Intelligence 57 (2016) 96–104

Contents lists available at ScienceDirect

Intelligence

Short-term memory for faces relates to general intelligence moderately☆

Gilles E. Gignac *, Mahesh Shankaralingam, Kipling Walker, Philippe Kilpatrick

University of Western, Australia

ARTICLE INFO

Article history:
Received 15 January 2016
Received in revised form 9 April 2016
Accepted 4 May 2016
Available online 21 May 2016

Keywords:
Intelligence
CHC theory
Face identity recognition
Prosopagnosia

ABSTRACT

The results associated with a small number of investigations suggest that individual differences in memory for faces, as measured by the Cambridge Face Memory Test (CFMT), are independent of intelligence. Consequently, memory for faces has been suggested to be a special construct, unlike other cognitive abilities. However, previous investigations have measured intelligence with only one or two subtests. Additionally, the sample sizes upon which previous investigations were based were relatively small ($N = 45$ to 80). Consequently, in this investigation, a battery of eight cognitive ability tests and the CFMT were administered to a relatively large number of participants ($N = 211$). Based on a correlated-factor model, memory for faces was found to be correlated positively with fluid intelligence (.29), short-term memory (.23) and lexical knowledge ability (.19). Additionally, based on a higher-order model, memory for faces was found to be associated with *g* at .34. The results are interpreted to suggest that memory for faces, as measured by the CFMT, may be characterised as a relatively typical narrow cognitive ability within the Cattell–Horn–Carroll (CHC) model of intelligence, rather than a special ability (i.e., independent of other abilities). Future research with a greater diversity in the measurement of face recognition ability is encouraged (e.g., long-term memory), as the CFMT is a measure of short-term face memory ability.

1. Introduction

The capacity for face identity recognition has been a significant source of research over the years (e.g. Carey, Diamond, & Woods, 1980, Galper & Hochberg, 1971, Tanaka & Farah, 1993), perhaps in part because of the sensational phenomenon of prosopagnosia: the incapacity of otherwise cognitively able individuals to recognise familiar faces (Duchaine, 2011). In recent years, a number of investigators have begun to investigate face identity recognition ability as an individual difference construct (e.g. Dennett, McKone, Edwards, & Susilo, 2012, Rhodes, Jeffery, Taylor, Hayward, & Ewing, 2014, Sekiguchi, 2011). The empirical evidence suggests that face recognition ability lies along a continuum, with some individuals in possession of relatively poor levels of face recognition ability to those who may be considered "super recognizers" (Russell, Duchaine, & Nakayama, 2009).

At least superficially, individual differences in the capacity to memorise and recall faces may be suggested to be a cognitive ability, given that it is similar in nature to other types of well-established cognitive abilities such as short-term memory (*Gsm*) and visual–spatial ability (*Gv*): two lower-order constructs known to be associated with general intelligence (*g*; Carroll, 1993). To-date, the empirical research relevant to the association between face recognition ability and intelligence is very mixed. Some research suggests that there is a substantial association between face recognition ability and other cognitive abilities (e.g., Hildebrandt, Wilhelm, Schmiedek, Herzmann, & Sommer, 2011). By contrast, others have contended that face identity recognition ability is a construct completely distinct from other cognitive abilities, including *g* (Wilmer, Germine, & Nakayama, 2014).

Arguably, previous investigations may be suggested to be limited, as they have not administered a comprehensive battery of cognitive ability tests, or they have not administered the most commonly administered measure of face recognition ability, the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006). Consequently, the purpose of this investigation was to estimate the latent variable association between face recognition ability and other cognitive abilities, including *g*, through administration of a battery of cognitive ability tests and the CFMT.

1.1. Face identity recognition ability and individual differences

Although it has been stated that all adult humans are experts at face recognition (Haxby, Hoffman, & Gobbini, 2000), the empirical research suggests that there are, nonetheless, a non-negligible amount of

☆ Thanks to Romina Palermo, Linda Jeffery, and Laura McLaughlin Engfors for comments during the preparation of this manuscript.

* Corresponding author at: School of Psychology, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia.
E-mail address: gilles.gignac@uwa.edu.au (G.E. Gignac).

http://dx.doi.org/10.1016/j.intell.2016.05.001
0160-2896/© 2016 Elsevier Inc. All rights reserved.

individual differences in the capacity to recognise faces. For example, individual differences in the capacity to recognise faces are apparent in the distribution of scores associated with the CFMT (Duchaine & Nakayama, 2006). The CFMT is the most commonly used test of face recognition ability (Cho et al., 2015). The items within the CFMT (72 in total) consist of photos of faces displayed on a computer monitor. The photos are ellipsoid in shape such that they exclude characteristics such as the model's hair and neck/clothes. Additionally, the models are not wearing make-up or jewelry. Consequently, the participant viewing the images cannot rely upon non-intrinsic characteristics of the face for the purposes of memorisation.[1] For each trial, the participant must first memorise three faces on a computer screen over a period of 20 s, after which the faces disappear from the screen. Then, another series of three faces appear on the screen and the participant must identify which one of the three faces was presented during the memorisation phase. Because the test phase within the CFMT occurs essentially immediately after the memorisation phase, a score on the CFMT is probably best considered as an indicator of short-term face memory, rather than long-term face memory. Also, note that for each CFMT item, there is a correct response alternative, and the participant must select one of the three faces. Thus, the CFMT is arguably not susceptible to response biases (e.g., tendency to respond “haven't seen”).

Short-term face recognition ability has been found to be a dimension associated with a moderate amount of variability. For example, based on a university sample ($N = 50$), the CFMT has been reported to be associated with a mean of 57.92 and a standard deviation of 7.91 (Duchaine & Nakayama, 2006), which corresponds to a coefficient of variation of .14 (7.91 / 57.92 = .14). Based on a larger sample recruited from the general community ($N = 107$), Bowles et al. (2009) reported a mean of 54.6 and a standard deviation of 9.4 on the CFMT, which corresponds to a coefficient of variation of .17. For the purposes of comparison, Gignac (2015) reported a coefficient of variation of .19 for digit span forward, across several normative samples. Consequently, with respect to variability, short-term face recognition ability, as measured by the CFMT, is very comparable to serial recall of digits — a cognitive capacity well-known to be associated positively with *g*. Thus, short-term face recognition ability may be considered as a possible correlate of *g*, as it shares approximately the same amount of variability in the normal population as other indicators of short-term memory.

1.2. Short-term memory and g

It has been well established that short-term memory capacity is related positively to *g* (Bachelder & Denny, 1977; Gignac & Watkins, 2015; Miller & Vernon, 1992). Within the Cattell–Horn–Carroll (CHC) three stratum model (McGrew, 2009), short-term memory (G_{sm}) is known as one of the nine broad (stratum II) factors, alongside fluid intelligence (*Gf*), crystallised intelligence (*Gc*) and processing speed (*Gs*), for example (Carroll, 2003). Based on the Wechsler Adult Intelligence Scale — IV (Wechsler, 2008) normative sample ($N = 2200$), Gignac (2014) found that a *G*sm lower-order factor was associated with *g* at .84; a result replicated closely with Wechsler Intelligence Scale for Children — V (Wechsler, 2014) normative sample (Gignac & Watkins, 2015). In another investigation based on a combination of the Differential Ability Scales (Elliott, 1990) and the Woodcock–Johnson Tests of Cognitive Abilities-III (Woodcock, McGrew, & Mather, 2001), Sanders, McIntosh, Dunham, Rothlisberg, and Finch (2007) reported an association of .61 between *G*sm and *g*. Additionally, a lower-order visual processing factor (*G*v) was reported to be associated with *g* at .76. Similarly, Reynolds, Keith, Fine, Fisher, and Low (2007) reported *G*sm and *G*v associations with *g* of .70 and .83, respectively based on the KABC — II (Kaufman & Kaufman, 2004). Thus, as it may be suggested that the completion of short-term face recognition ability tests involves short-term memory and visual processing processes, it is plausible to suggest that short-term face recognition ability may be related to *g*.

