest at least tacit acknowledgement of the problematic nature of omnibus IQ tests as a measure of human intelligence.

Efforts to isolate the different components of intelligence commonly distinguish between crystallized intelligence (gC, referring to one's accumulated knowledge and learned operations that can be applied to problem situations) and fluid intelligence (gF, referring to the ability to reason abstractly and problem solve in novel situations; Cattell, 1943, 1963; Nisbett et al., 2012). This distinction suggests that criticisms of the cultural and socioeconomic bias of IQ measures may pertain uniquely to gC, whereas gF provides a less biased index of the cognitive abilities underlying human intelligence. However, even measures of fluid intelligence have been criticized as involving culturally determined means of problem solving. For instance, it has been argued that although the abstract symbols commonly found in measures of fluid intelligence (e.g., Ravens Progressive Matrices) are largely experience-free, the rules required to decipher the patterns of change across these symbols are more commonly encountered by the Western world's middle class (in the form of timetables, spreadsheets, etc.; Richardson, 2002). In addition, research suggests that measures of fluid intelligence involve a range of cognitive

abilities (e.g., working memory capacity, processing speed; Nisbett et al., 2012; Pascual-Leone & Johnson, 2005; Redick, Unsworth, Kelly, & Engle, 2012), highlighting the inability of these measures to provide clarity regarding the constituent cognitive abilities underlying intelligence. Many researchers have therefore sought to further isolate the specific cognitive determinants of *gF*.

The link between working memory and intelligence

Some have suggested that intelligence might be better captured by measuring the cognitive processes that underlie performance on measures of fluid intelligence. Working memory (WM), conceptualized as a memory system for the short-term maintenance and manipulation of information (Baddeley & Hitch, 1974; Engle, 2010), has been identified as one such process, whose measurement may be less subject to cultural bias and prior learning (Nisbett et al., 2012). In fact, WM has been found to be one of the primary factors assessed by commonly employed measures of fluid intelligence (e.g., Ravens Progressive Matrices; Redick et al., 2012; Swanson, 2008, 2011). Research also has shown that WM processing is related to many of the same cognitive abilities as *gF*, such as learning, reasoning, comprehension, and cognitive control (Conway et al., 2005; Cowan & Alloway, 2009; Kane, Conway, Hambrick, & Engle, 2007; Swanson & Beebe-Frankenberger, 2004). Moreover, WM capacity has been identified as a similarly strong predictor of academic achievement (Alloway & Alloway, 2010; Swanson, 2004; Swanson & Howell, 2001).

Measures of WM thus hold promise for capturing much of the important cognitive variance underlying performance on IQ tests, as well as unique cognitive variance not captured by these tests (Alloway & Alloway, 2010). Nevertheless, common measures of WM capacity may conflate underlying cognitive processes with literacy (e.g., sentence span), numeracy (e.g., counting span, operation span), culture, or prior learning and experience (Mainela-Arnold, Evans, & Coady, 2010; Ostrosky-Solis & Lozano, 2006). For example, typical complex WM span measures (e.g., Operation Span) require short-term storage of information (e.g., letters) while one is engaged in secondary processing (e.g., solving a mathematical operation). Proficiency with this 'secondary' processing may be influenced, however, by culturally based prior experience. Engle, Santos, and Gathercole (2008), for instance, found that young Brazilian children from low versus high socioeconomic groups differed on vocabulary measures, but not on verbal WM measures or on the Ravens measure of fluid intelligence. There was, however, a trend toward higher performance on the counting span measure in the high socioeconomic group, suggesting that prior practice in counting might have impacted performance. Thus, scores on WM measures may reflect the limits of mental attention (considered to be the maturational component underlying the development of WM), but also may be impacted by insufficient requisite knowledge and/or strategies, or the added cognitive demand of literacy/numeracy skills that have not yet been automated (Pascual-Leone, 2000). This conflation of knowledge and processing is particularly problematic for measurement with children from non-English speaking backgrounds or those with learn-

ing disabilities. Their relatively weaker English-language literacy and/or numeracy skills may result in lower performance on select WM tasks, leading to the conclusion that these populations may be characterized by WM deficits (e.g., Alloway, Gathercole, Adams, & Willis, 2005; de Jong, 1998; Gathercole, Alloway, Willis, & Adams, 2005; Siegel & Ryan, 1989).

In addition, there is no common metric across WM tasks. That is, WM theorists tend to think of WM capacity in terms of the number of 'chunks' that can be simultaneously maintained in mind. A chunk, however, is an ill-defined construct that can vary in size and complexity across individuals, tasks, and even individual items within a task (e.g., changing as the task progresses or with repeated administrations). This metric also fails to account for the mental demand of the multiple cognitive processes that need to be carried out in, for example, complex WM span tasks. For this reason, researchers typically do not expect equivalent scores across different measures of WM (Pascual-Leone & Johnson, 2011). Inconsistency across measures and lack of a common metric are limitations for using WM measures to index intelligence. Furthermore, although scores on WM tasks generally increase with development in childhood (e.g., Barrouillet & Camos, 2001; Cowan et al., 2010; Cowan, Nugent, Elliott, Ponomarev, & Saults, 1999; Dempster, 1981; Gathercole, Pickering, Ambridge, & Wearing, 2004; Towse, Hitch, & Hutton, 1998), most researchers do not make predictions regarding age-typical performance levels.

An alternative to WM capacity: The TCO model of mental attention

Measures of mental attentional (*M-*) capacity derived from the Theory of Constructive Operators (TCO) seek to overcome these limitations. The TCO proposes the existence of domain-free cognitive operators (i.e., brain resources) that select, activate, coordinate, and manipulate information-carrying schemes, which together serve to co-determine performance (Pascual-Leone, 1970, 1984; Pascual-Leone & Goodman, 1979; Pascual-Leone & Johnson, 2005, 2011). Scheme is the basic psychological unit within the TCO (Pascual-Leone & Johnson, 1991, 2011). Schemes carry information, of varying levels of complexity; neurologically, they correspond to collections of neurons that are co-functional and often co-activated (Pascual-Leone & Johnson, 2011), vis-à-vis a goal. However, performance tends to be over-determined, such that all schemes that are particularly relevant or salient (as a result of learning, situational cues, perceptual salience, affect, etc.) will actively compete to apply in a given situation. According to the TCO, domain-general constructive operators, under the control of a task executive or goal, modulate the degree to which schemes are activated. Of particular relevance here is the mental attentional (*M-*) operator, which serves to effortfully boost the activation of schemes that are relevant for the task, but are not otherwise sufficiently activated.

TCO's model of mental attention parallels (albeit predates) aspects of contemporary WM theory (e.g., Cowan, 2005). Both perspectives envision controlled effortful attention as a resource for boosting activation of cognitive information and processes (i.e., schemes for the TCO, chunks for WM). Further, both acknowledge limits to this attentional resource,

which manifest as a capacity constraint on the number of schemes/chunks that can be concurrently activated in mind. Both Pascual-Leone and Cowan also recognize the need to minimize the influence of prior learning and rehearsal strategies to accurately measure this capacity. For other approaches to measurement of WM (e.g., complex span tasks; Conway et al., 2005) control of prior learning and rehearsal appears less of a concern. However, the TCO uniquely proposes that the capacity of the M-operator (i.e., an individual's M-capacity) grows with age in childhood. At 3 to 4 years of age a child is able to boost with M just one symbolic scheme. The capacity of M grows by one unit every other year (e.g., two symbolic schemes at 5 to 6 years, three at 7 to 8, etc.) until it reaches seven units at 15 to 16 years of age (Pascual-Leone, 1970; Pascual-Leone & Baillargeon, 1994). Thus, unlike WM theory, there is a predicted course of development in M-capacity, a prediction supported by much cross-sectional data (e.g., Pascual-Leone & Johnson, 2005, 2011).

