Intelligence 71 (2018) 85–91

Contents lists available at ScienceDirect

Intelligence

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

General Intelligence (*g*), ACT Scores, and Theory of Mind: (ACT)*g* Predicts Limited Variance Among Theory of Mind Tests

Thomas R. Coyle[a,*], Karrie E. Elpers[a], Miguel C. Gonzalez[a], Jacob Freeman[b], Jacopo A. Baggio[c,d]

[a] Department of Psychology, University of Texas at San Antonio, USA
[b] Department of Sociology, Social Work, and Anthropology, Utah State University, USA
[c] Department of Political Science, University of Central Florida, USA
[d] Sustainable Coastal Systems Cluster, National Center for Integrated Coastal Research, University of Central Florida, USA

ARTICLE INFO

Keywords:
General intelligence (*g*)
ACT
Theory of mind
Non-*g* residuals

ABSTRACT

This study is the first to examine relations between general intelligence (*g*), non-*g* factors, and theory of mind (ToM) using structural equation modeling with multiple indicators of *g* and ToM. *g* was based on the subtests of the ACT, a college admissions test that is strongly *g* loaded, and ToM was based on the Reading the Mind in the Eyes Test and the Short Story Test (SST). *g* correlated strongly with a latent ToM factor ($\beta = .65$) and moderately with the two ToM tests ($\beta \approx .34$), which correlated modestly with each other ($\beta = .27$). The modest correlation between the ToM tests indicates that *g* predicted a small amount of variance among the ToM tests (7%) and suggests that the ToM tests had little in common. In addition, non-*g* residuals of the ACT subtests, obtained after removing *g*, correlated negligibly with the ToM factor and the ToM tests ($|\beta| < 0.06$). Similar results were obtained for the ToM residuals, which correlated trivially with the ACT subtests. The trivial non-*g* effects suggest that *g*-ToM relations were attributable to "not much more than *g*." The results replicated with different combinations of ACT subtests, controls for possible confounds (reading comprehension on the SST), and another college admissions test (the SAT). The use of a convenience sample (college students) and the limited measures of *g* and ToM are discussed as limitations. Future research should examine the robustness of effects using different measures of *g* and ToM and also examine possible mediators of *g*-ToM relations (e.g., executive functions).

1. Introduction

This study examines relations between general intelligence (*g*) and theory of mind (ToM), two constructs that appear conceptually distinct but are empirically related in adults (e.g., Baker, Peterson, Pulos, & Kirkland, 2014; see also, Botting & Conti-Ramsden, 2008; Estes & Bartsch, 2017; Fernández-Berrocal, Cabello, & Gutiérrez-Cobo, 2017; Henry et al., 2009). Whereas *g* reflects the ability to handle complexity and novelty (Gottfredson, 1997), ToM reflects the ability to infer the thoughts and feelings of other people (Baron-Cohen, 1995). Both *g* and ToM are needed to navigate complex interactions in diverse situations (e.g., Astington, 1998; Gottfredson, 1997) and both constructs are related to diverse cognitive abilities (e.g., executive functions) (e.g., Royall & Palmer, 2014; see also, Bull, Phillips, & Conway, 2008; Peterson & Miller, 2012).

The aim of the study was to characterize relations of ToM with *g* and non-*g* factors. *g* represents variance common to cognitive tests and largely explains the predictive power of tests (e.g., Jensen, 1998, pp. 270–305). Non-*g* factors represent factors unrelated to *g* and include non-*g* residuals of tests, obtained after removing *g*. Non-*g* factors measure specific abilities such as math or verbal abilities on scholastic aptitude tests (e.g., Coyle, Purcell, Snyder, & Kochunov, 2013; see also, Coyle & Pillow, 2008). Compared to *g*, non-*g* factors generally have weak predictive validity for work and school outcomes (e.g., Jensen, 1998, pp. 270–305).

A relationship between *g* and ToM would be predicted by Spearman's principle of the indifference of the indicator (Jensen, 1998, pp. 32–34). The principle assumes that tests of all cognitive abilities are indicators of *g* (to varying degrees), regardless of the content or types of tests. The principle is bolstered by factor analysis of diverse tests, which typically correlate more strongly with *g* than with any other factor (e.g., Jensen, 1998, pp. 73–81). Spearman's principle suggests that *g* should be linked to ToM because tests of ToM correlate with tests of executive functions and other diverse abilities, which are indicators of *g* (e.g., Royall & Palmer, 2014; see also, Bull et al., 2008; Peterson & Miller, 2012).

Spearman's principle of the indifference of the indicator is related to what might be called the primacy of *g* hypothesis (cf. Jensen, 1984; see also, Reeve & Charles, 2008). The hypothesis assumes that relations among

* Corresponding author at: University of Texas at San Antonio, Department of Psychology, One UTSA Circle, San Antonio, TX 78249, USA.
E-mail address: thomas.coyle@utsa.edu (T.R. Coyle).

https://doi.org/10.1016/j.intell.2018.10.006
Received 29 December 2017; Received in revised form 15 October 2018; Accepted 15 October 2018

cognitive tests are primarily explained by *g* and that non-*g* factors (obtained after removing *g*) have limited effects. The primacy of *g* hypothesis extends Spearman's principle by identifying *g* as the primary source of variance that contributes to relations among tests. Support for the hypothesis comes from research showing that *g* explains most of the variance among diverse tests (e.g., verbal, math, spatial) and that non-*g* factors have relatively weak effects (e.g., Jensen, 1998, pp. 107–144). Based on the primacy of *g* hypothesis, non-*g* factors of tests were expected to correlate relatively weakly with ToM, which was expected to correlate strongly with *g*.

g and ToM were measured using multiple tests. *g* was based on the subtests of the ACT (formerly, American College Test), a college admissions test widely used in the United States. The ACT correlates moderately with college grade point average (GPA) (e.g., $r_{\text{ACT-GPA}} = 0.41$, Coyle & Pillow, 2008, p. 724) and strongly with other scholastic aptitude tests such as the PSAT (Preliminary SAT) and the SAT (formerly, Scholastic Aptitude Test) (e.g., $r_{\text{ACT-SAT}} = 0.87$, Coyle & Pillow, p. 723). The ACT also correlates strongly with IQ tests and with a *g* based on the Armed Services Vocational Aptitude Battery (e.g., $r = 0.77$, Koenig, Frey, & Detterman, 2008), a battery of 10 diverse tests. *g* factors based on diverse tests typically correlate strongly, and sometimes perfectly, with each other (e.g., $r \approx 1.00$, Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004, p. 103; see also, Johnson, te Nijenhuis, & Bouchard, 2008). In addition, the ACT includes math and verbal subtests (reading and English), which measure both *g* and specific abilities. The distinction between *g* and specific abilities is central to research linking ToM to verbal abilities (e.g., language, vocabulary, reading skill) but not to non-verbal abilities (e.g., Peterson & Miller, 2012; see also, Milligan, Astington, & Dack, 2007; Olderbak et al., 2015), suggesting that verbal abilities may be especially sensitive to ToM.

