Perspectives on Behavior Science (2020) 43:189–213
https://doi.org/10.1007/s40614-019-00197-z

The Relational Abilities Index+: Initial Validation of a Functionally Understood Proxy Measure for Intelligence

Dylan Colbert[1] · Aoife Malone[1] · Seafra Barrett[1] · Bryan Roche[1]

Published online: 24 May 2019

Abstract
The Relational Abilities Index (RAI) has shown considerable utility as a functional proxy measurement of intellectual performance by providing a metric of an important skill set known as relational skills, which are proposed to underlie much of what we conceive of as intellectual behavior. The Relational Abilities Index+ (RAI+) assesses performance across an extended range of relational skills (Same/Opposite, More/Less, Same/Different, Before/After, and Analogy), and has been designed to provide a more comprehensive and nuanced assessment of relational skills. The current study aims to investigate the validity and utility of the RAI+ by assessing its degree of correlation with well-established assessments of intelligence (WASI), numeracy (WAIS: Arithmetic), and educational attainment (WIAT-T-II). Results indicate that the RAI+ displays considerable efficacy in predicting intellectual performance and numeracy, but not educational attainment.

Keywords Relational frame theory · Relational abilities index · Relational skills · Intelligence · Academic attainment · Numeracy

The construct of intelligence has been an extensively debated topic among academics in the modern era, and consensus on a single definition remains elusive (Jensen, 1998; Mackintosh, 1998; Neisser et al., 1996). Perhaps the most prevalent conceptualization of the construct is rooted in Spearman's general factor theory, which posits that there is a latent factor or faculty ("*g*") that influences performance on all intelligence measures and that stays stable throughout a person's lifetime (Spearman, 1904). It is critical to note, however, that due to the nature of *g* the ability to measure this variable has yet to

This research has conducted by the first three authors under the supervision of the last author.

Aoife Malone and Seafra Barrett contributed equally to this work.

✉ Dylan Colbert
Dylan.colbert.2011@mumail.ie

[1] Department of Psychology, Maynooth University, Maynooth, Ireland

be clearly established, with much of its supposed measurement thus far being indirect, through the use of Intelligence Quotient (IQ) tests (Richardson, 2002; Richardson & Norgate, 2015). To this extent, intelligence has typically been operationalized in terms of what IQ tests measure (Boring, 1923; Richardson & Norgate, 2015; van der Maas, Kan, & Borsboom, 2014), insofar as IQ tests are said to merely "define the theory of intelligence that the test is intended to measure" (Naglieri, 2008, p. 68). In essence, by assessing an individual's performance on a range of mental tasks, IQ tests attempt to reduce a wide-ranging spectrum of intellectual behaviors into a unitary, quantitative factor (Cassidy, Roche, & O'Hora, 2010).

IQ test scores are argued to reflect a stable, invariant, and nonmalleable trait (Jensen, 1980; Juliano, Haddad, & Carroll, 1988; Locurto, 1991; Ramsden et al., 2011; Reynolds, Gutlin, Dappen, & Wright, 1979; Spearman, 1927). However, empirical research has increasingly suggested that IQ test scores may not be as immutable as once assumed, with the Flynn effect identifying substantial rises in IQ test performance throughout the 20th and 21st centuries (Flynn, 1984, 1998, 2007, 2009). It is interesting that there is emerging evidence to propose that the Flynn effect may have stalled or even reversed in recent times (Dutton & Lynn, 2015; Pietschnig & Voracek, 2015; Shayer & Ginsburg, 2009; Shayer, Ginsburg, & Coe, 2007; Sundet, Barlaug, & Torjussen, 2004; Teasdale & Owen, 2008; Woodley & Meisenberg, 2013). Nevertheless, the very fact that population IQ scores undergo such nonlinear fluctuations across time has been attributed to a wide variety of environmental (Dickens & Flynn, 2001; Flynn, 2007; Lynn, 1990, 2009; Wai & Putallaz, 2011), social (Blair, Gamson, Thorne, & Baker, 2005; Brand, 1987; Ceci, 1991), and genetic factors (Jensen, 1998; Mingroni, 2007), and suggests that the stability espoused by trait-based theories of intelligence may have been exaggerated. Indeed, there is now accumulating evidence within the literature (e.g., Olesen, Westerberg, & Klingberg, 2004; Schlinger, 1993; Sternberg, 2008) that argues that intellectual ability is in fact, a pliable concept that is influenced by the environment (Dickens & Flynn, 2001; Nisbett et al., 2012). Indeed, evidence from cognitive (Basak, Boot, Voss, & Kramer, 2008; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi, Buschkuehl, Jonides, & Shah, 2011; Jaeggi, Buschkuehl, Perrig, & Meier, 2010; Mackey, Hill, Stone, & Bunge, 2011; Stephenson & Halpern, 2013), educational (Brinch & Galloway, 2012; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Ceci, 1991; Jencks, 1972), neuroscientific (Mackey, Miller-Singley, & Bunge, 2013; Mackey, Whittaker, & Bunge, 2012), and behavior-analytic studies (Cassidy, Roche, Colbert, Stewart, & Grey, 2016; Cassidy, Roche, & Hayes, 2011; Colbert, Tyndall, Roche, & Cassidy, 2018; Hayes & Stewart, 2016) support the idea that intelligence can be improved through environmental interventions.