The TCO identifies four characteristics of good measures of M-capacity; these often are insufficiently addressed in construction of WM measures (Pascual-Leone & Johnson, 2011). First, good M-measures should minimize the executive demand of the task. This can be accomplished by providing executive pre-training, in order to ensure that all participants possess the required executive know-how to perform the task. Second, individual test items should index a single level of M-demand, while across the test, items should cover the full range of relevant M-demand levels (the M-demand of an item is the maximal amount of M-capacity needed to solve it). As such, M-measures typically are comprised of classes of items, where classes vary in M-demand, but within a class items share the same demand for M-capacity (Pascual-Leone & Johnson, 2011). Third, measurement of M-capacity should take place in a novel situation, to minimize variance due to prior learning.

Finally, M-measures should present suitably misleading situations. Misleading situations are those in which the schemes that are automatically elicited by the situation (due to salience, affect, over-learning, etc.) are task-inappropriate. According to the TCO, good performance in misleading situations requires effortful hyper-activation of task-appropriate schemes using mental attention, as well as inhibition of activated yet task-inappropriate schemes. To illustrate, consider the case of deciding whether to purchase a product on credit, for which considerations of sale price, immediate versus delayed gratification, relative value, affective response, current needs, and conflicting financial considerations might all come to bear. In many cases a heavily discounted sale price will outweigh other considerations (due to the salience of the discounted price sticker and/or an affective reaction to the sale), despite the fact that the credit interest can result in payments exceeding the full retail price. This illustrates that not all activated schemes are necessarily optimal for successful task performance (such as those that are particularly salient or over-learned, yet not optimally task-appropriate). Importantly, misleading situations ensure that all task-relevant schemes must be simultaneously hyper-activated by mental attention (rather than spontaneously activated by learning, affect, perceptual salience, etc.), making them particularly suited for measuring the capacity of mental attention (Arsalidou, Pascual-Leone, & Johnson, 2010; Pascual-Leone & Johnson, 2011; Pascual-Leone, Johnson, Baskind, Dworsky, & Severtson, 2000).

Because *M*-capacity is seen as a domain-general resource, it is possible to develop *M*-measures in different content domains. Such *M*-measures yield scores on an interval scale that corresponds to the power of the individual's *M*-capacity (Arsalidou et al., 2010; Pascual-Leone & Johnson, 2005, 2011). Thus, *M*-measures generate scores that have a clear theoretical meaning, have a common metric, and correspond to a predicted developmental trajectory. In addition, because task design minimizes other sources of activation, such as prior learning and affect, *M*-measures tend to be relatively culture-fair (Pascual-Leone & Johnson, 2005, 2011; Pascual-Leone et al., 2000).

In cross-sectional research, children's performance on *M*-measures has been shown to improve an average of one level (corresponding to one additional unit of information that can be concurrently coordinated by mental attention) for every other year of development, from 1 unit at around 3 years of age to 7 units at around 15 years of age (Morra, 1994; Pascual-Leone & Baillargeon, 1994; Pascual-Leone & Johnson, 2005, 2011). Furthermore, this developmental trajectory has been shown to be reasonably invariant across content domains (e.g., visuo-spatial, verbal, linguistic; Johnson & Pascual-Leone, 2011), cultures (e.g., Zulu-speaking children; Miller, Pascual-Leone, Campbell, & Juckes, 1989), and various other sample characteristics even when performance gaps are evident in WM scores (e.g., children from different socioeconomic classes; Globerson, 1983).

M-measures also correlate with standardized tests of ability and achievement (e.g., Bereiter & Scardamalia, 1979; Im-Bolter, Johnson, & Pascual-Leone, 2006). For instance, Pascual-Leone, Johnson, and Calvo (2004) found that in a large sample of grade 4 students, a measure of *M*-capacity correlated with scores on the Canadian Cognitive Abilities Test (CCAT; with correlations ranging from $r = 0.46$ on the verbal scale to $r = 0.61$ on the non-verbal subscale) and the Canadian Achievement Test (with correlations ranging from $r = 0.24$ for spelling to $r = 0.54$ for math). Subsequent investigation determined that correlation with CCAT composite score ($r = .59$) was not due to more efficient executive strategies, but rather was due to variance in the measured *M*-capacity of the participants. Similar results have also been found in cultural contexts outside of North America (e.g., Navarro et al., 2006). Further, children identified as cognitively gifted (based on performance in at least the 97th percentile on standardized intelligence tests) tend to perform an average of one level higher on *M*-measures than their non-gifted peers (corresponding to approximately 1 to 2 years of normal development; Johnson et al., 2003; Pascual-Leone & Johnson, 2011; Pascual-Leone, Johnson, Calvo, & Verilli, 2005). It may therefore be the case that heightened scores on *M*-measures relative to age norms is a characteristic of intellectual giftedness (Pascual-Leone & Johnson, 2011).

Taken together, these findings point to mental attention as a primary maturational component of WM and an important source of cognitive variance underlying performance on IQ measures. Further, measures of *M*-capacity may represent a more culture-fair means to capture this important cognitive variance. In light of these possibilities, the current study sought to test the TCO's theoretical predictions regarding the invariance of scores across *M*-measures and development of *M*-capacity in children characterized by differing levels of performance on standardized ability tests. Specifically, gifted and mainstream students' performance was examined across two common measures of *M*-capacity and at

differing levels of *M*-demand (defined as the maximal amount of *M*-capacity required to successfully solve the problem). It was expected that the gifted and mainstream groups would display similar patterns of performance across item classes, albeit with gifted students performing approximately one level higher than their age-matched mainstream peers. This superior performance also was expected to manifest in participants' *M*-scores (an index of their estimated *M*-capacity). Participating children were 9-10 versus 13-14 years of age. The former have a predicted *M*-capacity of four and the latter of six. Consistent with the TCO's theoretical predictions, it was expected that an age group would score above threshold on items whose *M*-demand is less than or equal to the group's predicted *M*-capacity, and below threshold on items whose demand is above the expected *M*-capacity.

Method

Participants

Participants were 91 elementary school students (45 girls and 46 boys) from a public school in the Greater Toronto Area of Canada. Students were differentiated on the basis of academic stream (gifted vs. mainstream) and age (grade 4 vs. grade 8), resulting in four groups: grade 4 mainstream (*n*=22), grade 4 gifted (*n*=28), grade 8 mainstream (*n*=22), and grade 8 gifted (*n*=19). Grades 4 (aged 9-10 years; $M = 9.81$, $SD = 0.33$) and 8 (aged 13-14 years; $M = 13.72$, $SD = 0.31$) were selected to capture two predicted stages in development of *M*-capacity. Gifted students were recruited from the school's gifted classes, placement in which requires a minimum achievement of 97th percentile on a standardized intelligence test. Mainstream students were recruited from the school's regular classes.