ToM was measured using the Reading the Mind in the Eyes Test (RMET; Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), and the Short Story Test (SST; Dodell-Feder, Lincoln, Coulson, & Hooker, 2013). The RMET is a social-perceptual test of ToM that presents participants with different pairs of eyes and asks them to identify the emotion being expressed (Baron-Cohen et al., 1997, 2001; Nettle & Liddle, 2008). In contrast, the SST is a social-cognitive test of ToM that presents participants with a short story and asks them to infer the thoughts and emotions of the characters (Dodell-Feder et al., 2013).

The two ToM tests are assumed to reflect different types of processing. The RMET is assumed to tap relatively automatic processing (Baron-Cohen et al., 2001), which is fast, intuitive, and effortless. In contrast, the SST is assumed to tap relatively effortful processing (Dodell-Feder et al., 2013), which is slower, reflective, and controlled. The SST and RMET correlate robustly with each other ($r = 0.49$, Dodell-Feder et al., 2013, p. 7), reflecting the fact that both tests tap ToM but measure different types of processing. In addition, both tests correlate positively with IQ and other *g* loaded measures (e.g., $rs \approx 0.24$; Dodell-Feder et al., 2013, p. 7; see also, Baker et al., 2014, p. 82).

The SST and RMET are among the most common ToM tasks for adults (e.g., Henry, Phillips, Ruffman, & Bailey, 2013, pp. 838–839). ToM tasks can be differentiated by type of task (e.g., Stories and Eyes), modality of task (e.g., verbal and visual), and mental process (e.g., cognitive and affective) (Henry et al., 2013, pp. 229–230). The SST is a “Stories” task, which measures the ability to infer cognitive or affective states of characters in a short story. The RMET is an “Eyes” task, which measures the ability to infer mental states (predominantly affective) from pictures of eyes. The RMET and SST predict social communication disorders such as autism spectrum disorder (e.g., Baron-Cohen et al., 2001), and social criteria such as faux pas, social communication errors, and mental states of fantasy characters (e.g., Dodell-Feder et al., 2013; Henry et al., 2013).

The current study differs from prior studies of ToM and intelligence in important ways. First, whereas prior studies have focused on the RMET or SST (Baker et al., 2014; Baron-Cohen et al., 2001), the current study estimated a latent ToM factor based on both tests. This latent ToM factor was correlated with a latent *g* factor, which was also correlated with the individual ToM tests. The approach provided an estimate of the relation between *g* and the ToM factor and between *g* and each ToM test.

Second, whereas prior studies have correlated measures of *g* (e.g., IQ) with ToM (e.g., Baker et al., 2014; Dodell-Feder et al., 2013), the current study also correlated the non-*g* residuals of tests, obtained after removing *g*, with ToM. In particular, the non-*g* residuals of the math and verbal ACT subtests were correlated with ToM. These non-*g* residuals measured the unique variance of the subtests, obtained after removing *g*. Unlike the non-*g* residuals of other tests, the non-*g* residuals of the ACT predict diverse criteria such as college GPAs, college majors, and jobs (Coyle & Pillow, 2008; Coyle et al., 2013; Coyle, Purcell, Snyder, & Richmond, 2014; Coyle, Snyder, Richmond, & Little, 2015). In particular, math residuals predict achievements and interests in science, technology, engineering, and math (STEM), while verbal residuals predict criteria in the humanities (e.g., English, art, music) (e.g., Coyle et al., 2015; see also, Coyle et al., 2013, 2014). Although non-*g* residuals typically have little or no predictive validity (Jensen, 1998, p. 30), the effect sizes of ACT non-*g* residuals for STEM and humanities criteria are substantial ($\beta \approx 0.30$, Coyle et al., 2013, p. 118).

The non-*g* residuals of the ACT permitted a test of the specificity doctrine, which assumes that specific skills and abilities predict theoretically related criteria beyond *g* (cf. Jensen, 1984). The non-*g* residuals of the ACT subtests represent specific math and verbal abilities, obtained after removing *g*. The verbal residuals of the ACT were expected to predict ToM, which has been linked to diverse language and verbal abilities (e.g., vocabulary and reading comprehension) (e.g., Peterson & Miller, 2012; see also, Milligan et al., 2007; Olderbak et al., 2015).

In the current study, the ToM measures (RMET and SST) were correlated with a *g* based on the ACT subtests. All measures were drawn from a larger study of relations among general intelligence, social intelligence, and cooperation. The relations were estimated using structural equation modeling (SEM). By convention, *g* (variance common to tests) was estimated using all tests, while a latent ToM factor (loading on *g*) was estimated using the ToM tests (RMET and SST). SEM estimated relations between *g*, the ToM factor, and the ToM tests, which were also correlated with the non-*g* residuals (i.e., unique variances) of the ACT subtests.

Based on prior research linking measures of *g* with ToM (e.g., Baker et al., 2014; Dodell-Feder et al., 2013), *g* was expected to correlate positively with the ToM factor. In addition, *g* was expected to correlate more strongly with the SST, a social-cognitive measure of ToM, than with the RMET, a social-perceptual measure of ToM. The latter prediction was based on the assumption that the SST and *g* primarily measure cognitive processing, whereas the RMET primarily measures perceptual processing (cf. Nettle & Liddle, 2008). In addition, the non-*g* residuals of the ACT subtests, obtained after removing *g*, were expected to show a pattern of differential validity. Based on research linking ToM with verbal ability (e.g., Peterson & Miller, 2012), the verbal residuals were expected to correlate (relatively) strongly with ToM, whereas the math residuals were expected to correlate weakly (and non-significantly) with ToM. Such a pattern would be the first to demonstrate that a *g* based on a widely used college admissions test (the ACT) predicts ToM, and that the verbal residuals of the test uniquely predict ToM.