Typically, individual tests of short-term memory capacity are observed to relate to *g* moderately (.30 to .50), rather than very appreciably in magnitude (.60 to .70). For example, digit span forward has been shown to relate to *g* at approximately .40, based on a bifactor model of the Wechsler Adult Intelligence Scale (Wechsler, 2008) normative sample (Gignac and Weiss, 2015). Thus, serial recall for digits (verbal memory) shared approximately 15% of its variance with *g*. Somewhat more appreciably, backward digit span, which is considered to involve some working memory capacity processing, was found to be related to *g* at .48.

Measures of visual memory have also been found to relate to *g* moderately. Consider, for example, the Rey–Osterrieth Complex Figure test (Osterrieth, 1944; Rey, 1941), which requires participants to copy a visually displayed complex figure on paper with a pencil. After a particular period of time, during which the participant completes other tasks, the participant is requested to re-draw the complex figure without forewarning (delayed recall). Higher scores are achieved contingent upon the accuracy with which a participant recreated the complex figure. Based on a higher-order model of cognitive abilities, Irwing, Booth, Nyborg, and Rushton (2012) found that the Rey-Osterrieth Complex Figure test was associated with *g* at .32 and .19 (Schmid–Leiman decomposed) as a cross-loading indicator of *GvGf* and *G*sm/*G*lr, respectively. Thus, visual memory, as measured by the Rey–Osterrieth Complex Figure test, shared approximately 14% of its variance with *g* ($.32^2 + .19^2 = .14$). In another investigation, Reynolds, Keith, Flanagan, and Alfonso (2013) reported the results associated with a cross-battery higher-order model of intelligence, which included the Picture Recognition subtest from the Woodcock–Johnson III (Woodcock et al., 2001). Picture Recognition was found to load onto an associative memory lower-order factor at .58. The associative memory lower-order factor loaded onto *g* at .82. Thus, based on a Schmid–Leiman decomposition of the higher-order effects, Picture Recognition was found to be associated with *g* at .48.

Given that face recognition ability may be, at least qualitatively, classified as a construct relevant to memory and visual cognitive processes, it may be suggested that individual differences in performance on face recognition ability tests (e.g., the CFMT) would be a representative of cognitive ability, at least to some degree. Theoretically, in order for face recognition ability to be classified as a cognitive ability, it would arguably have to be demonstrated to correlate positively with other well-known cognitive abilities (e.g., *Gf*, *Gc*, *G*sm). Additionally, and relatedly, it would be expected that face recognition ability would share variance with *g*. To-date, only a relatively small amount of empirical investigations have examined the association between face recognition ability and other cognitive abilities.

1.3. Short-term face recognition and intelligence

Davis et al. (2011) examined the association between face recognition ability and intelligence through administration of the CFMT and the Culture Fair Intelligence Test (CFIT; Cattell, 1963). Based on a sample of university students ($N = 63$), Davis et al. reported a correlation of −.08 between the CFMT and the CFIT. Thus, individual differences in face recognition ability were interpreted to be unrelated to non-verbal fluid intelligence. It will be noted, however, that the sample was found to be associated with a CFIT mean of 122 (the CFIT normative sample mean is 100 with an SD of 15). Consequently, performance on only a small number of items would have discriminated between many of the participants, as the CFIT subtests consist of only 10 to 14 items. Additionally, an estimate of intelligence based on, essentially, a single test

[1] The Kaufman Assessment Battery for Children — II (KABC; Kaufman & Kaufman, 2004) and the Wechsler Memory Scale III (Wechsler, 2008) include face recognition subtests. However, these subtests have been criticised as invalid indicators of face recognition ability, because the images include non-intrinsic characteristics such as hair, clothes, and a mixture of races (Dalrymple & Palermo, 2016).Consequently, research relevant to these subtests is not reviewed here.

may be considered rather limited. Jensen (1998) recommended the administration of nine tests for the purposes of estimating general intelligence. Finally, a sample size of only 63 is arguably insufficient to help support null hypothesis conclusions, convincingly.

In a very similar investigation, Palermo, O'Connor, Davis, Irons, and McKone (2013) reported a correlation of −.01 between the CFMT and the CFIT, based on a sample of mostly undergraduate university students ($N = 80$). Thus, again, the results were interpreted to suggest that there was an absence of evidence for an association between face recognition ability and non-verbal fluid intelligence. Consistent with Davis et al., however, the sample was found to be associated with a CFIT mean of 123, which suggests that the CFIT may not have been an appropriate discriminator of intelligence for the sample. Both the Davis et al. (2011) and the Palermo et al. (2013) samples were highly selected, as they were based on third-year university students at a highly rated university (R. Palermo, personal communication, December 7, 2015). Additionally, only one test of intelligence was administered, which precluded the possibility of estimating g.

In contrast to Davis et al. (2011) and Palermo et al. (2013); Peterson and Miller (2012) administered the CFMT and two subtests from the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999): Vocabulary and Matrix Reasoning. Based on a sample of 45 university students, Peterson et al. reported a correlation of .21 between the CFMT and the Matrix Reasoning subtest. However, the correlation was not significant statistically ($p = .166$), which may not be surprising, considering the lack of statistical power associated with an analysis based on a sample size of only 45 participants. The correlation between the CFMT and Vocabulary was reported at .01. Thus, again, there was no compelling evidence to suggest that face recognition ability was related to intelligence.

In a recent review of the above empirical research, Wilmer et al. (2014) reported a mean correlation of .01 between the CFMT and "general abilities" (pp. 2), despite the fact that none of the investigations measured general intelligence. Wilmer et al. (2014) concluded that face recognition ability "…dissociates almost completely from standardized IQ tests." Prior to the publication of the above empirical research, Bowles et al. (2009, pp. 452) expressed a very similar view, based principally upon indirect evidence: "…intelligence probably does not affect face memory (CFMT)…" Thus, the common view in the literature is that face recognition ability is completely distinct from other cognitive abilities, and, for this reason, may be considered "special" (Shakeshaft & Plomin, 2015, pp. 12890). Arguably, however, all of the investigations reviewed above may be suggested to have been seriously limited.

First and foremost, all of the investigations estimated intelligence with either one test or a maximum of two. Consequently, it would be unjustified to suggest that the previous investigations yielded an adequate estimate of *g*, as a minimum of nine relatively diverse cognitive ability tests has been suggested for such purposes (Jensen, 1998). An additional benefit associated with the administration of several cognitive ability tests is that it affords the opportunity to estimate associations between latent variables. Latent variables are not contaminated by measurement error (Fan, 2003). Consequently, the observed effects between latent variables (e.g., correlations) are not attenuated due to measurement error (Nunnally & Bernstein, 1994). Finally, none of the samples in the above investigations were large, nor were they representative of the population (university students). In light of these limitations, it may be contended that the conclusion that face recognition ability is completely dissociated from intelligence is premature.