Measures

Figural Intersections Task.   The Figural Intersections Task (FIT; Pascual-Leone & Baillargeon, 1994; Pascual-Leone & Johnson, 2011) is a paper-based measure of *M*-capacity. Items consist of two to eight overlapping geometric shapes, and participants are required to locate the one area of common intersection among the shapes. The number of overlapping shapes in an item corresponds to the item's *M*-demand (i.e., the requisite *M*-capacity to be able to concurrently coordinate the shapes in mind) and defines the item class. The FIT version used had 36 items, with five items in each of classes 2 through 8, with the exception of item class 4 (which had six items). Items were presented in a set random order.

The FIT was administered in class groups, although each participant completed the task independently. Each group received instruction prior to starting the task; this involved executive pre-training and completion of eight practice items that were ordered to build an understanding of the task rules. For each item, participants first placed a dot in each

of a set of discrete shapes on the right side of the page. On the left side of the page these shapes (possibly changed in size or rotation) appeared in an overlapping configuration, and the participant was instructed to place a single dot in the area of common intersection. Some items included an irrelevant shape on the left (indicated as such by its absence on the right side of the page), which the participant was required to ignore when finding the answer.

The FIT *M*-capacity score corresponded to the highest item class that a participant accurately completed at least 80% of the time (provided that all lower classes also met this 80% threshold, with one lower class permitted to fall to 60% accuracy). The proportion of items correct within each *M*-demand class also was calculated.

Direction Following Task. The Direction Following Task (DFT; Cunning, 2003; Im-Bolter et al., 2006; Pascual-Leone & Johnson, 2011) is a linguistic measure of *M*-capacity. In this task, verbal instructions (i.e., "place X on Y") direct participants to place prescribed cutouts ('shapes') onto prescribed goal locations ('spaces'). These shapes (circles and squares) and spaces (squares) vary in terms of colour (blue, green, red, white, or yellow) and size (small or large). As the task progresses, verbal directions become increasingly difficult. For instance, an item in level 1 requires participants to "place a blue square on a white space," whereas an item in level 5 requires participants to "place a red square and a white circle on a small yellow space." The *M*-demand of each item is defined by the number of elements in the verbal instruction that must be simultaneously maintained in mind. It is assumed that participants can chunk into one unit a single shape and one property. Thus, for the example from level 5, the participant would be required to remember *red-square*, *white*, *circle*, *small*, and *yellow*, for an estimated *M*-demand of five.

The task consisted of five items at each of eight levels (presented in order of increasing complexity). However, in the current study level 1 was not administered, leaving 35 test items. Testing was preceded by executive pre-training and five practice items. Although all participants were required to complete the first 20 items (levels 2 to 5), a stop rule required a minimum of 40% accuracy on level 5 in order to proceed to levels 6 and 7. A minimum of 40% accuracy also was required on either level 6 or level 7 to progress to level 8. The task was administered individually, and a trained tester read the directions to be carried out. Shapes and spaces were covered each time a direction was provided and then immediately made available for the participant to carry out the instruction. The experimenter recorded participants' responses manually.

The DFT *M*-capacity score corresponded to the *M*-demand of the highest item class for which the participant achieved at least 60% accuracy (provided all lower classes also meet this 60% threshold, with one lower class allowed to fall to 40% accuracy). The proportion of items correct within each *M*-demand class also was computed. Note that two of the DFT item levels have the same predicted *M*-demand, leaving six distinct *M*-demand classes in the current task (*M*-demand of three through eight).

Procedure

The two measures of *M*-capacity were administered to all participants in the same order. The first measure (DFT) was administered in an individual testing session in an empty classroom within the school. The second measure (FIT) was administered in a group session in students' homeroom. Each testing session took about 40 minutes. Sessions were administered on separate days.

Results

We first examine FIT and DFT *M*-scores by stream (gifted vs. mainstream students) and grade (4 vs. 8) to test the TCO's theoretical predictions regarding the development of *M*-capacity in these groups. We then examine accuracy as a function of item *M*-demand, to investigate more closely the patterns of performance across the two tasks. Initial screening of the data indicated that the assumption of sphericity was consistently violated. Therefore, a Greenhouse-Geisser adjusted degrees of freedom ANOVA was conducted for all within-subjects main effects and interactions. In addition, in cases where extreme observations were evident (as identified by boxplots), analyses were conducted with and without these observations. Because patterns of results did not differ as a result of excluding outliers, all observations were included in the analyses reported below.

M-Scores

M-scores for the FIT and DFT were analyzed using a 2 (stream) x 2 (grade) x 2 (task) ANOVA. In line with theoretical predictions, it was expected that older children would outperform younger children and gifted students would outperform mainstream students, yet *M*-scores would not differ significantly across the two tasks. Mean *M*-scores are plotted in Figure 1.

There were main effects for stream, $F(1, 86) = 17.65$, $p < .001$, partial $\eta^2 = .17$; grade, $F(1, 86) = 43.86$, $p < .001$, partial $\eta^2 = .34$; and task, $F(1, 86) = 6.59$, $p = .012$, partial $\eta^2 = .07$. As predicted, gifted students ($M = 5.80$, $SD = 1.05$) outperformed their mainstream peers ($M = 5.10$, $SD = 1.12$), and grade 8 students ($M = 6.12$, $SD = 1.04$) outperformed grade 4 students ($M = 4.94$, $SD = 0.90$). Contrary to expectations, however, performance was higher on the FIT ($M = 5.65$, $SD = 1.57$) than the DFT ($M = 5.28$, $SD = 1.04$). This was conditioned by a Task x Grade interaction, $F(1, 86) = 6.20$, $p = .015$, partial $\eta^2 = .07$. *M*-scores on the FIT and DFT did not differ in grade 4 (FIT: $M = 4.96$, $SD = 1.40$; DFT: $M = 4.92$, $SD = 0.80$), $t(49) = -0.20$, $p = .841$, $\eta^2 < .01$; however FIT *M*-scores were higher than DFT *M*-scores in grade 8 (FIT: $M = 6.50$, $SD = 1.34$; DFT: $M = 5.71$, $SD = 1.13$), $t(39) = -3.46$, $p = .001$, $\eta^2 = .23$.

As illustrated in Figure 1, *M*-scores on both tasks were close to the theoretically predicted value of 4 for mainstream grade 4 students. Consistent with previous research on the
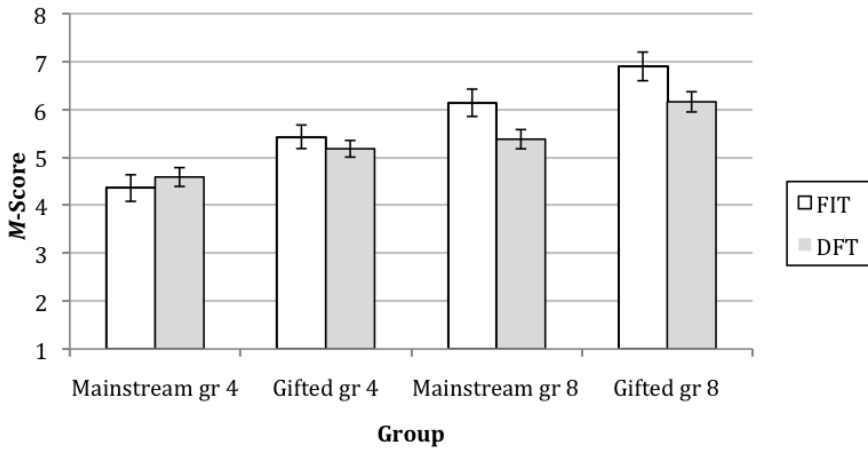
Figure 1:
Mean *M*-scores on FIT and DFT, by grade (4 vs. 8) and stream (gifted vs. mainstream).