2. Method

2.1. Subjects

Subjects were 374 college students who participated in a study of human intelligence, social intelligence, and cooperation.[1] 249 subjects had scores for three ACT subtests (math, reading, English), which were used in the primary analyses. A small number of subjects (29 of 374) had RMET and SST scores but no ACT scores and another set of subjects (96 of 374) had SAT scores for only two subtests (math and verbal). (The SAT is another college admissions

[1] The current study used only the measures of *g* (ACT subtests) and ToM (RMET, SST) from the original study, which also included measures of cooperation on a multiplayer video game. The measures of cooperation were not relevant for the current study and were excluded from the analyses.

test used in the United States.) Because three or more subtests were needed to identify the SEM model and estimate the relations for non-*g* residuals, the SAT data were not used in the primary analyses. To address the possibility that missing data distorted the estimates in the primary analyses (using only cases with ACT scores), supplemental analyses were performed using the full sample with all 374 subjects (see *Statistical analyses*).

2.2. Variables

The RMET and SST were presented on computers to small groups of subjects, with 8 to 20 subjects per group. ACT scores were obtained from university records. Descriptive statistics for all variables are reported in Table 1.

ACT scores (possible range = 0 to 36) were available for thee subtests: math (ACTm), reading (ACTr), and English (ACTe) ($N = 249$). The reading and English subtests were used to measure verbal ability. ACT scores showed no signs of distributional anomalies (|skewness| < 0.12, |kurtosis| < 0.79) (Table 1).

RMET scores (possible range = 0 to 36) were available for all subjects ($N = 374$). The RMET was based on the task developed by Baron-Cohen et al. (2001). Subjects were presented with 36 pairs of eyes (one at a time) and were shown four verbal descriptions of the emotion conveyed by the eyes (e.g., happy, sad, angry, depressed). Subjects had to select the emotion conveyed by each pair of eyes.

SST scores (possible range = 0 to 16) were missing for eight subjects ($N = 366$). (Missing data were due to response omissions.) The SST was based on the task developed by Dodell-Feder et al. (2013). Subjects had to a read a short story by Ernest Hemingway (*The End of Something*) and answer questions about the mental states (e.g., beliefs, intentions) of the characters in the story. In addition, subjects had to answer a series of questions about non-mental states in the stories, which were designed to control for reading comprehension. (The comprehension questions were used to check the robustness of results in the primary analyses.) SST and RMET scores showed no signs of distributional anomalies (|skewness| < 0.55, |kurtosis| < 0.35) (Table 1).

2.3. Statistical analyses

Statistical analyses were divided into primary and supplemental analyses. The primary analyses analyzed subjects with ACT scores ($N = 249$). Supplemental analyses analyzed all subjects ($N = 374$), including subjects with no ACT scores or ToM scores (RMET, SST), as well as a small sample of subjects with SAT scores ($N = 96$). Missing data were handled using full information maximum likelihood (FIML).

SEM with maximum likelihood estimation estimated relations between *g* and ToM. By convention, *g* was based on all tests and measured variance common to all tests. A ToM factor measured variance common to the ToM tests. Fig. 1 depicts the base model with ACT scores and ToM tests.[2] In this model, the three ACT subtests (ACTm, ACTr, ACTe) loaded directly on *g*, and the two ToM tests (RMET, SST) loaded on the ToM factor, which in turn loaded on *g*.

The fit of the base model (Fig. 1) was tested using subjects with ACT scores ($N = 249$). Model fit was evaluated using the chi square test (χ^2),

Table 1
Means and standard deviations for all variables.

	N	*M*	*SD*	Min	Max	Skewness	Kurtosis
Primary analyses							
ACTm	249	22.56	4.43	13.00	34.00	0.12	−0.68
ACTr	249	24.83	5.66	12.00	36.00	−0.03	−0.79
ACTe	249	24.02	5.41	10.00	36.00	0.06	−0.28
RMET	249	27.55	3.51	18.00	34.00	−0.55	−0.35
SST	245	9.29	2.61	1.00	15.00	−0.49	0.05
Supplemental analyses							
RMET	374	27.52	3.39	18.00	35.00	−0.47	−0.31
SST	366	9.24	2.66	0.00	15.00	−0.55	0.16

Note. Primary Analyses = subjects with ACT scores. Supplemental Analyses = all subjects. The descriptive statistics for the ACT scores were the same in both analyses and therefore are reported only in the Primary Analyses. Min = minimum observed score. Max = maximum observed score. ACTm = ACT math. ACTr = ACT reading. ACTe = ACT English. RMET = Reading the Mind in the Eyes Test. SST = Short Story Test.

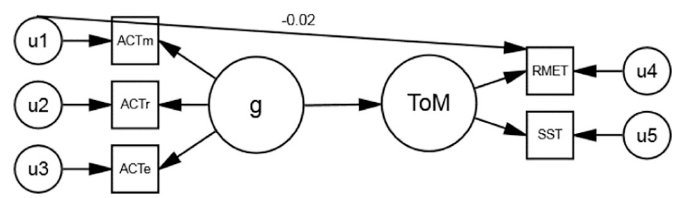

Fig. 1. SEM model with *g* and ToM factor. Correlations were computed between *g* and the ToM factor, and between the non-*g* residuals of the ACT subtests (u1, u2, u3) and the ToM variables. The model depicts u1→RMET effect (β = −0.02), which represents the path coefficient between the ACTm residual and RMET (Table 4). (The base model omitted the u1→RMET path.)

comparative fit index (CFI), root mean square error of approximation (RMSEA), and Akaike information criterion (AIC) (Kline, 2005, pp. 133–145). Model fit for the base model was excellent ($\chi^2 = 0.35$, $df = 4$, $p = .99$; CFI = 1.00, RMSEA = 0.00, AIC = 32.35) and was better than for a model in which the ToM tests loaded directly on *g* with no ToM factor ($\chi^2 = 7.75$, $df = 5$, $p = .17$; CFI = 0.994, RMSEA = 0.047, AIC = 37.75).