In addition to the above intelligence and CFMT research, there are also a small number of studies that have examined the association between intelligence and measures of face recognition ability less well-known than the CFMT. For example, Hildebrandt et al. (2011) investigated the age moderating effects on the association between face recognition ability and *g*, as measured by their own developed face recognition tasks ($N = 448$). Unfortunately, Hildebrandt et al. (2011) did not report the standardised effect between their general intelligence latent variable and the face recognition ability latent variable. However, they stated in the discussion that general intelligence "…accounted for only about half of the variance of face perception and face memory factors" (p. 711). Such a statement would imply that the standardised effect (standardised beta weight) was approximately $\beta = .70$. Similarly, based on a bifactor model of intelligence and a latent variable defined by three face memory tests, Wilhelm et al. (2010) found that 48% ($\beta = .69$) of the variance associated with a memory for faces latent variable was accounted for by a combination of *g*, object cognition, and immediate and delayed memory. These results are in stark contrast to the *N*-weighted mean correlation of .01 between cognitive abilities and the CFMT reported by Wilmer et al.'s (2014) review, which did not include the Hildebrandt et al. (2011) or the Wilhelm et al. (2010) studies.

The Hildebrandt et al. (2011) results suggest that face recognition ability may be a very appreciably associated with *g*. It is worth noting, however, that the *g* factor modeled by Hildebrandt et al. (2011) was composed of nine subtests, five of which were tests of memory, and three of which were tests of processing speed. The remaining test was a short-form (16 items) of Raven's (Raven, Court, & Raven, 1979). Consequently, given the heavy weighting toward memory tests in the battery, an effect of .70 between the general intelligence factor and face recognition ability, which is itself a test of memory, may be considered substantially upwardly biased. Nonetheless, the results of Hildebrandt et al. (2011) do suggest that face recognition ability may possibly be related to *g* positively.

In light of the above, the firm conclusion drawn by Wilmer et al. (2014) in relation to face recognition ability and intelligence may be suggested to be premature. By contrast, the effects reported by Hildebrandt et al. (2011) may be considered excessively large. Consequently, the purpose of this investigation was to estimate the association between face recognition ability, as measured by the CFMT, and a relatively diverse battery of cognitive ability tests via latent variable modeling. Based on previous research relevant to relatively narrow memory tests and intelligence, it was hypothesized that face recognition ability would associate with *g* positively and moderately (≈.30 to .50).

2. Method

2.1. Sample

The total sample consisted of 211 participants (68% female) who spoke English as a first language. The age range of the participants was 17 to 35 ($M = 19.8$, SD = 2.9). The original sample contained an additional 12 participants over the age of 35 years. However, as age-based norms were not available for most of the tests used in this investigation, they were omitted from the final sample. The participants were recruited principally from a first-year undergraduate unit within a psychology program at a large university in Australia. Participants received partial course credit for participation. The investigation was approved by the university's ethics committee.

2.2. Measures

A battery of nine cognitive ability tests were administered for the purposes of estimating a general factor, one of which included a test of face recognition ability.

Gf was measured with the following four tests:

2.2.1. Culture fair intelligence test — series (Scale 3, Form A)

A test designed to measure the ability to detect a pattern amongst a series of figural images (Cattell, 1963). The participant must identify the missing figure to complete the series amongst six alternatives. In accordance with the manual, participants were given 3 min to complete the test (13 items). Internal consistency reliability was estimated at $\alpha = .28$ ($\alpha_{upper} = .89$)[2] in this sample.

2.2.2. Culture fair intelligence test — matrices (Scale 3, Form A)

A test designed to measure the ability to detect a pattern within a matrix (Cattell, 1963). The participant must identify the missing figure to complete the matrix amongst six alternative figs. In accordance with the manual, participants were given 3 min to complete the test (13 items). Internal consistency reliability was estimated at $\alpha = .31$ ($\alpha_{upper} = .90$) in this sample.

2.2.3. Quickie Gf/Gc number series

A test designed to measure the ability to reason with numbers (Stankov, 1997). For each item, a sequence of 11 single-digit numbers is presented in a row. The participant must identify the progressive pattern amongst the numbers in order to identify the missing number at the end of the row. The test consists of 15 items. Participants were given 6 min to complete as many items as possible. Internal consistency of $\alpha = .92$ ($\alpha_{upper} = .97$) was estimated in this investigation's sample.

2.2.4. Vandenberg mental rotation task

A re-drawn version published by Peters et al. (1995) of a well-known test designed to measure the ability to rotate figural shapes in 2-D space (Vandenberg & Kuse, 1978). In this investigation, Form A was used which consists of 24 items. Participants were given 6 min to complete as many items as possible. Internal consistency reliability was estimated at $\alpha = .88$ ($\alpha_{upper} = .96$) in this investigation's sample.

*G*sm was measured with following three tests:

2.2.5. Digit span backward

A test adapted from the WAIS — IV (Wechsler, 2008). Participants were presented via audio presentation a series of single-digit numbers (one second pause between each digit). Participants were required to recite the numbers in the reverse order with which they were presented. The first two trials consisted of only two numbers and were considered practice trials. Following correct completion of the practice trials, the participants were presented with two trials per series of digits progressively from two digits up to a maximum of nine digits. The test was terminated if the participant recalled incorrectly two trials in a row. The maximum possible score on this test was 18 (1 point per trial). The internal consistency reliability in this sample was $\alpha = .74$ ($\alpha_{upper} = .85$).

2.2.6. Word span backward

A tested adapted from La Pointe and Engle (1990). Participants were presented via audio presentation a series of three syllable words (one second pause between each word; the words were selected from the Appendix in La Pointe & Engle, 1990). Participants were required to recite the words in the reverse order with which they were presented. The first two trials consisted of only two words and were considered practice trials. Following correct completion of the practice trials, the participants were presented with two trials per series of words progressively from two words up to a maximum of nine words. The test was terminated if the participant recalled incorrectly two trials in a row. The maximum possible score on this test was 18 (1 point per trial). The internal consistency reliability in this sample was $\alpha = .59$ ($\alpha_{upper} = .78$).

2.2.7. Visual span backward

A newly created tested for the purposes of this investigation. Participants were presented on a computer monitor (PowerPoint Animation) a series of basic shapes with three lines and two corners. The shapes were developed so as to not be similar to other well-known shapes. As an example, the series of shapes presented in series 5, trial 2 consisted of the following shapes:

For each trial, a 'Ready' prompt was shown for 750 ms to begin the testing for that trial. Next, a blank screen was shown for 1000 ms. Shapes were then presented for 1000 ms with 1000 ms blank screen intervals. After the final shape for that trial was shown, a 'Draw' prompt was depicted on the screen. Participants were required to draw the shapes on a response sheet in the reverse order with which they were presented on the computer monitor. Two trials per series of shapes were presented progressively to the participants from two shapes up to a maximum of eight shapes. The test was terminated if the participant recalled incorrectly two trials in a row. The maximum possible score on this test was 14 (1 point per trial). The internal consistency reliability in this sample was $\alpha = .62$ ($\alpha_{upper} = .83$).