One noteworthy research stream that has proved particularly efficacious in improving intellectual performance has emerged from a set of behavior-analytic principles that allow researchers to target a key repertoire of skills (known as relational skills) that are proposed to underlie most, or even all, intellectual performance (e.g., Cassidy et al., 2011, 2016; Colbert, Dobutowitsch, Roche, & Brophy, 2017). From this behavior-analytic or *functional* perspective (De Houwer, Barnes-Holmes, & Moors 2013), trait definitions of intelligence are considered to fall victim to the errors of reification and circular reasoning (Gottfredson, 1998; Howe, 1990; Schlinger, 2003), and therefore are wholly incongruent with the behaviorist tradition (Schlinger, 2003; Skinner, 1974). Thus, behavior analysts embrace a more functional account of "intelligence," in which the term merely refers to a measurable quality of a set of actions, which are intricately linked to their context and are amenable to experimental manipulation (Cassidy et al., 2011; Hayes & Stewart, 2016; Schlinger, 2003). In effect, intellectual abilities are viewed as malleable, with IQ tests functioning solely to provide an index of the fluency of the skills involved. The stability of IQ scores across time does not imply the existence of an underlying trait, but merely reflects stability in the learning environment and the unfolding of intellectual development at a typical rate. Among the key tasks of the psychologist in this domain, viewed from a behaviorist perspective, is to identify extraordinary environmental contexts that can accelerate intellectual development.

Recent developments within the field of behavior analysis, most notably under the rubric of Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001), have begun to progressively explore the utility of this functional approach to conceiving intellectual behavior (Dymond & Roche, 2013; O'Hora, Pelaez, & Barnes-Holmes, 2005; O'Toole, Barnes-Holmes, Murphy, O'Connor, & Barnes-Holmes, 2009). In particular, RFT highlights the *derived* and *generative* nature of human language and cognition (O'Toole et al., 2009) and proposes that much of what we consider intelligent behavior can be reconsidered as a form of behavior known as *derived relational responding* (DRR; Barnes-Holmes et al., 2001; Cassidy et al., 2010). DRR can be defined as the act of responding to one stimulus in accordance with its derived relation to another stimulus (Stewart & McElwee, 2009). Such responding is therefore under the contextual control of specific relational "frames," such as coordination (e.g., A is the same as B; Hayes et al., 2001), distinction (e.g., A is different to B; Roche & Barnes, 1997), opposition (e.g., A is opposite to B; Barnes-Holmes, Barnes-Holmes, & Smeets, 2004a), comparison (e.g., A is greater than/less than B; Dymond & Barnes, 1995; O'Hora, Roche, Barnes-Holmes, & Smeets, 2002), temporality (e.g., A is before/after B; O'Hora et al., 2002), analogy (e.g., A is to B as C is to D; Stewart, Barnes-Holmes, & Roche, 2004), hierarchy (e.g., A subsumes/belongs to B; Griffee & Dougher, 2002), and deixis (e.g., A is here and B is there; McHugh, Barnes-Holmes, & Barnes-Holmes, 2004). Following the acquisition of a large number of these relational frames involving a large number of stimuli, an individual can respond to an almost infinite number of established or novel (i.e., derived) relations between stimuli. For example, if an individual is explicitly taught that relatum A is greater than relatum B, and relatum B is the same as relatum C, it is possible to derive the relation between relata A and C (i.e., A is more than C) based on their respective relation to relatum B, even though this relation has never been trained directly.