The path from *g* to ToM (*g*→ToM) estimated relations between *g* and ToM (Fig. 1). In addition, relations among the non-*g* residuals of the ACT math and verbal subtests, obtained after removing *g*, were estimated by adding a path from a subtest's residuals (u's in Fig. 1) to a ToM variable (ToM factor, RMET, SST). Based on prior research (e.g., Peterson & Miller, 2012), the non-*g* effects were expected to be relatively strong (and significant) for the verbal residuals but weak (and nonsignificant) for the math residuals. All effects are reported as standardized coefficients, with the mean of multiple coefficients reported in parentheses (e.g., M_r and M_β). Following Cohen (1988, pp. 77–81), effects (*r* or β) around 0.10, 0.30, and 0.50 were labeled weak, moderate, and strong, respectively. Significant effects are reported at $p < .05$.

3. Results

3.1. Preliminary analyses

Table 2 reports bivariate correlations between the five manifest variables (ACTm, ACTr, ACTe, RMET, SST) using data from subjects with ACT scores ($N = 249$), which were also used in the primary analyses. The correlations showed positive manifold (i.e., positive correlations) ($M_r = 0.39$), which is the basis of *g* (i.e., variance common to tests). The positive correlations indicate that people who performed well on one test generally performed well on all others. ACT scores correlated more strongly with each other ($M_r = 0.67$) than with the ToM tests ($M_r = 0.28$). The ToM tests correlated moderately with each other ($r = 0.26$) and with the ACT subtests ($M_r = 0.23, 0.28, 0.32$, math, reading, and English, respectively). The

[2] The base model (Fig. 1) was supported by a factor analysis (principal axis factoring, direct oblimin rotation) of ACT scores (ACTm, ACTr, ACTe) and ToM tests (RMET, SST), using subjects with ACT scores ($N = 249$). The results yielded a large first factor (Factor 1) (eigenvalue = 2.65, variance = 53%), and a second factor (Factor 2) accounting for non-trivial variance (eigenvalue = 0.95, variance = 19%). The *g* and ToM factors were supported by the pattern of factor loadings. *g* was supported by the Factor 1 loadings, which were positive and robust (loadings > 0.30) but somewhat lower for the ToM tests (loadings = 0.70, 0.95, 0.79, 0.37, 0.34, ACTm, ACTr, ACTe, RMET, SST, respectively). A ToM factor was supported by the Factor 2 loadings, which were robust for the ToM tests but negligible for the ACT subtests (loadings = −0.10, −0.13, −0.08, 0.35, 0.35, ACTm, ACTr, ACTe, RMET, SST, respectively).

Table 2
Correlations among manifest variables.

	1	2	3	4	5
1. ACTm	–				
2. ACTr	0.56	–			
3. ACTe	0.69	0.77	–		
4. RMET	0.24	0.30	0.34	–	
5. SST	0.22	0.25	0.30	0.26	–

Note. All correlations $p < .01$. Correlations computed using only subjects with ACT scores ($N = 249$). ACTm = ACT math. ACTr = ACT reading. ACTe = ACT English. RMET = Reading the Mind in the Eyes Test. SST = Short Stories Test.

Table 3
SEM implied correlations among g, ToM factor, ToM tests, and ACT subtests.

	1	2	3	4	5	6	7
1. *g*	–						
2. ToM	0.65	–					
3. RMET	0.35	0.55	–				
4. SST	0.32	0.49	0.27	–			
5. ACTm	0.71	0.46	0.25	0.23	–		
6. ACTr	0.80	0.51	0.28	0.25	0.57	–	
7. ACTe	0.97	0.62	0.34	0.31	0.69	0.77	–

Note. All correlations $p < .01$. Correlations estimated using SEM and subjects with ACT scores ($N = 249$). g = latent factor based on all tests. ToM = latent factor based on RMET and SST. RMET = Reading the Mind in the Eyes Test. SST = Short Story Test. ACTm = ACT math. ACTr = ACT reading. ACTe = ACT English.

Table 4
SEM relations between g, non-g residuals of ACT subtests, and ToM variables.

Analysis and effect	*N*	Effect	χ^2 (*df*, *p*)	CFI	RMSEA (*p*)	AIC
Primary analyses ($N = 249$)						
1. *g*→ToM	249	0.65	0.35 (4, 0.99)	1.00	0.00 (1.00)	32.35
2. u ACTm→ToM	249	−0.03	0.25 (3, 0.97)	1.00	0.00 (0.99)	34.25
3. u ACTr→ToM	249	0.05	0.09 (3, 0.99)	1.00	0.00 (1.00)	34.09
4. u ACTe→ToM	249	−0.06	0.30 (3, 0.96)	1.00	0.00 (0.98)	34.30
5. u ACTm→RMET	249	−0.02	0.27 (3, 0.97)	1.00	0.00 (0.99)	34.27
6. u ACTr→RMET	249	0.03	0.11 (3, 0.99)	1.00	0.00 (1.00)	34.11
7. u ACTe→RMET	249	−0.03	0.30 (3, 0.96)	1.00	0.00 (0.98)	34.30
8. u ACTm→SST	249	−0.01	0.34 (3, 0.95)	1.00	0.00 (0.98)	34.34
9. u ACTr→SST	249	0.01	0.33 (3, 0.95)	1.00	0.00 (0.98)	34.33
10. u ACTe→SST	249	−0.01	0.35 (3, 0.95)	1.00	0.00 (0.98)	34.35
Supplemental analyses ($N = 374$)						
11. *g*→ToM	374	0.64	0.35 (4, 0.99)	1.00	0.00 (1.00)	32.35
12. u ACTm→ToM	374	−0.03	0.25 (3, 0.97)	1.00	0.00 (0.99)	34.25
13. u ACTr→ToM	374	0.05	0.09 (3, 0.99)	1.00	0.00 (1.00)	34.09
14. u ACTe→ToM	374	−0.06	0.30 (3, 0.96)	1.00	0.00 (1.00)	34.30
15. u ACTm→RMET	374	−0.02	0.27 (3, 0.97)	1.00	0.00 (0.97)	34.27
16. u ACTr →RMET	374	0.03	0.12 (3, 0.99)	1.00	0.00 (1.00)	34.12
17. u ACTe→RMET	374	−0.03	0.30 (3, 0.96)	1.00	0.00 (1.00)	34.30
18. u ACTm→SST	374	−0.01	0.34 (3, 0.95)	1.00	0.00 (0.98)	34.34
19. u ACTr→SST	374	0.01	0.33 (3, 0.95)	1.00	0.00 (0.98)	34.34
20. u ACTe→SST	374	−0.01	0.35 (3, 0.95)	1.00	0.00 (0.98)	34.35

Note. All non-*g* residual effects (u's) are nonsignificant. Primary Analyses = subjects with ACT scores. Supplemental Analyses = all subjects. g = variance common to all tests. ToM = variance common to ToM tests. u = non-*g* residual (unique) variance of ACT subtests, obtained after removing g. ACTm = ACT math. ACTr = ACT reading. ACTe = ACT English. RMET = Reading the Mind in the Eyes Test. SST = Short Story Test.

modest correlation between the ToM tests indicates that the ToM tests shared a small amount of common variance (7%).