*G*c was measured with a single test:

2.2.8. Advanced Vocabulary Test

An original multiple-choice vocabulary test designed for the purposes of this investigation. Although multiple-choice vocabulary tests have been developed previously (e.g., Multidimensional Aptitude Battery; Jackson, 1998), a test more appropriate for a high-ability sample was considered required in this investigation, as the participants in the sample were recruited from students enrolled in a well-established university. The AVT consists of 21 multiple-choice items with five response alternatives. Participants were given 5 min to complete as many items as possible. Correct responses were given one point and there was no penalty for guessing incorrectly.

The main difference between the AVT and other vocabulary tests is that there are no easy items. For example, the first two items to be defined within the AVT are 'camaraderie' (response alternatives: wild party, impersonal chatting, collection of colors, disastrous outcome, friendship) and 'irate' (response alternatives: very excited, great anger, restricted, visible from a distance, evasive). Consequently, participants spend more time completing items closer to their ability level, which, theoretically, should yield scores which can discriminate between high-ability participants more precisely. Participants were given a maximum of 5 min to complete as many items as possible. In this sample, the internal consistency reliability was estimated at $\alpha = .73$ ($\alpha_{upper} = .97$). As will be reported below, the AVT was associated positively with a *g* factor, which suggests that the scores were associated with some validity.

2.2.9. Cambridge Face Memory Test

Face identity recognition ability was measured with the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006). The CFMT is a 72-item untimed test designed to measure the ability to recall 2-D images (digital photographic images) of non-famous faces presented on a computer screen. The faces are relatively expressionless emotionally. Consequently, the CFMT is considered a measure of face identify recognition, rather than a face emotion recognition (Palermo et al., 2013). The items are segmented into three categories: same images (18 items), novel images (30 items), and novel images with noise (24 items). With respect to the novel images with noise items, the 2-D images of the faces are pixelated ("noise") in such a way as to render the faces more difficult to perceive and recall. For each correctly recalled item, one point is given. Thus, 72 points is the maximum that can be achieved on the CFMT.

[2] Reliability estimation based on the conventional phi coefficient for dichotomously scored items has been shown to underestimate true reliability. For example, based on a simulation, Sun et al. (2009) demonstrated that use of the upper-bound phi coefficient tends to yield much more accurate estimates of internal consistency reliability. Consequently, in this investigation, the upper-bound phi-based coefficient alpha is reported (in parentheses, α_{upper}), in addition to the conventional phi-based coefficient alpha.

Although the CFMT is clearly a measure of short-term memory, CFMT scores may be considered to be associated with some generalisability across the sampling domain. For example, based on a sample of 1471 adult participants from the community, the CFMT has been reported to correlated .55 with a Famous Faces Memory Test (Wilmer et al., 2012). Additionally, average performance on the CFMT has been reported to correspond to 80% in the normal population (Bowles et al., 2009). By contrast, individuals suffering from prosopagnosia (self-reported significant problems with everyday face recognition) have been reported to perform at approximately 51% on the CFMT (33% is chance level performance). Thus, scores on the CFMT may be regarded, to some degree, as more broadly valid, as individuals with prosopagnosia suffer from general face memory difficulties (i.e., not only short-term memory). Finally, the internal consistency reliabilities were $\alpha = .57$ ($\alpha_{\text{upper}} = .98$), $\alpha = .83$ ($\alpha_{\text{upper}} = .98$), and $\alpha = .79$ ($\alpha_{\text{upper}} = .98$), for the same images, novel images, and novel images with noise subscales, respectively.

Almost invariably, researchers who use the CFMT calculate a total score from the 72 items, rather than model a latent variable defined by the three CFMT subtests. The possibility of modeling a face recognition ability latent variable was evaluated by testing a single-factor model defined by the same images subtest, the novel images subtest, and the novel images with noise subtest via maximum likelihood estimation within Amos 21 (Arbuckle, 2012). The maximum likelihood completely standardised solution yielded the following loadings: .48 (same images), .99 (novel images), and .79 (novel images with noise). The loading of .99 associated with the novel images subtest suggested that it was an essentially perfect indicator of the face recognition ability construct, as measured by the CFMT. Consequently, for the remainder of the analyses, only the novel images subtest was used as a measure of individual differences in face recognition ability.

In addition to the tests above, forward span versions of the digit, word, and visual span tests were administered for the purposes of a separate investigation. Only the backward versions of the tests were included in this investigation, as backward memory tests are known to be better indicators of *g* (Jensen & Figueroa, 1975; Gignac & Weiss, 2015). Additionally, there was concern that the *g* factor would be unbalanced in favor of memory tests, if all six memory span tests were included in the analyses.

2.3. Procedure

Participants first completed an informed consent form. All participants completed the testing individually through a combination of tests administered on a computer and face-to-face testing. Participants completed the memory span tests first. The order of the memory span test administration (numerical, verbal, visual) was randomised across the sample. Participants then completed the remaining tests in the following order: CFMT, CFIT-Series, CFIT-Matrices, Number Series Test, the Mental Rotation Test, and, finally, the Advanced Vocabulary Test. All participants were debriefed at the end of the testing session. Total testing time was between 55 and 60 min.

2.4. Data analysis

In order to evaluate the association between face recognition ability and other cognitive ability factors, a latent variable modeling strategy was employed. First, a correlated four-factor model was tested. The *Gf* first-order factor within the correlated-factor model was defined by the CFIT-series, CFIT-matrices, numerical reasoning, and mental rotation subtests. The *Gsm* first-order factor within the correlated-factor model was defined by the digit span backward, word span backward, and visual span backward subtests. As the stimuli associated with the mental rotation test and the visual span backward test were similar in nature (shapes), it was considered defensible theoretically to allow their uniquenesses to correlate. Face recognition ability in the correlated-factor model was represented by the CFMT novel faces subtest. Because the novel images CFMT subtest was the only indicator used as a measure of face recognition ability, and it was associated with imperfect levels of internal consistency reliability (i.e., .83), it was included in the correlated factor model as a single-indicator latent variable (see Jöreskog & Sörbom, 1982). Specifically, the novel images' indicator uniqueness in the model was fixed to S^2(1-alpha), where S^2 and alpha corresponded to the subtest's variance and coefficient alpha, respectively. Thus, in this case, the novel images CFMT subtest uniqueness was fixed to $25.512 * (1–.832) = 4.286$.

Similarly, as the Advanced Vocabulary Test was the only indicator of *Gc* in the intelligence battery, it was modeled as a single-indicator latent variable representation of the stratum I lexical knowledge ability (VL; see Newton & McGrew, 2010) within the correlated-factor model. As per face recognition ability, the VL ability single-indicator latent variable's uniqueness in the model was be fixed to S^2(1-alpha), or $14.824(1–.732) = 3.973$.

In addition to the correlated-factor model, a corresponding higher-order model was tested to estimate the association between the first-order factors and *g*. Finally, for thoroughness, a single-factor model defined by all 11 observed variables was tested to help justify a multidimensional interpretation of the data. Specifically, it was expected that the correlated-factor model and the higher-order model would fit better than the single-factor model.

For the purposes of identification/scaling, each latent variable's variance within the correlated-factor model was fixed to 1. In the higher-order model, the second-order latent variable variance was fixed to 1. Additionally, all of the first-order latent variables had one unstandardized loading fixed to 1 for scaling purposes. As scaling options are known to interact with point estimate standard errors and *p*-values (Gonzalez & Griffin, 2001), the model solution was estimated via the standardised bootstrap (2000 resamples). In addition to the chi-square test, model-fit was evaluated with CFI, RMSEA and BIC. As BIC incorporates a significant penalty for model complexity (Marsh, Hau, & Grayson, 2005), the models were compared via BIC, with smaller BIC values indicative of a preferred model. All analyses were performed with Amos 21 (Arbuckle, 2012).