Several studies have revealed high levels of correlations between measures of relational responding and various tests of verbal ability (Barnes, McCullagh, & Keenan, 1990; Colbert et al., 2017; Dugdale & Lowe, 2000). In the domain of academic attainment, relational responding has been shown to be of key importance to reading (de Rose, de Souza, Rossito, & de Rose, 1992; Farrington-Flint & Wood, 2007; Goswami, 1986; Mackay, 1985; Sidman, 1971), vocabulary (Edwards, Figueras, Mellanby, & Langdon, 2011; McHugh et al., 2004; Nippold & Sullivan, 1987), grammar (Hock, 1991, 2003), and even spelling (Brown, Sinatra, & Wagstaff, 1996; Goswami, 1988; Mackay, 1985). These findings are further complemented by a range of studies that indicate that the ability to derive relations can be enhanced (e.g., Barnes-Holmes et al., 2004; Berens & Hayes, 2007; Carpentier, Smeets, & Barnes-Holmes, 2003; Rosales, Rehfeldt, & Lovett, 2011), and that doing so can result in increases in both verbal and intellectual ability (e.g., Cassidy et al., 2011, 2016; Colbert et al., 2018; Dixon, Whiting, Rowsey, & Belisly, 2014; Moran, Stewart, McElwee, & Ming, 2010). Such findings would appear to highlight the importance of an established proficiency in relational responding as a key contributor to literacy, among other intellectual skills. It is critical to note that RFT has produced dozens of studies that have shown that DRR interventions usher in language ability (e.g., Cowley, Green, & Braunling-McMorrow, 1992; Cullinan, Barnes, Hampson, & Lyddy 1994; de Rose et al., 1992; Hayes & Hayes, 1992; Matos & d'Oliveira, 1992; Murphy & Barnes-Holmes, 2010b, 2010a, 2011; Murphy, Barnes-Holmes, & Barnes-Holmes, 2003; Rehfeldt & Root, 2005; Wulfert & Hayes, 1988), and so it is widely argued that the former produces the latter rather than vice versa (Barnes-Holmes, Finn, McEnteggart, & Barnes-Holmes, 2017).

Sophistication in relational responding may also comprise a key facet of numeracy and mathematical fluency (Berens & Hayes, 2007; Carpenter, Franke, & Levi, 2003; Colbert et al., 2016; O'Hora et al., 2005). Molina, Castro, and Ambrose (2005) found that the encouragement of "relational thinking" (i.e., analyzing the relationships specified in mathematical problems before engaging in mathematical computation) afforded a meaningful and comprehensive learning of arithmetic and provided a foundational basis for the study of algebra in a sample of primary-school–aged children. Indeed, the mathematical symbols that receive such focus in these relational thinking interventions are conceived of as contextual cues from an RFT perspective. In addition, several RFT studies conducted by Ninness et al. (2005a, 2005b, 2006, 2009) have demonstrated the utility of training students to derive relations as a means of improving advanced mathematical skills.

Relational Ability and Measures of Intellectual and Academic Ability

Several authors have now argued that standardized IQ tests can be conceived as tests of DRR proficiency (e.g., Cassidy et al., 2010; O'Hora et al., 2005). Perhaps the most evident example of the IQ test items that "tap into" DRR proficiency comes in the form of assessments of vocabulary, in which participants are required to define a given word. Such subtests can be considered tests of either word–word and word–object coordination relations, and are included in a number of gold-standard IQ assessments such as the Wechsler Adult Scale of Intelligence (WAIS; Wechsler, 1955, 1981, 1997, 2008) and Stanford-Binet Intelligence Scale (SB; Roid, 2003; Terman, 1916; Terman & Merrill, 1937, 1960; Thorndike, Hagen, & Sattler, 1986). Subtests such as Wechsler Similarities, Woodcock-Johnson Verbal Comprehension (Schrank, McGrew, Mather, & Woodcock, 2014) and Differential Ability Scales Similarities (Bennett, Seashore, & Wesman, 1990; Elliot, 2007) require participants to identify the common characteristics and/or categorizations of two words (e.g., "How are a bus and a plane alike?"), thereby implicating coordination and/or hierarchical relational responding. In addition, comparison-based relational responding may be viewed as a contributor to performance on subtests assessing numeracy, such as the arithmetic subtest commonly included in various Wechsler IQ scales (e.g., "Chris has two times as much as Robert. Chris has 99 pounds. How much money does Robert have?"). The importance of analogical relational responding is evinced by the inclusion of both verbal and numerical analogical reasoning tasks in a wide variety of testing batteries, including the Stanford-Binet, Woodcock-Johnson, and Cognitive Abilities Test. (For further detail outlining the relevance of relational responding skill to specific IQ test items, see Cassidy et al., 2010).