3.2. Primary analyses

Table 3 reports the SEM implied correlations between g, ToM factor, ToM tests (RMET, SST), and ACT subtests (ACTm, ACTr, ACTe). The correlations were estimated using the SEM model (Fig. 1) with cases with ACT scores ($N = 249$). Consistent with predictions, g correlated strongly with the ToM factor ($\beta = 0.65$) and moderately with the ToM tests ($M_\beta = 0.34$). There was no discernable difference in the relation of g with the SST ($\beta = 0.32$) and RMET ($\beta = 0.32$). In addition, the ToM factor correlated strongly with the ToM tests ($M_\beta = 0.52$) and with all ACT subtests ($M_\beta = 0.53$). The ToM factor and ToM tests correlated moderately with the verbally-loaded ACT reading and English subtests ($M_\beta = 0.39$) and with the ACT math subtest ($M_\beta = 0.31$).

Table 4 (Analysis 1–10) reports the SEM relations between the non-*g* residuals of the ACT subtests, ToM factor, and ToM tests. (Analysis 1 reports the effect between g and ToM, which was also reported in Table 3.) The analyses were performed separately for the ToM variables (ToM factor, RMET, SST) and ACT subtests (ACTm, ACTr, ACTe), yielding nine effects (3 ToM variables x 3 ACT residuals). Contrary to predictions, the ToM factor and ToM tests were negligibly (and non-significantly) related to the non-*g* residuals of all ACT subtests ($M_\beta = -0.01$).[3,4] In addition, all ToM variables correlated negligibly with the residuals of the ACT math subtest ($M_\beta = -0.02$) and with the residuals of the ACT reading and English subtests ($M_\beta < -0.01$).[5,6]

[3] The sample was comprised of students who were admitted to college and were likely to have high ACT scores (compared to non-admitted students), which could produce range restriction and attenuate effects. To address this possibility, the SEM model (Fig. 1) analyzed disattenuated ACT correlations, which were corrected for range restriction using the standard deviations for the ACT normative sample (Wiberg & Sundström, 2000; see also, Lawley, 1943; Ree, Carretta, Earles, & Albert, 1994). The standard deviations for the normative sample were 5.5, 6.6, and 6.9, for ACTm, ACTr, and ACTe, respectively (ACT, 2017a), which were higher than those for the current sample (cf. Table 1). The results extended the prior analyses (cf. Table 4). The *g*-ToM effect increased in strength (0.79 vs. 0.65), owing to the disattenuated correlations. In addition, the relations of the ACT non-*g* residuals with the ToM variables (ToM factor, RMET, SST) were still negligible for the ACTm (−0.02, −0.02, 0.01), ACTr (−0.01, 0.01, −0.02), and ACTe (0.07, 0.03, 0.02). The average difference in analogous effects between the supplemental and prior analyses (cf. Table 4) was trivial (*M* difference = 0.03).

[4] Supplemental analyses examined g and non-*g* effects after correcting ACT correlations for unreliability (DeShon, 1998, Eq. 5). The internal consistency reliabilities of the ACT subtests were obtained for the ACT normative sample (r_t = 0.91, 0.87, 0.92, for ACTm, ACTr, ACTe, respectively; ACT, 2017b, pp. 10.1–10.2). Using these reliabilities, the sample ACT correlations were corrected for unreliability and analyzed using the SEM model (Fig. 1) after constraining to equality the g loadings of the two verbal tests (ACTr and ACTe). (The constraint prevented the g loading of ACTe from exceeding 1.0, which produced a Heywood case.) The results were very similar to the prior analyses (cf. Table 4). The *g*-ToM effect was still strong ($r = 0.63$). In addition, the relations of the ACT non-*g* residuals with the ToM variables (ToM factor, RMET, SST) were still non-significant and low for the ACTm (−0.05, −0.03, −0.01), ACTr (0.00, 0.01, −0.03), and ACTe (0.15, 0.04, 0.09). The average difference in analogous effects between the supplemental and prior analyses (cf. Table 4) was trivial (*M* difference = 0.02).

[5] Additional analyses examined the effects of the ACT residuals after reversing the direction of effect between g and the ToM factor in the SEM model (Fig. 1). The new model was identical to the prior model except that the effect between g and the ToM factor was changed from g→ToM to g←ToM. As before, the analyses were performed separately for each ACT subtest (ACTm, ACTr, ACTe) and ToM variable (ToM factor, RMET, SST), yielding a total of nine effects (3 ToM variables x 3 ACT subtests). Consistent with the prior analyses (cf. Table 4, Analyses 2 to 10), the ACT residuals (ACTm, ACTr, ACTe) correlated negligibly with the ToM factor (−0.05, 0.09, −0.10), RMET (−0.02, 0.03, −0.03), and SST (−0.01, 0.01, −0.01).

[6] Additional analyses examined whether the non-*g* residuals of the ToM tests (RMET and SST) predicted the ACT subtests (ACTm, ACTr, ACTe). The analyses were analogous to those of the ACT residuals (Table 4, Analyses 5 to 10), except that the ToM residuals were analyzed. The analyses were performed separately for each ToM test and each ACT subtest, yielding a total of six effects (2 ToM tests x 3 ACT subtests). Consistent with the analyses of ACT residuals (cf. Table 4, Analyses 5 to 10), the ToM residuals (RMET, SST) correlated negligibly with ACTm (−0.01, 0.00), ACTr (0.02, 0.01), and ACTe (−0.01, 0.00).

3.3. Supplemental analyses

Table 4 (Analysis 11–20) reports supplemental analyses of the full sample ($N = 374$) using the same methods and the base model (Fig. 1). (These analyses included 125 subjects without ACT scores.) These analyses estimated the relations between *g* and the ToM factor (Analysis 11), and the relations between the non-*g* residuals of the ACT subtests and the ToM variables (ToM factor, RMET, SST) (Analysis 12–20). The results replicated the primary analyses (cf. Table 4, Analysis 1 to 10). *g* correlated strongly with the ToM factor ($\beta = 0.64$). In addition, the ToM factor and ToM tests correlated negligibly with the non-*g* residuals of all ACT subtests ($M_\beta = -0.01$) and with the residuals of the ACT math ($M_\beta = -0.02$) and verbal ($M_\beta < -0.01$) subtests. The mean difference in analogous effects for primary and supplemental analyses was trivial (M difference $< |0.01|$).