3. Results

As can be seen in Table 1, most of the correlations between the CFMT subtests and the other cognitive ability subtests were positive in direction and statistically significant ($p < .05$). In particular, it was noted that the correlations between the CFMT2 subtest (novel faces) and the two Culture Fair Intelligence subtests were $r = .14$ ($p = .040$) and $r = .19$ ($p = .006$), respectively. Additionally, the largest numerical correlation with the CFMT2 subtest (novel faces) was with the visual spatial span subtest ($r = .23$, $p = .001$). Thus, higher levels of face recognition ability tended to be associated with higher scores on the other cognitive ability tests.

The correlated four-factor model was found to be associated with acceptable levels of model-fit, $\chi^2(22, N = 211) = 32.51$, $p = .069$, CFI $= .950$, RMSEA $= .048$, BIC $= 155.60$. As can be seen in Table 2, all of the inter-latent variable correlations associated with the correlated-factor model were positive and statistically significant. In particular, it will be noted that face recognition ability correlated positively with VL ($r = .19$, $p = .038$), *Gf* ($r = .29$, $p = .009$), and *Gsm* ($r = .23$, $p = .048$).

The corresponding higher-order model with four first-order lower-order factors was also found to be associated with good levels of model-fit, $\chi^2(24, N = 211) = 33.49$, $p = .094$, CFI $= .955$, RMSEA $= .043$, BIC $= 145.88$. As can be seen in Fig. 1, the higher-order model was associated with positive and statistically significant second-order loadings. In particular, it will be noted that the face recognition lower-order factor was associated with a second-order *g* loading of $\lambda = .34$ ($\lambda^2 = .12$), $p = .007$. Thus, approximately 12% of the variance in face recognition ability, as measured by the CFMT, was related to *g*. By

Table 1
Descriptive statistics and inter-subtest correlations.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	M	SD	Skew	Kurtosis
1. AVT	1.0											10.69	3.85	.03	−.73
2. CFS	.21	1.0										8.10	1.59	−.60	1.09
3. CFM	.13	.39	1.0									7.10	1.47	−.54	.54
4. NS	.07	.19	.32	1.0								6.09	4.84	.48	−1.11
5. MR	.13	.16	.29	.22	1.0							10.77	5.01	.54	−.10
6. DSB	.19	.33	.21	.15	.23	1.0						9.47	2.26	.28	−.24
7. WSB	.19	.20	.15	.04	.22	.41	1.0					3.84	1.31	.88	2.18
8. VSB	.16	.21	.23	.19	.34	.29	.28	1.0				1.37	1.35	1.41	2.52
9. CFMT1	.03	.12	.16	−.05	.17	.03	.10	.13	1.0			17.63	.88	−3.73	19.27
10. CFMT2	.14	.14	.19	.01	.19	.07	.11	.23	.47	1.0		22.43	5.05	−.53	−.49
11. CFMT3	.12	.03	.18	−.05	.14	.03	.13	.19	.38	.78	1.0	15.37	4.55	−.11	−.77

Note. N = 211; correlations greater than .13, $p < .05$; AVT = Advanced Vocabulary Test; CFS = culture fair intelligence test — series; CFM = culture fair intelligence test — matrices; NS = number series; MR = mental rotation; DSB = digit span backward; WSB = word span backward; VSB = visual span backward; CFMT1 = Cambridge Face Memory Test — same images; CFMT2 = Cambridge Face Memory Test — novel images; CFMT3 = Cambridge Face Memory Test — novel images with noise.

contrast, *Gf* and *Gsm* were associated with more substantial *g* loadings of $\lambda = .80$ ($\lambda^2 = .64$, $p = .001$) and $\lambda = .81$ ($\lambda^2 = .66$, $p = .002$), respectively. Finally, the VL latent variable was associated with *g* at $\lambda = .43$ ($\lambda^2 = .18$), $p = .001$.

For thoroughness, a single-factor model defined by all 11 subtests was tested and not found to be associated with acceptable levels of model-fit, $\chi^2(28, N = 211) = 48.94$, $p = .004$, CFI = .891, RMSEA = .065, BIC = 150.62. Furthermore, based on Raftery's (1995) guidelines, the higher-order model was found to be associated with better model fit than the single-factor model (ΔBIC = −4.74).

4. Discussion

The results of this investigation suggest that face recognition ability, as measured by the CFMT, was related positively to other cognitive abilities, including *Gf*, *Gsm*, and VL. Additionally, based on a higher-order model, face recognition ability was related to *g* moderately at .34. Thus, the hypothesis that face recognition ability would relate to *g* in manner similar to other memory span abilities was supported.

In contrast to previous studies (Davis et al., 2011; Palermo et al., 2013; Peterson & Miller, 2012), this investigation observed a moderate, positive association between face recognition ability and intelligence. Arguably, a statistically significant positive association was observed in this investigation, because the investigation was associated with a number of strengths. In particular, intelligence was measured with a battery of nine cognitive ability tests. Because general intelligence is a broad construct, one which is theorised to account for the positive correlations between a diversity of cognitive ability tests (Spearman, 1904), it is necessary to administer more than one or two cognitive ability tests to presume an adequate estimate of general cognitive functioning has been obtained (Jensen, 1998). Additionally, a relatively large sample size ($N = 200+$) increased the chances of rejecting the null hypothesis of no association, in comparison to previous investigations in the area ($N = 45$ to 80). In particular, it will be noted that the CFMT correlated with the Culture Fair Intelligence subtests at approximately .15 to .20, which requires a sample size of 200+ in order to achieve power of .80. Finally, the diverse combination of subtests and the relatively large sample size facilitated the implementation of latent variable modeling, a statistical technique which allows for the estimation of effects disattenuated for imperfect reliability (Fan, 2003).

Table 2
Latent Variable Correlations Associated with the Correlated Factor Model.

	VL	*Gf*	*Gsm*	FIR
VL	1.0			
Gf	.31	1.0		
Gsm	.36	.65	1.0	
FIR	.19	.29	.23	1.0

Note. N = 211; all latent variable correlations statistically significant ($p < .05$); VL = lexical knowledge ability; *Gf* = fluid intelligence; *Gsm* = short-term memory; FIR = face identity recognition ability.

4.1. Face recognition ability within the CHC model of intelligence

According to Newton and McGrew (2010), there may be as many as 100 stratum I abilities within the sampling domain of cognitive tests. Examples of stratum I abilities include lexical knowledge (VL), word fluency (FW), and quantitative reasoning (RQ). In comparison to stratum II abilities (e.g. *Gf*, *Gc*, *Gsm*, *Gs*), stratum I abilities tend to be relatively narrow in breadth. Theoretically, given the relative lack of breadth associated with face recognition ability, face recognition ability may be considered a stratum I ability within the CHC model of cognitive abilities (Carroll, 1993; McGrew, 2009). Although face recognition ability does not appear to have ever been considered within the CHC model previously, this may be the case only because measures of face recognition ability are not well-known to intelligence researchers.