A small number of correlational analyses (Colbert et al., 2017; Dixon et al., 2014; O'Hora et al., 2005, 2008) have further underlined the relationship between relational responding and intellectual performance by assessing performance on traditional IQ test performance and assessments of relational skills. For instance, across two separate analyses, O'Hora et al. (2005) and O'Hora et al. (2008) reported significant correlations between performance on a temporal relations task (before/after) and all three WAIS-III indices (Full Scale, Verbal, and Performance IQ), two of four WAIS-III subindices (Verbal Comprehension and Perceptual Organization) and two WAIS-III subtests (Vocabulary and Arithmetic). In addition, a recent study by Dixon, Belisle, and Stanley (2018) found that an assessment of DRR across numerous sensory modalities, the PEAK-E-PA (Promoting the Emergence of Advanced Knowledge Equivalence Pre-Assessment, Dixon, 2015) displayed a high level of correlation with performance on two IQ subtests, Vocabulary and Block Design.

The Relational Abilities Index (RAI), initially developed by Cassidy (2008) and later extended by Cassidy et al. (2016), is in its current form a 55-item syllogistic reasoning assessment that measures proficiency in coordination, opposition, and comparison relational responding, and is now regarded as an acceptable proxy measure of IQ (Colbert et al., 2017). Colbert et al. (2017) carried out the most in-depth analysis of the RAI to date, reporting medium to strong correlations between RAI scores and all three WAIS-III indices (Full Scale, Verbal and Performance IQ), all four subscales (Verbal Comprehension, Working Memory, Perceptual Organization and Processing Speed), as well as 10 of 13 IQ subtests. In addition, in the second of the Colbert et al.'s studies, RAI scores predicted performance on number of other measures of cognitive ability, including verbal ability (National Adult Reading Test; Nelson, 1982), visuo-spatial function (the Trail Making Test; Lezak, 1995), and memory (Rey Auditory Visual Learning Tests; Rey, 1958; English version: Taylor, 1959). Although the RAI demonstrated considerable predictive validity across the test battery, closer investigation indicated the RAI's relatively limited utility in discriminating performance for high IQ participants due to a potential ceiling effect. As such, the authors concluded that the inclusion of a wider range of relational tasks, such as temporality and analogy, may be beneficial in parsing out individual differences across a greater diversity of trial types and providing a more comprehensive account of relational ability and how its various aspects relate differentially to various aspects of intelligence, as typically assessed.

Indeed, one of the primary functions, and great utilities, of intelligence testing is the prediction of educational achievement (Kaufman, Reynolds, Liu, Kaufman, & McGrew, 2012), and correlations between the two constructs have been well-established (Deary, Strand, Smith, & Fernandes, 2007; Jensen, 1998; Lynn & Meisenberg, 2010; Mackintosh, 1998; Naglieri & Bornstein, 2003; Rindermann, 2007). Despite this, the direct relationship between relational responding and traditional measures of academic attainment remains relatively unelucidated. However, in one study, Cassidy et al. (2016) reported a significant correlation between RAI scores and overall Educational Aptitude, Verbal Ability, and Numerical Ability as assessed by a widely used measure of scholastic ability, the Differential Aptitude Test (DAT; Bennett et al., 1990).

The Current Study

Given the theoretical objections toward IQ espoused by many behaviorists, a relational abilities assessment harbors considerable promise of providing a functional alternative (i.e., one that is understood technically and does not necessitate hypothetical constructs) to traditional IQ testing batteries as a measure of skills underlying intellectual performance. Despite displaying considerable utility as a proxy measure of IQ, the original Relational Abilities Index is somewhat limited in scope, due to the relatively narrow compendium of relational frames included. The four relational frames included in the RAI (same, opposite, more, and less) were originally selected due to their apparent importance in standardized IQ tests (Cassidy et al., 2010; Stewart, Tarbox, Roche, & O'Hora, 2013), along with their prominence in language acquisition and logistical reasoning (Barnes-Holmes, Barnes-Holmes, Smeets, Cullinan, & Leader, 2004b; Hayes et al., 2001; Rehfeldt & Barnes-Holmes, 2009). However, a relational abilities index may benefit from comprehensive expansion in order to assess a wider range of relational skills, which have previously been associated with intellectual behavior, namely distinction, temporality, and analogy (Hayes & Stewart, 2016; O'Hora et al., 2005, 2008). Such a development should improve the utility of a relational abilities index in providing a more sensitive and nuanced differentiation of performances, in particular at the higher end of the performance spectrum (Colbert et al., 2017; Gore, Barnes-Holmes, & Murphy, 2010). As such, the Relational Abilities Index+ (RAI+) has been developed for this purpose, and assesses performance across five modules of relational responding (Same/Opposite, More/Less, Same/Different, Before/After, and Analogy) in comparison to the two modules assessed by the original RAI (Same/Opposite and More/Less).