TCO (Johnson et al., 2003; Pascual-Leone et al., 2004; Pascual-Leone & Johnson, 2005, 2011), gifted grade 4 students scored about one stage higher on both tasks (i.e., mean *M*-score of about 5). For the FIT, mean *M*-scores also were consistent with predictions at grade 8: Mainstream students scored at the predicted theoretical level of 6, and gifted students scored about one unit higher. On the DFT, however, grade 8 students underperformed relative to the TCO's theoretical predictions and their performance on the FIT.

Proportional accuracy by *M*-demand

To further examine patterns of performance on the two *M*-tasks, the proportion correct as a function of grade, stream, and item *M*-demand was analyzed separately for each task. *M*-demand values varied from three to eight, and performance on both tasks was expected to decrease in parallel as *M*-demand increased. For mainstream students, grade 4s were expected to perform well on items up to and including those with an *M*-demand of four, and grade 8s were expected to perform well on items with an *M*-demand of up to six. Gifted students were expected to maintain good performance on one *M*-demand class higher (i.e., classes five and seven, for grade 4 and 8, respectively).

Note that criteria for assigning *M*-scores for the FIT and DFT carry the expectation that performance may be lower on the DFT, when examined in terms of proportion pass. Recall that 80% constitutes passing of a FIT item level, whereas 60% is required to pass a DFT item level. These different criteria apply for two reasons: 1) participants have as much time as needed to complete FIT items, and the item stays visible throughout; in contrast, DFT items are read out only once, and if the participant's attention should drift during this time information will be irretrievably lost. 2) In the FIT there is the potential

for chance correct performance on items above the child's *M*-capacity, whereas chance correct performance on higher *M*-class DFT items is less likely. In the FIT, if the child places his or her dot in an area of high intersection of figures, there is increased probability of hitting the correct spot. There is no parallel strategy on the DFT. For these reasons, the pass criterion is more stringent on FIT than DFT.

FIT. Proportion correct scores on the FIT were analyzed using a Greenhouse-Geisser 2 (stream) x 2 (grade) x 6 (*M*-demand) ANOVA. Results indicated main effects for stream, $F(1, 86) = 15.28$, $p < .001$, partial $\eta^2 = .15$; grade, $F(1, 86) = 37.00$, $p < .001$, partial $\eta^2 = .30$; and *M*-demand, $F(3.36, 288.93) = 224.64$, $p < .001$, partial $\eta^2 = .72$. Overall, gifted students ($M = .77$, $SD = .15$) scored higher than mainstream students ($M = .64$, $SD = .14$), and grade 8 students ($M = .80$, $SD = .15$) outperformed grade 4 students ($M = .61$, $SD = .15$). Performance also decreased with increasing item *M*-demand (see Figure 2). As expected, main effects were conditioned by a number of interactions, because not all *M*-demand levels were expected to differentiate between grades and streams.

An *M*-demand x Stream interaction, $F(3.36, 288.93) = 4.23$, $p = .004$, partial $\eta^2 = .05$, indicated that gifted children were differentiated ($p < .01$) only at *M*-demand levels five through eight (η^2 ranged from .05 to .09 at these levels). Thus, consistent with theoretical predictions, a gifted advantage emerged only when item *M*-demand began to exceed the expected *M*-capacity of the mainstream grade 4 students (see Figure 2).

There also was an *M*-demand x Grade interaction, $F(3.36, 288.93) = 13.88$, $p < .001$, partial $\eta^2 = .14$, indicating that grade 8 students scored higher than grade 4 students at *M*-demand levels five through eight ($p < .05$). Effects sizes (η^2) were .02, .04, .06, .31, .24, and .17 for *M*-demand levels three through eight, respectively. Thus, effect sizes tended
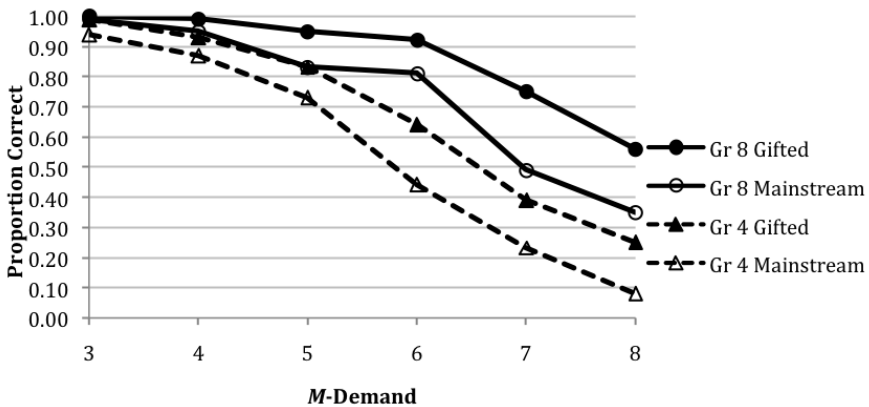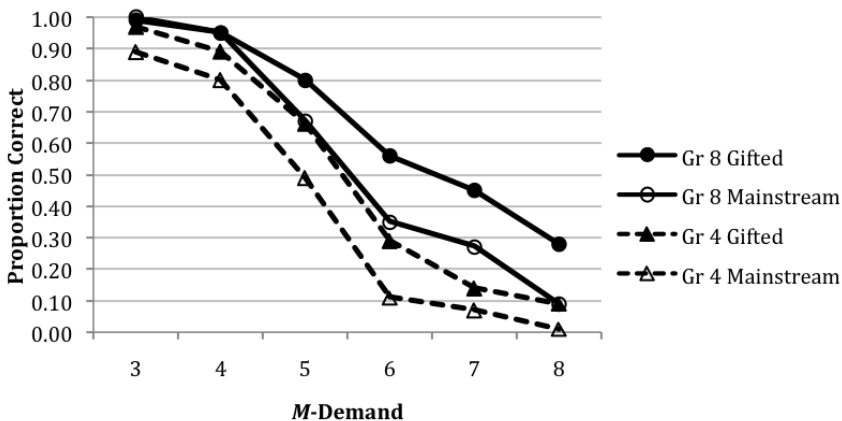
Figure 2:

Proportional accuracy scores on FIT, by grade (4 vs. 8), stream (gifted vs. mainstream), and item *M*-demand.

to be greater for higher *M*-demand items. As would be expected based on theoretical predictions, grade 8s maintained good performance at an *M*-demand of six, whereas grade 4 performance fell off sharply at this point.