3.4. Robustness checks

Additional analyses controlled for SST comprehension of non-mental states (SSTC). These analyses were identical to the prior analyses (Table 4, Analyses 1–20), with one exception: An observed variable for SSTC was added to the SEM model (Fig. 1), which also included a path from the comprehension measure to SST reasoning (SST). The path from SSTC to SST controlled for story comprehension. The results replicated the prior analyses (Table 4, Analyses 1–20). *g* correlated strongly with the ToM factor ($M_\beta = 0.63$), and the ACT non-*g* residuals correlated negligibly with the ToM factor and ToM tests ($|M_\beta| < 0.04$). The mean difference in analogous effects with and without the SSTC in the model was trivial ($|M|$ difference < 0.002).

Other analyses examined whether the relations between *g* and the ToM factor (*g*→ToM) were affected by the set of ACT subtests. The set included one math subtest and two verbal subtests (reading and English), one of which (English) had a very strong *g* loading (loading = 0.97, Table 3). To examine whether the verbally-loaded ACT subtests affected *g*→ToM relations, additional analyses estimated *g* using the ACT math subtest plus either the reading or English subtest. These analyses were identical to the primary analyses, except that only two ACT subtests loaded on *g*: the math subtest plus one verbal subtest (reading or English). (Like the primary analyses, these analyses were performed using only subjects with ACT scores.) Consistent with the primary analyses (cf. Table 3), *g*→ToM relations were strong when *g* was based on the math and reading subtests ($\beta = 0.65$) or when *g* was based on the math and English subtests ($\beta = 0.63$).[7] The average difference in analogous effects for the current analyses (using two subtests) versus the original analyses (using three subtests) was trivial (M difference $= -0.01$) (cf. Table 3).

4. Discussion

This study was the first to examine *g*-ToM relations using a ToM factor based on multiple ToM tests (SST, RMET) and a *g* based on the ACT, a widely-used college admissions test. *g* correlated strongly with the ToM factor ($\beta = 0.65$) and moderately with the ToM tests ($M_\beta = 0.34$), which correlated modestly with each other ($\beta = 0.27$) (Table 3). The modest correlation between ToM tests indicates that *g* predicted limited variance among the ToM tests and suggests that the ToM tests had little in common. Moreover, the non-*g* residuals of the ACT math and verbal subtests, obtained after removing *g*, correlated negligibly with the ToM factor and ToM tests ($M_\beta = -0.01$), with trivial differences in effects for the math and verbal residuals (M difference $= -0.02$) (Table 4). Similar results were obtained for the ToM residuals, which correlated negligibly (and non-significantly) with the ACT subtests ($|M_\beta|$ effects < 0.01) (Footnote 6). The results support the primacy of *g* hypothesis, which predicts that *g* largely explains the relations among cognitive abilities (cf. Jensen, 1984; see also, Reeve & Charles, 2008).[8]

The trivial non-*g* effects replicated with (a) all three ACT scores (math, reading, English), (b) ACT math scores and one verbal score (reading or English), and (c) SAT scores. In addition, the trivial effects replicated when the path between *g* and the ToM factor was reversed (Footnote 5) and when the ToM residuals predicted the ACT subtests (Footnote 6). Together, the results suggest that the trivial non-*g* effects were robust to variation in the statistical model and residual source (ACT residuals vs. ToM residuals).

It should be emphasized that the strong *g*-ToM relation ($\beta = 0.65$, Table 3) was based on a modest correlation between the two ToM tests ($\beta = 0.27$, Table 3), indicating that *g* correlated with a small amount of shared variance between the tests (7%). The modest correlation between the ToM tests indicates that the tests were largely independent and argues against a strong general ToM factor based on variance common to the tests. In addition, the ToM factor was based on two ToM tests (RMET, SST), which may not be representative of other ToM tests, potentially limiting the generalizability of the findings. Given the modest correlation between the ToM tests, the interpretation of the strong *g*-ToM relation must acknowledge the limited variance among the ToM tests and the possibility that different ToM tests may yield different results. The issue of the weak relation between the ToM tests will be revisited below in the section on *Limitations*.

The results extend research on *g*-ToM relations. First, whereas prior research estimated *g*-ToM relations using IQ tests (e.g., Baker et al., 2014; Dodell-Feder et al., 2013), the current study estimated *g*-ToM relations using the ACT, a college admissions test widely used in the United States. The ACT correlates strongly with *g* (e.g., corrected $r = 0.77$, Koenig et al., 2008) and with the non-verbal Raven's Matrices (corrected $r = 0.75$, Koenig et al., 2008). In the current study, a *g* based on the ACT correlated strongly with the ToM factor ($\beta = 0.65$, Table 3). The strong *g*-ToM relation is consistent with the hypothesis that *g* and its proxies (the ACT) facilitate the ability to make complex inferences in everyday life (Gottfredson, 1997). Such inferences include the ability to infer the mental states of others, which is a key component of ToM (Baron-Cohen, 1995).

Second, whereas prior research estimated *g*-ToM relations using the RMET (Baker et al., 2014) or SST (Dodell-Feder et al., 2013), the current study estimated *g*-ToM relations using a latent ToM factor based on the RMET and SST. The ToM factor measured variance common to both the RMET and SST (after removing variance unique to the tests). The correlation between the ToM factor and *g* was strong ($\beta = 0.65$, Table 3) and was much stronger than a meta-analytic correlation between IQ and the RMET ($r_{\text{IQ-RMET}} \approx 0.24$, Baker et al., 2014). The strong correlation in the current study can be attributed to the use of multiple ToM tests and the removal of unique variance from the individual ToM tests. The use of multiple tests provides a more valid estimate of *g*-ToM relations compared to individual ToM tests (RMET or SST), which are loaded with more unique variance and less common

[7] Supplemental analyses using similar methods examined the independent group of subjects with SAT scores ($N = 96$), who received the same ToM tests. (The results should be interpreted with caution, given the relatively small number of subjects with SAT scores compared to number with ACT scores [$N = 249$].) These subjects had two SAT scores (math and verbal). The two scores were used to estimate relations between *g* and the ToM factor but not relations with the SAT non-*g* residuals, which would require three or more SAT scores to identify the SEM model and estimate effects of the non-*g* residuals. Consistent with analyses involving two ACT scores (i.e., ACT math plus reading or English), the *g* based on the two SAT scores correlated strongly with the ToM factor ($\beta = 0.52$), a large effect according to Cohen's (1998) criteria.