In addition to the evidence reported in this investigation, there is also indirect evidence to suggest that individual differences in face recognition ability may be characterised as a cognitive ability similar to others found within the CHC model. For example, based on cross-sectional data with participants aged 18 to 89 ($N = 241$), Bowles et al. (2009) reported that performance on the CFMT begins to decrease between the ages of 35 and 55, and then decreases much more significantly after the age of 60. Correspondingly, in a very large-scale ($N = 10{,}394$) cross-sectional investigation (10 to 69 years of age), Hartshorne and Germine (2015) reported that digit span performance peaks at the age of approximately 30 years, and then decreases moderately until the age of approximately 60 years, after which reductions in performance are much more substantial. Thus, with respect to age effects, face recognition ability appears to behave in a manner similar to other *Gsm* abilities.

It should be acknowledged, however, that there may be evidence to suggest that face recognition ability may be, to some degree, distinct in its nature as a cognitive ability. In particular, it is interesting to note that, in this investigation, the face recognition ability latent variable did not associate with the *Gsm* latent variable ($r = .23$) to any degree stronger than its association with *Gf* ($r = .29$). Theoretically, one would expect a visual short-term memory task such as the CFMT to correlate more substantially with a short-term memory latent variable defined by digit, word, and visual span items, in comparison to a *Gf* latent variable. Thus, the results of this investigation suggest that face recognition ability may not be exclusively an indicator of the *Gsm* stratum II ability.

According to Newton and McGrew (2010), visual memory (MV) is an indicator of the visual processing (*Gv*) stratum II ability, rather than *Gsm* stratum II ability. Thus, it may be the case that face recognition ability is better aligned empirically with *Gv* than with *Gsm*. For reasons that

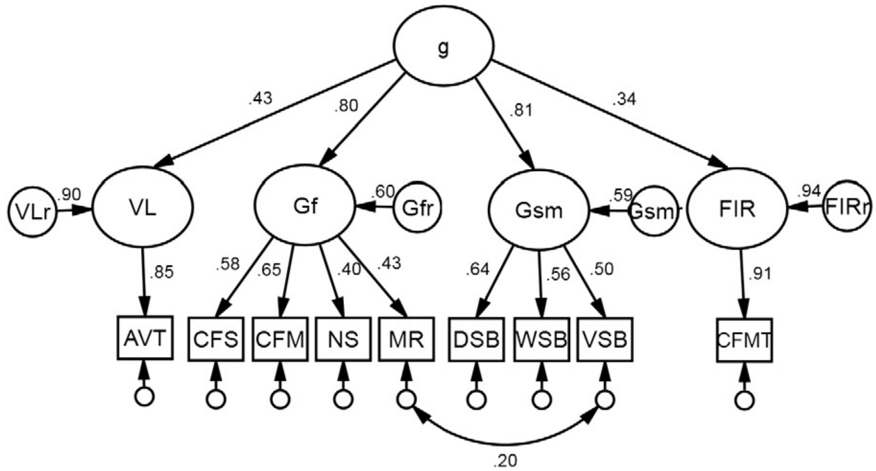

Fig. 1. Higher-order model with completely standardised factor loadings; all loadings and coefficients $p < .01$; g = general intelligence; VL = lexical knowledge ability; *Gf* = fluid intelligence; *Gsm* = short-term memory; FIR = face identity recognition ability; CFMT = Cambridge Face Memory Test (novel images); see Table 1 for full subtest spellings.

remain unknown and relatively unexplored, memory span tasks tend to inter-correlate with each other relatively weakly, in comparison to the inter-correlations between other tests from the same domain (Dang, Braeken, Ferrer, & Liu, 2012; Mackintosh & Bennett, 2003). Future research is encouraged to uncover more precisely the nature of face recognition ability's associations with stratum II abilities, as discussed in the limitations further below.

4.2. Implications for understanding prosopagnosia

Some researchers may have formed the impression that face recognition ability and general intelligence must be unrelated, given that that there are cases of prosopagnosia in individuals who are very intelligent (e.g., Oliver Sacks; Sacks, 2010). However, the phenomenon of prosopagnosia is entirely possible within a conventional multifactorial conception of intelligence.

Specifically, the *g* factor may be considered to be partly reflective of individual differences in the ability for sustained concentration (Lykken, 2005; Unsworth, 2015) — a capacity that may be expected to influence performance on all cognitive ability tests to varying degrees. However, in addition to *g*, there are sources of variance unique to stratum II and stratum I abilities that affect test score performance in a substantial manner. In the case of face recognition ability, the variance that was shared with *g* was equal to approximately 12%, in this investigation. By contrast, the reliable variance that was unique to face recognition ability, as measured by the CFMT, was approximately 88%. Within the higher-order model of intelligence depicted in Fig. 1, the reliable variance unique to face recognition ability was represented by the FIRr residual term associated with the face recognition first-order factor.

Theoretically, a very profound deficit in either *g* or FIRr may be expected to yield performance on the CFMT suggestive of a clinical disorder such as prosopagnosia (i.e. two standard deviations below the mean; Bowles et al., 2009). Furthermore, as *g* and FIRr are completely uncorrelated terms in the higher-order model, it is conceivable that an individual could have an isolated deficit in FIRr, but a largely intact *g*. Such an isolated deficit may occur when there is damage to an isolated part of the brain (e.g., fusiform gyrus; Furl, Garrido, Dolan, Driver, & Duchaine, 2011). In a simplistic way, a relatively high score associated with FIRr within the higher-order model may be considered to represent the health and integrity of a relatively specific portion of the brain principally responsible for mediating face recognition ability, whereas a high score on *g* may be considered to represent the health and functioning of a more distributed neuroanatomical system (Basten, Hilger, & Fiebach, 2015; Haier, White, & Alkire, 2003; Luders, Narr, Thompson, & Toga, 2009).

Finally, it will be noted that learning disabilities such as dyslexia (Vellutino, Fletcher, Snowling, & Scanlon, 2004), aculculia (specific inability to perform arithmetic operations; Benson & Weir, 1972) and alexia (word blindness; Greenblatt, 1973) are similar to prosopagnosia in that there is a profound impairment in a specific ability, but general intellectual functioning is relatively intact. The observation of such disabilities does not invalidate the contention that there is an association between *g* and a particular specific cognitive ability. Instead, an intact *g* is considered a necessary but not sufficient condition for the possibility of a learning disability diagnosis. That is, learning disabilities are defined as a substantial discrepancy between a specific ability and general intellectual functioning (Wilmshurst, 2012). If the moderate, positive correlation between face recognition ability and *g* observed in this investigation is found to be replicable, clinical diagnoses of prosopagnosia would be more justifiably based on a difference in CFMT performance and FSIQ performance equal to two standard deviations or more (CFMT < FSIQ). Such a recommendation would be consistent with the conventional approach used by clinicians to help with the diagnosis of specific learning disabilities (Wilmshurst, 2012).

4.3. Limitations

The sample used in this investigation was based principally upon first-year university students. It was anticipated that the sample would be somewhat more representative than the highly selected samples used in Davis et al. (2011) and Palermo et al. (2013). However, based on the CFIT scores, the sample was estimated to be associated with an approximate mean of 121 (normative sample mean = 100, SD = 15), which is very comparable to the levels reported by Davis et al. and Palermo et al. Therefore, it is difficult to specify with any confidence the degree to which the results reported in this investigation may generalise to the normal population. Unfortunately, corrections for range restriction were not feasible in this investigation, as not all of the subtests were associated with a sufficient amount of normative sample information. Future research in the area which includes participants within the lower spectrum of intelligence is sorely needed, as none of the current published work has included such participants.