DFT. A parallel analysis of the DFT also yielded main effects for stream, $F(1, 88) = 21.31$, $p < .001$, partial $\eta^2 = .20$; grade, $F(1, 88) = 38.91$, $p < .001$, partial $\eta^2 = .31$; and *M*-demand, $F(4.09, 359.52) = 479.05$, $p < .001$, partial $\eta^2 = .85$. Again gifted students ($M = .59$, $SD = .12$) outperformed mainstream students ($M = .47$, $SD = .12$), grade 8 students ($M = .61$, $SD = .12$) scored higher than grade 4 students ($M = .45$, $SD = .12$), and performance decreased with each increase in *M*-demand (see Figure 3).

Paralleling the FIT results, there was an *M*-demand x Stream interaction, $F(4.09, 359.52) = 4.06$, $p = .003$, partial $\eta^2 = .04$, which indicated that gifted children were differentiated ($p < .05$) beginning at an *M*-demand level of five (η^2 ranged from .04 to .10 at levels five through eight). Similar to the FIT analyses, and consistent with theoretical predictions, this gifted advantage was evident when item *M*-demand exceeded the expected *M*-capacity of the mainstream grade 4 students (see Figure 3).

There was an *M*-demand x Grade interaction, $F(4.09, 359.52) = 5.15$, $p < .001$, partial $\eta^2 = .06$, which indicated that grade 8 children scored higher ($p < .01$) at all *M*-demand levels (η^2 ranged from .06 to .21, with larger effect sizes for higher *M*-demand levels). Recall that 60% is the threshold for passing on the DFT. As would be expected based on theoretical predictions, grade 4 students maintained good performance at an *M*-demand of four. Grade 8 students scored well up to *M*-demand of five, but performance fell off sharply after that. Thus, grade 8 performance was lower than would be expected at *M*-demand of six. As also reflected in the *M*-scores, grade 8 students underperformed on the DFT relative to the TCO's theoretical predictions and their performance on the FIT.

Figure 3:

Proportional accuracy scores on DFT, by grade (4 vs. 8), stream (gifted vs. mainstream), and item *M*-demand.

There also was an *M*-demand x Grade x Stream interaction, $F(4.09, 359.52) = 2.41$, $p = .047$, partial $\eta^2 = .03$. As illustrated in Figure 3, the gifted and mainstream performance curves were fairly parallel for grade 4 students, with gifted showing a consistent advantage across *M*-demand levels. In contrast, the mainstream grade 8 students showed a sharp decline in performance from *M*-demand four to six, as compared with grade 8 gifted students.

Taken together, these results demonstrate that for both *M*-measures: (1) performance generally decreased as demand for *M*-capacity increased; (2) grade 8 students displayed higher proportional accuracy scores than grade 4 students for all item classes, with this performance gap increasing with increasing item difficulty; and (3) gifted students scored higher than mainstream in all but the lowest *M*-demand classes. However, (4) proportional accuracy scores were higher on the FIT than the DFT, except in the two lowest *M*-demand classes. This last finding is not entirely unexpected, because passing criteria are more stringent for FIT than for DFT. Proportion pass levels on the FIT were as predicted in both grades (see Figure 2). On the DFT, proportion pass levels were as predicted at grade 4, but fell below expectation at grade 8 (see Figure 3). These findings parallel those obtained using *M*-scores.

Discussion

The aim of the current study was to investigate the TCO's theoretical predictions regarding the invariance of *M*-measures and development of *M*-capacity among children characterized by their divergent performance on standardized intelligence measures. Previous research suggests that *M*-measures might represent a more culture-fair means to capture important cognitive variance underlying measures of human intelligence. The current study serves to extend these findings in its consideration of additional age groups of gifted and mainstream students, as well as looking at the specific patterns of performance (as a complement to *M*-scores) across two common *M*-measures.

Results provide support for the TCO's predictions regarding development of *M*-capacity (e.g., Pascual-Leone & Johnson, 2011). Children aged 9-10 years (grade 4) are predicted to have an *M*-capacity of 4 symbolic schemes, and children aged 13-14 years (grade 8) to have an *M*-capacity of 6. In line with these predictions, grade 4 mainstream students in the current study obtained *M*-scores of about 4, and grade 8 mainstream students obtained *M*-scores of about 6 (although the grade 8 group appeared to underperform on the DFT relative to theoretical predictions and the FIT; see Figure 1). A closer examination of participants' underlying patterns of performance similarly showed that grade 4 and grade 8 mainstream students' accuracy typically fell below threshold for item classes in which *M*-demand was predicted to exceed the *M*-capacity available at their age (i.e., class 5 items for grade 4 mainstream students and class 7 items for grade 8 mainstream students, although again grade 8s underperformed on the DFT).

Importantly, gifted students consistently outperformed their mainstream peers on both *M*-measures, with these students obtaining *M*-scores approximately one unit higher (corresponding to one to two years of normal development) than their mainstream peers. This pattern was manifest also in the proportion accuracy data on the FIT. Whereas accuracy decreased sharply on items beyond the *M*-demand predicted to be within the capacity of mainstream students, accuracy of gifted students remained at or just below threshold for one additional *M*-demand class (i.e., *M*-demands of five and seven for grade 4 and 8 gifted, respectively, see Figure 2). This pattern was mirrored in the DFT accuracy data for mainstream and gifted grade 4 students. However, for the DFT, grade 8 mainstream and gifted students remained at (or just below) threshold at *M*-demand of five and six, respectively – a pattern that theoretically would be predicted for grade 6 (11-12 years) students. Nevertheless, a gifted advantage remained in all the higher item classes of the DFT in the grade 8 group.

Thus, the results largely support predictions regarding both developmental and group differences. Results also indicate, however, underperformance of grade 8 students (both gifted and mainstream) on more complex items of the DFT relative to theoretical predictions and performance on the FIT. A possible explanation for this result may relate to the DFT's increased requirement for sustained attention. That is, later items of the DFT require the maintenance of more complex (and therefore longer) instructions in mind in order to subsequently carry out these instructions. For instance, a lower level item of the DFT requires participants to "place a yellow circle on a small green space" (*M*-demand of three), whereas an advanced item requires participants to "place a green circle on a yellow space and a blue square on a red space" (*M*-demand of six, including one scheme for keeping track of the prescribed order of placing shapes). Participants' attention may be more likely to wander during the longer items (which also appear in the later stages of the task), with a detrimental effect on their ability to maintain the entire instruction in mind. Note that attentional lapses would be less detrimental in the FIT, because FIT items remain perceptually available at all times. That is, if participants' attention wanders while performing the FIT, they are able to return their attention to the item and begin again. This is not the case for the DFT, because instructions are presented only once.

Note that this disparate requirement for sustained attention is factored into the thresholds for calculating *M*-scores (80% for FIT and 60% for DFT), perhaps explaining why *M*-scores for the two tasks were reasonably equivalent for the grade 4 students. It may therefore be the case that these different thresholds were sufficient to compensate for the sustained attention requirements of the shorter DFT items presented earlier in the task (those that grade 4 students were expected to complete successfully), yet insufficient for the longer items presented later in the task (many of which grade 8 students were expected to successfully complete). As such, the apparent underperformance of grade 8 students on the DFT may be a function of an increased demand for sustained attention conflicting with participants' decreasing attentional vigilance in the later stages of the task. If this is the case, it might argue in favor of revising the DFT so that items are presented in random order (as in the FIT). Research with adults has shown no difference in

performance for groups tested with graded versus random order of DFT items (Pascual-Leone & Johnson, 2011), however, further research with adolescents is warranted.