[8] It is worth noting that the average ratio of unique variance (based on the ACT non-*g* residual effects with ToM variables) over common variance (based on the *g*-ToM effect) was very small (0.04), further highlighting the strong effects of *g* and the weak effects of the non-*g* residuals (Table 4, Primary analyses).

variance (cf. Baker et al., 2014; see also, Jensen, 1998, pp. 30–31).

It is worth noting that the RMET and SST measure different aspects of ToM. The RMET measures social-perceptual ToM (i.e., inferring emotions from eyes), which is assumed to be fast and automatic (Baron-Cohen et al., 1997, 2001). In contrast, the SST measures social-cognitive ToM (e.g., inferring mental states of story characters), which is assumed to be slow and effortful (Dodell-Feder et al., 2013). Because of its cognitive orientation, the SST was expected to correlate more strongly than the RMET with *g*, a measure of cognitive ability. This hypothesis was not supported. The relations of *g* with each test (SST or RMET) were similar and not appreciably different (difference = 0.03, Table 3). The trivial difference in the relations of *g* with the SST and RMET suggests that, whatever their conceptual differences, both tests have similar loadings on *g*.

Third, whereas prior research estimated *g*-ToM relations (e.g., Baker et al., 2014; Dodell-Feder et al., 2013), the current study also estimated ToM relations with the non-*g* residuals of the ACT math and verbal subtests, obtained after removing *g* (cf. Coyle et al., 2013, 2015). Because ToM has been found to be especially sensitive to verbal abilities (e.g., Peterson & Miller, 2012), the verbal residuals were expected to predict ToM better than the math residuals. This hypothesis was not supported. The ACT verbal and math residuals showed negligible correlations with all ToM variables ($|\beta| < 0.06$) (Table 4), suggesting that specific cognitive abilities, obtained after removing *g*, contribute little to *g*-ToM relations. The negligible non-*g* effects, coupled with the strong *g*-ToM relations, indicates that ACT is linked to ToM via *g* (i.e., variance common to all tests). Stated differently, the link between ACT and the ToM is attributable to "not much more than *g*" (cf. Stauffer, Ree, & Carretta, 1996).

4.1. Limitations and future research

The current study has limitations that can guide future research. First, as noted above, the strong *g*-ToM relation ($\beta = 0.65$, Table 3) is based on a modest correlation between the two ToM tests ($\beta = 0.27$, Table 3), which indicates that *g* correlates with limited variance among the ToM tests (7%).[9] The relatively small amount of variance shared by the two ToM tests suggests that the tests are empirically distinct and measure different aspects of ToM (e.g., cognitive and perceptual). Moreover, the modest relation among the ToM tests may reflect a weak relation among ToM tests in general. ToM tests measure diverse ToM components (e.g., cognitive or affective; verbal or visual) (Henry et al., 2013, pp. 829–830), which may share limited variance and be empirically distinct. To test this possibility, future research should analyze a broader set of ToM tests, which would yield a better estimate of ToM variance and increase the validity of the *g*-ToM effects.

A second limitation concerns the use of the ACT as a measure of *g* and specific abilities. The ACT correlates strongly with *g* (e.g., $r = 0.77$, Koenig et al., 2008) and with specific math and verbal abilities, which predict school and work criteria (e.g., Coyle et al., 2014). However, the ACT excludes non-academic and other abilities (e.g., spatial abilities), which may also predict ToM. Future research should examine models of intelligence that sample a broader range of abilities. Two such models are the Cattell-Horn-Carroll (CHC) model (McGrew, 2009) and the verbal-perceptual-rotational (VPR) model (Johnson & Bouchard, 2007). The VPR and CHC models incorporate spatial abilities, which may predict ToM (cf. Cook & Saucier, 2010), and emotional intelligence (e.g., MacCann, Joseph, Newman, & Roberts, 2014), which may also predict ToM (e.g., Ferguson & Austin, 2010).

A third limitation concerns the sample of college students, who were admitted to college after taking the ACT, a selection test used to screen out low ability applicants. Compared to non-admitted applicants, admitted applicants would be expected to have higher levels of cognitive ability and possibly ToM, which correlates with *g*. In the current study, higher levels of cognitive ability and ToM could restrict variance on *g* and ToM, which could attenuate *g*-ToM effects. Although the strong *g*-ToM effects replicated after correcting for range restriction (Footnote 3), future research should use samples with a wider range of ability and age, which predicts cognitive functions related to ToM and *g* (e.g., executive functions) (e.g., Bull et al., 2008). Such samples would increase variance on *g* and ToM and could be used to examine whether *g*-ToM effects are mediated by other factors (e.g., executive functions).

A final limitation, related to the use of higher ability subjects, concerns Spearman's Law of Diminishing Returns (SLODR) (e.g., Jensen, 1998, pp. 585–588). SLODR is based on Spearman's (1932) observation that the correlations among mental tests decrease at higher ability levels, a pattern confirmed by a recent meta-analysis (Blum & Holling, 2017). The decrease (in correlations) is related to decreases in the *g* loadings of tests and the predictive validity of tests (Jensen, 1998, pp. 274–294). Based on SLODR, the use of higher ability subjects would be expected to depress *g*-ToM relations, which should decline at higher ability levels. Future research should estimate *g*-ToM relations at different ability levels and examine whether *g*-ToM relations decline at higher ability levels, which is predicted by SLODR.

5. Conclusion

This research examined *g*-ToM relations using an SEM model based on multiple measures of *g* and ToM (ACT, SST, RMET). *g* was based on the subtests of the ACT, which is strongly *g* loaded, and ToM was based on the SST and RMET, which measure different aspects of ToM (viz., perceptual and cognitive). The results indicated that *g* correlated strongly with ToM ($\beta = 0.65$, Table 4), and that the non-*g* residuals of the ACT subtests correlated trivially with ToM ($|\beta| < 0.06$, Table 4). However, the strong *g*-ToM relation was based on a modest correlation between the ToM tests ($\beta = 0.27$, Table 3), indicating that *g* correlated with limited variance among the ToM tests (7%). Future research should examine the robustness of effects using different measures of *g* and ToM and also examine possible mediators of *g*-ToM relations (e.g., executive functions).