Although the battery of cognitive ability subtests used in this investigation was much larger than what has been used in previous investigations with the CFMT, the higher-order model of intelligence was, nonetheless, far from complete. In particular, *Gc* was measured with only one stratum I ability (VL), rather than the more ideal three to four subtests per stratum II factors. Thus, the *g* factor was likely to some degree under-saturated with respect to *Gc* variance.

Additionally, the battery did not include any specific processing speed subtests, consequently, the well-known *Gs* stratum II factor was not included in the higher-order model. Based on previous inter-battery research (e.g., Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004), it may be suggested that the *g* factor measured in this investigation would correlate .90+ with a *g* factor defined by a more broad combination of stratum II factors. Consequently, the .34 association reported in this investigation between face recognition ability and *g* may be suggested to be essentially accurate, notwithstanding the high-ability samples used in this investigation. However, a very representative higher-order model of intelligence which included stratum II factors such as *Gv* (visual–spatial abilities), for example, would potentially facilitate a more detailed decomposition of face recognition ability's variance. Specifically, it is possible that face recognition ability, as measured by the CFMT, is composed of 15–20% *g*, 5–10% *Gv*, and 5–10% *Gsm* variance, with a remaining 60–75% construct specific variance (some of which would be measurement error). Although based on an arguably compromised face recognition ability subtest within the KABC-II (i.e., images which include non-intrinsic face characteristics), Potvin, Keith, Caemmerer, and Trundt (2015) reported a face recognition factor loading of .31 on a *Gv* first-order factor. The *Gv* first-order factor was associated with a second-order general factor at .95. Thus, based on a Schmid-Leiman decomposition, the face recognition subtest was associated with *g* at .29 and unique *Gv* (controlling for *g*) at .10.

It will also be acknowledged that face identity recognition ability was essentially measured with only one test in this investigation. Although the CFMT appears to be the most popular test in the area, a representation of face identity recognition ability based on several types of tests would be more ideal from both a psychometric and theoretical perspective. Furthermore, as the CFMT is a measure of short-term memory capacity, the implications of the results reported in this investigated are restricted to that dimension. It would be potentially useful to determine whether individual differences in face recognition ability for own-race and other-race faces (see Zhao & Bentin, 2008, for example) are correlated perfectly or not.[3] Additionally, a face recognition ability test based on photos of animal faces is conceivable. Finally, although the CFMT has been found to correlate positively ($r = .55$) with a long-term measure of face recognition ability (Famous Faces Memory Test; Wilmer et al., 2012), it is not necessarily the case that the results reported in this investigation would generalise to long-term memory for faces. Ultimately, if researchers in the area face recognition could agree upon three valid approaches to the measurement of face recognition ability, the possibility of a face recognition ability latent variable would be made available, which would allow for greater generalisation of results.

Finally, although the CFIT-Series and CFIT-Matrices subtests evidenced reasonable factorial validity in this investigation, the subtest scores were associated with rather poor levels of internal consistency reliability, as estimated via the conventional phi-based coefficient alpha. The upper-bound phi-based coefficient alpha estimates were much higher. Based on supplementary item-level analyses, there was some evidence to suggest that the last three items from each subtest were associated with weaker psychometric properties. Furthermore, anecdotal evidence through administration of the two subtests suggested that the recommended three minute time limit for each subtest created an excessively time-pressured testing scenario for some participants. Consequently, the association between face recognition ability *Gf* may have been attenuated to some degree for this reason. It will be noted that other investigations have reported low levels of internal consistency reliability for the CFIT. For example, Furlow, Armijo-Prewitt, Gangestad & Thornhill (1997) reported a coefficient alpha of only .46, based on a sample of 112 university students. Consequently, future research in the area should consider using alternative indicators of *Gf* such as Matrix Reasoning from the WASI and Raven's, for example.

[3] The CFMT is based on photos of Caucasian faces. Although race was not measured in this investigation, the Australian university from which the sample was drawn (University of Western Australia) is known to be substantially homogenous (i.e., Caucasian).

5. Conclusion

Individual differences in face recognition ability is an interesting construct, in part, because the task of recalling previously perceived faces is performed on a very regular basis in a typical person's day-to-day life. Furthermore, a high level of face recognition ability may be speculated to be associated with both social and career-related benefits, perhaps even independently of *g*. Researchers are encouraged to investigate such hypotheses (and more) with face recognition ability conceptualised as relatively narrow cognitive ability within the CHC model of intelligence.

References

Arbuckle, J. L. (2012). *Amos (version 21.0)*. Chicago: IBM SPSS.

Bachelder, B. L., & Denny, M. R. (1977). A theory of intelligence: I. Span and the complexity of stimulus control. *Intelligence*, *1*(2), 127–150.

Basten, U., Hilger, K., & Fiebach, C. J. (2015). Where smart brains are different: A quantitative meta-analysis of functional and structural brain imaging studies on intelligence. *Intelligence*, *51*, 10–27.

Benson, D. F., & Weir, W. F. (1972). Acalculia: Acquired anarithmetia. *Cortex*, *8*(4), 465–472.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., ... Yovel, G. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant–stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology*, *26*(5), 423–455.

Carey, S., Diamond, R., & Woods, B. (1980). Development of face recognition: A maturational component? *Developmental Psychology*, *16*(4), 257.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press.

Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 5–21). New York: Pergamon Press.

Cattell, R. B. (1963). *The IPAT culture fair intelligence scales 1, 2 and 3* (2nd ed.). Champaign, IL: Institute for Personality and Ability Test.

Cho, S. J., Wilmer, J., Herzmann, G., McGugin, R. W., Fiset, D., Van Gulick, A. E., ... Gauthier, I. (2015). Item response theory analyses of the Cambridge Face Memory Test (CFMT). *Psychological Assessment*, *27*(2), 552–566.

Dalrymple, K. A., & Palermo, R. (2016). Guidelines for studying developmental prosopagnosia in adults and children. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(1), 73–87.

Dang, C. P., Braeken, J., Ferrer, E., & Liu, C. (2012). Unitary or non-unitary nature of working memory? Evidence from its relation to general fluid and crystallized intelligence. *Intelligence*, *40*(5), 499–508.

Davis, J. M., McKone, E., Dennett, H., O'connor, K. B., O'Kearney, R., & Palermo, R. (2011). Individual differences in the ability to recognise facial identity are associated with social anxiety. *PloS One*, *6*(12), e28800.

Dennett, H. W., McKone, E., Edwards, M., & Susilo, T. (2012). Face aftereffects predict individual differences in face recognition ability. *Psychological Science*, *23*(11), 1279–1287.

Duchaine, B. (2011). Developmental prosopagnosia: Cognitive, neural and developmental investigations. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of psychology* (pp. 821–838) (1st ed.). Oxford, UK: Oxford University Press.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576–585.

Elliott, C. D. (1990). *Differential ability scales*. San Antonio, TX: Psychological Corporation.

Fan, X. (2003). Two approaches for correcting correlation attenuation caused by measurement error: implications for research practice. *Educational and Psychological Measurement*, *63*(6), 915–930.

Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus faces electivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, *23*(7), 1723–1740.

Furlow, F. B., Armijo-Prewitt, T., Gangestad, S. W., & Thornhill, R. (1997). Fluctuating asymmetry and psychometric intelligence. *Proceedings of the Royal Society of London B: Biological Sciences*, *264*(1383), 823–829.

Galper, R. E., & Hochberg, J. (1971). Recognition memory for photographs of faces. *The American Journal of Psychology*, 351–354.