Although the current findings can be interpreted as supporting both the TCO model of mental attention and WM theories that posit a central role for controlled effortful attention, insofar as both see controlled effortful attention as an important source of cognitive variance underlying human intellect, it is unclear whether WM theory could generate these same results. That is, WM theories do not make precise predictions regarding the developmental course of increase in WM. They also lack task analytical methods that would allow accurate estimate of the mental demand of classes of items in a WM task. The TCO provides both, allowing one to examine the trade-off between task M-demand and participant M-capacity. This results in the ability to make fairly precise predictions about both overall M-scores and detailed performance patterns expected at different ages. In addition, M-measures are constructed to reflect a common metric, allowing one to directly average M-scores across M-tasks with the expectation that the resulting average will carry greater validity than the individual tasks (e.g., Im-Bolter et al., 2006).

The current results suggest that intellectually gifted children might be characterized by a superior M-capacity. It could be argued, however, that characterizing giftedness in terms of heightened performance on task-based measures of mental attention is not without problems. That is, the measure of M-capacity conflates an individual's true or structural (i.e., organismic) M-capacity with their functional (here-and-now used) M-capacity as exhibited in performance. The latter is reflective of M-capacity, but also of the participant's repertoire of executive schemes (e.g., the problem-solving strategies being used). It may therefore be the case that gifted students are in fact advantaged in terms of a superior endogenous M-capacity and/or a superior repertoire of learned executive schemes. The current results do not allow us to choose between these two options. Previous research suggests, however, that at least some cognitively gifted students may be advantaged in both endogenous (i.e., true M-capacity) and experiential (i.e., learned executive know-how) developmental factors (Johnson, Howard, & Pascual-Leone, 2011; Pascual-Leone & Johnson, 2005; Pascual-Leone et al., 2004). This importantly suggests that gifted students can neither be characterized purely in terms of endogenous development nor purely in terms of learning.

It must be noted that the scope of the current study does not allow for critical engagement with the ongoing debate regarding how to most appropriately operationalize 'giftedness' for the purposes of education. That is, giftedness was operationalized in the current study strictly on the basis of established school board criteria, which required extremely high performance on standardized ability tests. Numerous researchers, however, have argued that this conception of giftedness is overly narrow. Instead, many researchers and educational boards have moved to broader and more inclusive definitions of giftedness (e.g., Gagné, 1995; Renzulli & Reis, 1985), which often incorporate additional domains of gifted performance, a continuum that acknowledges the need to develop exceptional potential into exceptional performance, and suggest that varying sources of information (beyond standardized intelligence measures) be sought for the identification of gifted students (e.g., Subotnik, Olszewski, & Worrell, 2011). Although these are important and necessary expansions to the gifted construct, only an intellectual form of

giftedness (as evidenced by exceptionally high performance on standardized intelligence measures) was considered here.

In summary, the current study provides further support for the TCO's developmental predictions. The current results demonstrate a clear and consistent gifted advantage on both the FIT and DFT (in M-scores and proportional accuracy by item class), which supports the idea that superior performance on M-measures may be a characteristic of intellectual giftedness. Moreover, this points to M-measures as a potentially viable means (which previous research has suggested is more precise and culture-fair) to measure important cognitive variance underlying intelligent performances. This may be particularly the case for the FIT, given its minimal requirements for literacy and numeracy skills, as well as its reduced requirement for sustained attention. The current results also contrast common notions of attention held by contemporary WM researchers, who often avoid quantifying the capacity of attention from an organismic perspective. Instead, the current results provide evidence that M-measurement, as informed by the TCO, generates clear, consistent, and predictable developmental trends. Taken together, these results give clear support for the TCO model of mental attention, including its unique predictions regarding the superior performance of intellectually gifted children on M-measures.

Author note

References

Alloway, T.P., & Alloway, R.G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Psychology, 106*(1). 20-29.

Alloway, T.P., Gathercole, S.E., Adams, A-M., & Willis, C. (2005). Working memory abilities in children with special educational needs. *Educational & Child Psychology, 22*(4), 56-67.

Arsalidou, M., Pascual-Leone, J., & Johnson, J. (2010). Misleading cues improve developmental assessment of working memory capacity: The color matching tasks. *Cognitive Development, 25*(3), 262-277. doi: 10.1016/j.cogdev.2010.07.001

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation, vol. 8* (pp. 47-89). New York, NY: Holt, Rinehart and Winston.

Baltes, P.B., & Reinert, G. (1969). Cohort effects in cognitive development of children as revealed by cross-sectional consequences. *Developmental Psychology, 1*(2), 169-177. doi: 10.1037/h0026997

Barrouillet, P., & Camos, V. (2001). Developmental increase in working memory span: Resource sharing or temporal decay? *Journal of Memory and Language, 45*, 1-20. doi: 10.1006/jmla.2001.2767

Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *Quarterly Journal of Economics, 121*(4), 1437-1472.

Bereiter, C., & Scardamalia, M. (1979). Pascual-Leone's M construct as a link between cognitive-developmental and psychometric concepts of intelligence. *Intelligence, 3*, 41-63. doi: 10.1016/0160-2896(79)80005-7

Bourne, V.J., Fox, H.C., Deary, I.J., & Whalley, L.J. (2007). Does childhood intelligence predict variation in cognitive change in later life? *Personality and Individual Differences, 42*(8), 1551-1559. doi: 10.1016/j.paid.2006.10.030

Brinch, C.N., & Galloway, T.A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences, 109*(2), 425-430. doi: 10.1073/pnas.1106077109

Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist, 52*(10), 1046-1050. doi: 10.1037/0003-066X.52.10.1046

Campbell, F.A., Pungello, E.P., Miller-Johnson, S., Burchinal, M., & Ramey, C.T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37*(2), 231-242. doi: 10.1037/0012-1649.37.2.231

Capron, C., & Duyme, M. (1989). Assessment of effects of socio-economic status on IQ in a full cross-fostering study. *Nature, 340*(6234), 552-554. doi: 10.1038/340552a0

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

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

Ceci, S.J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*(5), 703-722. doi: 10.1037/0012-1649.27.5.703

Ceci, S.J., & Gilstrap, L.L. (2000). Determinants of intelligence: Schooling and intelligence. In A.E. Kazdin (Ed.), *Encyclopedia of psychology, vol. 2* (pp. 496-498). New York, NY: Oxford University Press.

Ceci, S.J., & Williams, W.M. (1997). Schooling, intelligence, and income. *American Psychologist, 52*(10), 1051-1058. doi: 10.1037/0003-066X.52.10.1051

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

Cowan, N. (2005). *Working memory capacity.* East Sussex, UK: Psychology Press. doi: 10.4324/9780203342398

Cowan, N., & Alloway, T. (2009). Development of working memory in childhood. In M.L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 303-342). New York, NY: Psychology Press.