Author note

This research was supported by a grant from the National Science Foundation's Interdisciplinary Behavioral and Social Science Research Competition (IBSS-L 1620457). The authors thank three anonymous reviewers for their constructive comments.

References

ACT (2017a). *ACT profile report – national: Graduating class 2017.* Iowa City, Iowa: ACT.
ACT (2017b). *ACT technical manual.* Iowa City, Iowa: ACT.
Astington, J. W. (1998). Theory of mind, Humpty Dumpty, and the icebox. *Human Development, 41*, 30–39.
Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and "Reading the Mind in the Eyes". *Intelligence, 44*, 78–92.
Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Boston: MIT Press/Bradford Books.
Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry, 38*, 813–822.
Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*, 241–251.

[9] The weak correlation between the two ToM measures (RMET and SST) may be attributed to the low internal reliabilities of the SST ($\alpha = 0.54$, Dodell-Feder et al., 2013) and the RMET (e.g., $\alpha = 0.60$, Mar, Oatley, Hirsh, dela Paz, & Peterson, 2006; $\alpha = 0.48$, Meyer & Shean, 2006), which would attenuate the correlation between the tests. The low reliabilities of the RMET and SST could distort the *g*-ToM effects, which were strong but based on two ToM measures with limited shared variance.

Blum, D., & Holling, H. (2017). Spearman's law of diminishing returns. A meta-analysis. *Intelligence, 65*, 60–66.

Botting, N., & Conti-Ramsden, G. (2008). The role of language, social cognition, and social skill in the functional social outcomes of young adolescents with and without a history of SLI. *British Journal of Developmental Psychology, 26*, 281–300.

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of theory of mind and executive function. *Cognition, 107*, 663–672.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, C. M., & Saucier, D. M. (2010). Mental rotation, targeting ability and Baron-Cohen's Empathizing-Systemizing Theory of Sex Differences. *Personality and Individual Differences, 49*, 712–716.

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA after removing *g*. *Intelligence, 36*, 719–729.

Coyle, T. R., Purcell, J. M., Snyder, A. C., & Kochunov, P. (2013). Non-*g* residuals of the SAT and ACT predict specific abilities. *Intelligence, 41*, 114–120.

Coyle, T. R., Purcell, J. M., Snyder, A. C., & Richmond, M. C. (2014). Ability tilt on the SAT and ACT predicts specific abilities and college majors. *Intelligence, 46*, 18–24.

Coyle, T. R., Snyder, A. C., Richmond, M. C., & Little, M. (2015). SAT non-*g* residuals predict course specific GPAs: Support for investment theory. *Intelligence, 51*, 57–66.

DeShon, R. P. (1998). A cautionary note on measurement error corrections in structural equation models. *Psychological Methods, 3*, 412–423.

Dodell-Feder, D., Lincoln, S. H., Coulson, J. P., & Hooker, C. I. (2013). Using fiction to assess mental state understanding: A new task for assessing theory of mind in adults. *PLoS One, 8*, 1–14.

Estes, D., & Bartsch, K. (2017). Theory of mind: A foundational component of human general intelligence. *Behavioral and Brain Sciences, 40*, 28–29.

Ferguson, F. J., & Austin, E. J. (2010). Associations of trait and ability emotional intelligence with performance on theory of mind tasks in an adult sample. *Personality and Individual Differences, 49*, 414–418.

Fernández-Berrocal, P., Cabello, R., & Gutiérrez-Cobo, M. (2017). Understanding the relationship between general intelligence and socio-cognitive abilities in humans. *Behavioral and Brain Sciences, 40*, 29–30.

Gottfredson, L. S. (1997). Why *g* matters : The complexity of everyday life. *Intelligence, 24*, 79–132.

Henry, J. D., Phillips, L. H., Beatty, W. W., McDonald, S., Longley, W. A., Joscelyne, A., & Rendell, P. G. (2009). Evidence for deficits in facial affect recognition and theory of mind in multiple sclerosis. *Journal of the International Neuropsychological Society, 15*, 277–285.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28*, 826–839.

Jensen, A. R. (1984). Test validity: *g* versus the specificity doctrine. *Journal of Social and Biological Structures, 7*, 93–118.

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

Johnson, W., & Bouchard, T. J. (2007). Sex differences in mental abilities: *g* masks the dimensions on which they lie. *Intelligence, 35*, 23–39.

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.

Johnson, W., te Nijenhuis, J., & Bouchard, T. J. (2008). Still just 1 *g*: Consistent results from five test batteries. *Intelligence, 36*, 81–95.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guildford.

Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence, 36*, 153–160.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh, 62*(Section A, Pt. 1), 28–30.

MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second-stratum factor of intelligence: Evidence from hierarchical and bifactor models. *Emotion, 14*, 358–374.

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality, 40*, 694–712.

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.

Meyer, J., & Shean, G. (2006). Social-cognitive functioning and schizotypal characteristics. *Journal of Psychology, 140*(3), 199–207.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*, 622–646.

Nettle, D., & Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *European Journal of Personality, 22*, 232–335.

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: Toward a brief form for research and applied settings. *Frontiers in Psychology, 6*, 1–14.

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*, 1–6.

Ree, M. J., Carretta, T. R., Earles, J. A., & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formula. *Journal of Applied Psychology, 79*, 298–301.

Reeve, C. L., & Charles, J. E. (2008). Survey of opinions on the primacy of *g* and social consequences of ability testing: A comparison of expert and non-expert views. *Intelligence, 36*, 681–688.

Royall, D. R., & Palmer, R. F. (2014). "Executive functions" cannot be distinguished from general intelligence: Two variations on a single theme within a symphony of latent variance. *Frontiers in Behavioral Neuroscience, 8*, 1–10.

Spearman, C. (1932). *The abilities of man: Their nature and measurement.* New York, NY: AMS Press.

Stauffer, J. M., Ree, M. J., & Carretta, T. R. (1996). Cognitive-components tests are not much more than *g*: An extension of Kyllonen's analyses. *Journal of General Psychology, 123*, 193–205.

Wiberg, M., & Sundström, A. (2000). A comparison of two approaches to correction of restriction of range in correlation analysis. *Practical Assessment, Research & Evaluation, 14*, 1–9.