Gignac, G. E. (2014). Fluid intelligence shares closer to 60% of its variance with working memory capacity and is a better indicator of general intelligence. *Intelligence*, *47*, 122–133.

Gignac, G. E. (2015). The magical numbers 7 and 4 are resistant to the Flynn effect: No evidence for increases in forward or backward recall across 85 years of data. *Intelligence*, *48*, 85–95.

Gignac, G. E., & Watkins, M. W. (2015). There may be nothing special about the association between working memory capacity and fluid intelligence. *Intelligence*, *52*, 18–23.

Gignac, G. E., & Weiss, L. G. (2015). Digit span is (mostly) related linearly to general intelligence: Every extra bit of span counts. *Psychological Assessment*, *27*(4), 1312–1323.

Gonzalez, R., & Griffin, D. (2001). Testing parameters in structural equation modeling: Every "one" matters. *Psychological Methods*, *6*(3), 258–269.

Greenblatt, S. H. (1973). Alexia without agraphia or hemianopsia. *Brain*, *96*(2), 307–316.

Haier, R. J., White, N. S., & Alkire, M. T. (2003). Individual differences in general intelligence correlate with brain function during nonreasoning tasks. *Intelligence*, *31*(5), 429–441.

Hartshorne, J. K., & Germine, L. T. (2015). *When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life*. Psychological Science: Span (in press).

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

Hildebrandt, A., Wilhelm, O., Schmiedek, F., Herzmann, G., & Sommer, W. (2011). On the specificity of face cognition compared with general cognitive functioning across adult age. *Psychology and Aging*, *26*(3), 701–715.

Irwing, P., Booth, T., Nyborg, H., & Rushton, J. P. (2012). Are *g* and the general factor of personality (GFP) correlated? *Intelligence*, *40*(3), 296–305.

Jackson, D. N. (1998). *Multidimensional aptitude battery–II: manual*. Port Huron, MI: Sigma Assessment Systems.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Greenwood, Westport, CT: Praeger.

Jensen, A. R., & Figueroa, R. A. (1975). Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology*, *67*(6), 882–893.

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.

Jöreskog, K. G., & Sörbom, A. D. (1982). Recent developments in structural equation modeling. *Journal of Marketing Research*, *19*, 404–416.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children—2nd ed*. Circle Pines, MN: AGS.

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1118–1133.

Luders, E., Narr, K. L., Thompson, P. M., & Toga, A. W. (2009). Neuroanatomical correlates of intelligence. *Intelligence*, *37*(2), 156–163.

Lykken, D. T. (2005). Mental energy. *Intelligence*, *33*(4), 331–335.

Mackintosh, N. J., & Bennett, E. S. (2003). The fractionation of working memory maps onto different components of intelligence. *Intelligence*, *31*(6), 519–531.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.

Miller, L. T., & Vernon, P. A. (1992). The general factor in short-term memory, intelligence, and reaction time. *Intelligence*, *16*(1), 5–29.

Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: Current research in Cattell–Horn–Carroll-based assessment. *Psychology in the Schools*, *47*(7), 621–634.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw–Hill.

Osterrieth, P. A. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie*, *30*, 206–356.

Palermo, R., O'Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion and their association with ability to recognise vocal emotions and facial identity. *PloS One*, *8*(6), 1–16.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain and Cognition*, *28*(1), 39–58.

Peterson, E., & Miller, S. F. (2012). The eyes test as a measure of individual differences: how much of the variance reflects verbal IQ? *Frontiers in Psychology*, *3*, 220. http://dx.doi.org/10.3389/fpsyg.2012.00220.

Potvin, D. C., Keith, T. Z., Caemmerer, J. M., & Trundt, K. M. (2015). Confirmatory factor structure of the Kaufman assessment battery for children—Second edition with preschool children too young for differentiation? *Journal of Psychoeducational Assessment*, *33*(6), 522–533.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Raven, J. C., Court, J. H., & Raven, J. (1979). *Manual for Raven's progressive matrices and vocabulary scales*. London: H. K. Lewis & Co.

Rey, A. (1941). L'examen psychologique dans les cas d'encépalopathie traumatique. *Archives de Psychologie*, *28*, 286–340.

Reynolds, M. R., Keith, T. Z., Fine, J. G., Fisher, M. E., & Low, J. A. (2007). Confirmatory factor structure of the Kaufman assessment battery for children—Second edition: Consistency with Cattell-Horn-Carroll theory. *School Psychology Quarterly*, *22*(4), 511–539.

Reynolds, M. R., Keith, T. Z., Flanagan, D. P., & Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology*, *51*(4), 535–555.

Rhodes, G., Jeffery, L., Taylor, L., Hayward, W. G., & Ewing, L. (2014). Individual differences in adaptive coding of face identity are linked to individual differences in face recognition ability. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 897–903.

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review*, *16*(2), 252–257.

Sacks, O. (2010). August). Face-blind: why are some of us terrible at recognizing faces? *The New Yorker*, *36–43*.

Sanders, S., McIntosh, D. E., Dunham, M., Rothlisberg, B. A., & Finch, H. (2007). Joint confirmatory factor analysis of the differential ability scales and the Woodcock–Johnson Tests of cognitive abilities—Third edition. *Psychology in the Schools*, *44*(2), 119–138.

Sekiguchi, T. (2011). Individual differences in face memory and eye fixation patterns during face learning. *Acta Psychologica*, *137*(1), 1–9.

Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences*, *112*(41), 12887–12892.

Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–292.

Stankov, L. (1997). *The Gf/Gc Quickie Test Battery. Unpublished test battery from the School of Psychology*. Australia: University of Sydney.

Sun, W., Chou, C. -P., Stacy, A. W., Ma, H., Unger, J., & Gallaher, P. (2009). SAS and SPSS macros to calculate standardized Cronbach's alpha using the upper bound of the phi coefficient for dichotomous items. *Behavior Research Methods*, *39*(1), 71–81.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*(2), 225–245.

Unsworth, N. (2015). Consistency of attentional control as an important cognitive trait: A latent variable analysis. *Intelligence*, *49*, 110–128.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2–40.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence (WASI)*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale/Wechsler memory scale-fourth edition: Administration and scoring manual*. San Antonio, TX: Pearson Assessment.

Wechsler, D. (2014). *Wechsler intelligence scale for children—Fifth edition*. San Antonio, TX: Pearson Assessment.

Wilhelm, O., Herzmann, G., Kunina, O., Danthiir, V., Schacht, A., & Sommer, W. (2010). Individual differences in perceiving and recognizing faces—One element of social cognition. *Journal of Personality and Social Psychology*, *99*(3), 530–548.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Gerbasi, M., & Nakayama, K. (2012). Capturing specific abilities as a window into human individuality: The example of face recognition. *Cognitive Neuropsychology*, *29*(5–6), 360–392.

Wilmer, J. B., Germine, L. T., & Nakayama, K. (2014). Face recognition: A model specific ability. *Frontiers in Human Neuroscience*, *8*, 769.

Wilmshurst, L. (2012). *Clinical and educational psychology: An ecological transactional approach to understanding child problems and interventions*. Hoboken: Wiley.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III tests of cognitive abilities*. Itasca, IL: Riverside.

Zhao, L., & Bentin, S. (2008). Own-and other-race categorization of faces by race, gender, and age. *Psychonomic Bulletin & Review*, *15*(6), 1093–1099.