Cowan, N., Hismjatullina, A., AuBuchon, A.M., Saults, J.S., Horton, N., Leadbitter, K., & Towse, J. (2010). With development, list recall includes more chunks, not just larger ones. *Developmental Psychology, 46*, 1119-1131. doi: 10.1037/a0020618

Cowan, N., Nugent, L.D., Elliott, E.M., Ponomarev, I., & Saults, J.S. (1999). The role of attention in the development of working memory: Age differences in the verbal span of apprehension. *Child Development, 70*(5), 1082-1097. doi: 10.1111/1467-8624.00080

Cunning, S. (2003). The direction-following task: Assessing mental capacity through the linguistic domain. Unpublished doctoral dissertation, York University, Toronto, Canada.

de Jong, P.F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology, 70*(2), 75-96. doi: 10.1006/jecp.1998.2451

Deary, I.J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13-21. doi: 10.1016/j.intell.2006.02.001

Dempster, F.N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*(1), 63-100. doi: 10.1037/0033-2909.89.1.63

Duckworth, A.L., Quinn, P.D., Lynam, D.R., Loeber, R., & Stouthamer-Loeber, M. (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences, 108*(19), 7716-7720. doi: 10.1073/pnas.1018601108

Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of "dull children"? A late adoption study. *Proceedings of the National Academy of Sciences, 96*, 8790-8794. doi: 10.1073/pnas.96.15.8790

Edwards, O.W., & Fuller, D.P. (2005). Implications of the Cattell-Horn-Carroll theory on ethnic differences in IQ. *Psychological Reports, 97*(3), 891-897. doi: 10.2466/pr0.97.3.891-897

Engle, P.M.A., Santos, F.H., & Gathercole, S.E. (2008) Are working memory measures free of socioeconomic influence? *Journal of* Speech, Language, and Hearing Research, 51, 1580-1587. doi: 10.1044/1092-4388(2008/07-0210)

Engle, R.W. (2010). Role of working-memory capacity in cognitive control. *Current Anthropology, 51*(Suppl. 1), S17-S26. doi: 10.1086/650572

Firkowska-Mankiewicz, A. (2011). Adult careers: Does childhood IQ predict later life outcome? *Journal of Policy and Practice in Intellectual Disabilities, 8*(1), 1-9. doi: 10.1111/j.1741-1130.2011.00281.x

Gagne, F. (1995). From giftedness to talent: A developmental model and its impact on the language of the field. *Roeper Review, 18*(2), 103-111. doi: 10.1080/02783199509553709

Gathercole, S.E., Alloway, T.P., Willis, C., & Adams, A-M. (2005). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*(3), 265-281. doi: 10.1016/j.jecp.2005.08.003

Gathercole, S.E., Pickering, S.J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177-190. doi: 10.1037/0012-1649.40.2.177

Globerson, T. (1983). Mental capacity and cognitive functioning: Developmental and social class differences. *Developmental Psychology, 19*(2), 225-230. doi: 10.1037/0012-1649.19.2.225

Gottfredson, L.S. (1997). Why g matters: The complexity of everyday life. *Intelligence, 24*(1), 79-132. doi: 10.1016/S0160-2896(97)90014-3

Gottfredson, L.S. (2004). Life, death, and intelligence. *Journal of Cognitive Education and Psychology, 4*(1), 23-46. doi: 10.1891/194589504787382839

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H Brookes Publishing.

Herrnstein, R.J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York, NY: Free Press.

Hoard, M.K., Geary, D.C., Byrd-Craven, J., & Nugent, L. (2008). Mathematical cognition in intellectually precocious first graders. *Developmental Neuropsychology, 33*(3), 251-276. doi: 10.1080/87565640801982338

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77*(6), 1822-1841. doi: 10.1111/j.1467-8624.2006.00976.x

Jaeggi, S.M., Buschkuehl, M., Jonides, J., & Perrig, W.J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829-6833. doi: 10.1073/pnas.0801268105

Jausovec, N. (1998). Are gifted individuals less chaotic thinkers? *Personality and Individual Differences, 25*(2), 253-267. doi: 10.1016/S0191-8869(98)00039-7

Johnson, D.L., Swank, P.R., Howie, V.M., Baldwin, C.D., & Owen, M. (1996). Breastfeeding and children's intelligence. *Psychological Reports, 79*, 1179=-1185. doi: 10.2466/pr0.1996.79.3f.1179

Johnson, J., Howard, S.J., & Pascual-Leone, J. (2011) Attentional processes subserving working memory: Mental versus perceptual attention. Unpublished manuscript, York University, Toronto, ON.

Johnson, J., Im-Bolter, N., & Pascual-Leone, J. (2003). Development of mental attention in gifted and mainstream children: The role of mental capacity, inhibition, and speed of processing. *Child Development, 74*(6), 1594-1614. doi: 10.1046/j.1467-8624.2003.00626.x

Kane, M.J., Conway, A.R.A., Hambrick, D.Z., & Engle, R.W. (2007). Variation in working memory capacity as variation in executive attention and control. In A.R.A. Conway, C. Jarrold, M.J. Kane, A. Miyake, & J.N. Towse (Eds.), *Variation in working memory* (pp. 21-46). New York, NY: Oxford University Press.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*(6), 781-791. doi: 10.1076/jcen.24.6.781.8395

Mackey, A.P., Hill, S.S., Stone, S.I., & Bunge, S.A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science, 14*(3), 582-590. doi: 10.1111/j. 1467-7687.2010.01005.x

Mainela-Arnold, E., Evans, J.L., & Coady, J. (2010). Beyond capacity limitations II: Effects of lexical processes on word recall in verbal working memory tasks in children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(6), 1656-1672. doi: 10.1044/1092-4388(2010/08-0240)

Mercer, J.R. (1988). Ethnic differences in IQ scores: What do they mean? *Hispanic Journal of Behavioral Sciences, 10*(3), 199-218. doi: 10.1177/07399863880103002

Miller, R., Pascual-Leone, J., Campbell, C., & Juckes, T. (1989). Cross-cultural similarities and differences on two neo-Piagetian cognitive tasks. *International Journal of Psychology, 24*(3), 293-313.

Moore, E.G. (1986). Family socialization and the IQ test performance of traditionally and transracially adopted Black children. *Developmental Psychology, 22*(3), 317-326. doi: 10.1037/0012-1649.22.3.317

Morra, S. (1994). Issues in working memory measurement: Testing for M capacity. *International Journal of Behavioral Development, 17*(1), 143-159.

Murray, C. (1996). Murray's précis. *Current Anthropology, 37*(Suppl. 1), S143-S151.

Navarro, J.I., Ramiro, P., Lopez, J.M., Aguilar, M., Acosta, M., & Montero, J. (2006). Mental attention in gifted and nongifted children. *European Journal of Psychology of Education, 21*(4), 401-411. doi: 10.1007/BF03173510

Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77-101. doi: 10.1037/0003-066X.51.2.77

Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D.F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist, 67*(2), 130-159. doi: 10.1037/a0026699

Nokes C., & Bundy, D. (1994). Does Helminth Infection affect mental processing and educational achievement? *Parasitology Today, 10*(1), 14-18. doi: 10.1016/0169-4758(94)90348-4

Ostrosky-Solis, F., & Lozano, A. (2006). Digit span: Effect of education and culture. *International Journal of Psychology, 5*, 333-341. doi: 10.1080/00207590500345724

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*, 301-345. doi: 10.1016/0001-6918(70)90108-3

Pascual-Leone, J. (1984). Attentional, dialectic, and mental effort: Toward an organismic theory of life stages. In C. Armon, M.L. Commons, & F. Richards (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 182-215). New York, NY: Praeger.

Pascual-Leone, J. (2000). Reflections on working memory: Are the two models complementary? *Journal of Experimental Child Psychology*, *77,* 138-154. doi: 10.1006/jecp.2000.2593

Pascual-Leone, J., & Baillargeon, R. (1994). Developmental measurement of mental attention. *International Journal of Behavioral Development, 17*, 161-200.

Pascual-Leone, J., & Goodman, D. (1979). Intelligence and experience: A neopiagetian approach. *Instructional Science, 8*(4), 301-367. doi: 10.1007/BF00117011

Pascual-Leone, J., & Ijaz, H. (1989). Mental capacity testing as a form of intellectual-development assessment. In R. Samuda, S. Kong, J. Cummins, J. Pascual-Leone, & J. Lewis (Eds.), *Assessment and placement of minority students* (pp. 143-171). Toronto, ON: Hogrefe & Huber.

Pascual-Leone, J., & Johnson, J. (1991). The psychological unit and its role in task analysis. A reinterpretation of object permanence. In M. Chandler & M. Chapman (Eds.), *Criteria for competence: Controversies in the assessment of children's abilities* (pp. 153-187). Hillsdale, NJ: Erlbaum.

Pascual-Leone, J., & Johnson, J. (2005). A dialectical constructivist view of developmental intelligence. In O. Wilhelm & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 177-201). London, UK: Sage Publications. doi: 10.4135/9781452233529.n11

Pascual-Leone, J., & Johnson, J. (2011). A developmental theory of mental attention. In P. Barrouillet & V. Gaillard (Eds.), *Cognitive development and working memory: From neo-Piagetian to cognitive approaches* (pp. 13-46). New York, NY: Psychology Press.

Pascual-Leone, J., Johnson, J., Baskind, S., Dworsky, S., & Severtson, E. (2000). Culture-fair assessment and the processes of mental attention. In A. Kozulin & Y. Rand (Eds.), *Experience of mediated learning: An impact of Feuerstein's theory in education and psychology* (pp. 191-214). Elmsford, NY: Pergamon Press.

Pascual-Leone, J., Johnson, J., & Calvo, A. (2004, June). *Can mental attentional capacity predict the Canadian Cognitive Abilities score of school children?* Poster session presented at the meeting of the Jean Piaget Society, Toronto, Canada.

Pascual-Leone, J., Johnson, J., Calvo, A., & Verilli, E. (2005, August). *Can latent giftedness be indexed by mental-attentional capacity?* Poster session presented at the 16th Biennial World Conference of the World Council for Gifted and Talented Children, New Orleans, LA.

Redick, T.S., Unsworth, N., Kelly, A.J., & Engle, R.W. (2012). Faster, smarter? Working memory capacity and perceptual speed in relation to fluid intelligence. *Journal of Cognitive Psychology, 24*(7), 844-854. doi: 10.1080/20445911.2012.704359

Renzulli, J.S., & Reis, S.M. (1985). *The schoolwide enrichment model: A comprehensive plan for educational excellence*. Mansfield, CT: Creative Learning Press.

Richardson, K. (2002). What IQ tests test. *Theory & Psychology, 12*(3), 283-314. doi: 10.1177/0959354302012003012

Rueda, M.R., Rothbart, M.K., McCandliss, B.D., Saccomanno, L., & Posner, M.I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences, 102*(41), 14931-14936. doi: 10.1073/pnas.0506897102

Rushton, J.P. (1995). *Race, evolution and behavior: A life history perspective*. New Brunswick, NJ: Transaction.

Rushton, J.P., & Jensen, A.R. (2010). Race and IQ: A theory-based review of the research in Richard Nisbett's Intelligence and How to Get It. *The Open Psychology Journal, 3,* 9-35. doi: 10.2174/1874350101003010009

Saccuzzo, D.P., Johnson, N.E., & Guertin, T.L. (1994). Information processing in gifted versus nongifted African American, Latino, Filipino, and White children: Speeded versus nonspeeded paradigms. *Intelligence, 19*(2), 219-243. doi: 10.1016/0160-2896(94)90014-0

Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection methods in personal psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262-274. doi: 10.1037/0033-2909.124.2.262

Schmidt, F.L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*(1), 162-173. doi: 10.1037/0022-3514.86.1.162

Schoenthaler, S.J., Amos, S.P., Eysenck, H.J., Peritz, E., & Yudkin, J. (1991). Controlled trial of vitamin-mineral supplementation: Intelligence and performance. *Personality and Individual Differences, 12*(4), 351-362. doi: 10.1016/0191-8869(91)90287-L

Siegel, L.S., & Ryan, E.B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*(4), 973-980. doi: 10.2307/1131037

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*(5), 401-426. doi: 10.1016/j.intell.2006.09.004

Subotnik, R.F., Olszewski, P., & Worrell, F.C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest, 12*, 3-54. doi: 10.1177/1529100611418056

Swanson, H.L. (2004). Working memory and phonological processing as predictors of children's mathematical problem solving at different ages. *Memory & Cognition, 32*(4), 648-661. doi: 10.3758/BF03195856

Swanson, H.L. (2008). Working memory and intelligence in children: What develops? *Journal of Educational Psychology, 100*(3), 581-602. doi: 10.1037/0022-0663.100.3.581

Swanson, H.L. (2011). Intellectual growth in children as a function of domain specific and domain general working memory subgroups. *Intelligence, 39*(6), 481-492. doi: 10.1016/j.intell.2011.10.001

Swanson, H.L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*(3), 471-491. doi: 10.1037/0022-0663.96.3.471

Swanson, H.L., & Howell, M. (2001). Working memory, short-term memory, and speech rate as predictors of children's reading performance at different ages. *Journal of Educational Psychology, 93*(4), 720-734. doi: 10.1037/0022-0663.93.4.720

Towse, J.N., Hitch, G.J., Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language, 39*, 195-217. doi: 10.1006/jmla.1998.2574

van Ijzendoorn, M.H., Juffer, F., & Poelhuis, C.W.K. (2005). Adoption and cognitive development: A meta-analytic comparison of adopted and nonadopted children's IQ and school performance. *Psychological Bulletin, 131*(2), 301-316. doi: 10.1037/0033-2909.131.2.301

Wahlsten, D. (1997). The malleability of intelligence is not constrained by heritability. In B. Devlin (Ed.), *Intelligence, genes, and success: Scientists respond to The Bell Curve* (pp. 71-87). New York, NY: Springer-Verlag.

Wasik, B.H., Ramey, C.T., Bryant, D.M., & Sparling, J.J. (1990). A longitudinal study of two early intervention strategies: Project CARE. *Child Development, 61*(6), 1682-1696. doi: 10.2307/1130831

Weinberg, R.A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist, 44*(2), 98-104. doi: 10.1037/0003-066X.44.2.98