The primary aim of the current study is to investigate the potential utility of a more expansive assessment of relational responding as a functional measure of intellectual performance by analyzing the relationship between scores on the RAI+ and scores on widely administered, gold-standard assessments of intelligence (Wechsler Abbreviated Scale of Intelligence; WASI), numeracy (Wechsler Adult Intelligence Scale: Arithmetic subtest), literacy and educational attainment (Wechsler Individual Achievement Test: Teacher Edition). This objective does not entail a motivation to validate the RAI+ as a "better" measure of IQ, but rather attempts to evaluate the efficacy of a functional assessment of intellectual performance by studying its relevance to performance on tasks generally perceived to reflect intellectual ability by those outside the field of behavior analysis. In addition, this design will permit an elucidation of the relative contribution of each of the relational frames being assessed to intellectual performance.

Method

Participants

A total of 97 individuals (50 female) participated in this study. Participants' ages ranged from 18 to 45 years (mean = 25.42, SD = 7.87). All participants were fluent English speakers with no incidence of any cognitive disorders or impairments that could have affected the current results. The vast majority of participants (n = 85) were attending third-level education across a range of disciplines at the time of participation.

Materials

Relational Abilities Index (RAI+) A revised version (RAI+) of the RAI employed in Colbert et al. (2017) and Cassidy et al. (2016) was administered through the website proprofs.org to assess participants' relational abilities. The RAI+ consists of a battery of 67 syllogistic relational puzzles, assessing proficiency in responding in accordance with Same/Opposite (15 trials), Same/Different (14 trials), More/Less (13 trials), Before/After (13 trials), and Analogy (12 trials) frames in that order (see Appendix Table 3 for examples). The RAI+ required approximately 20–25 minutes to complete. Full details of each of the 67 RAI+ trials are included in Table 3 in the Appendix.

The general format of trials utilized in the RAI+ mirrored that of the original RAI. Each task consisted of between one and three relational premise(s) in which relations between nonsense words were stated (e.g., "CUG is the same as TOF"), followed by a question based on the relationship(s) specified in the premise(s) (e.g., "Is TOF the same as CUG?"). A total of 227 stimuli comprised of three-letter nonsense words (e.g., "CUG," "TOF," "JOS") in the format "consonant-vowel-consonant" (to ensure pronounceability) were presented with no stimulus being repeated throughout the assessment. Participants indicated their response by using the computer mouse to click on either a "YES" or "NO" button onscreen. Positional responding was controlled for by switching the positions of the response options throughout the assessment. A countdown timer was also visible on the page at all times, imposing a limit of 34 min to complete the assessment (30 sec per question; see Figure 1).

Task complexity was therefore controlled by modifying; 1) the number of relational premises (1–3); 2) the order of relational premises (sequential or random); 3) the directionality of the relational question (i.e., whether or not the relational question probes for first-term–last-term relations, or last-term–first-term relations as specified in the premises); 4) the number of relation types presented in each trial (e.g., only "same" relations, or a combination of "same" and "opposite"); and 5) the presence/absence of the relational cue used in the question in the relational premise(s), (e.g., "CUG is same as LER, is CUG same as LER?").

With the exception of Analogy trials, the first trial for each relational frame included a single premise, followed by a relational question and as such assessed the participant's ability to derive mutually entailed relations. This involved either changing the directionality of the relational statement or switching the relational frame to its inverse in the relational question. Each block then progressed to 10 two-premise trials that included three relata in which every possible derived relation within this network was probed for. Finally, each block then included a number of three-premise trials (4–6) that specified a relational network across four relata. Any additional relations entailed by the presentation of the fourth relational premise (e.g., between stimulus A/B/C and stimulus D, and vice versa) were assessed during these trials. For the Analogy block, 12 two-premise trials were included. Each premise stated the relation between two stimuli in accordance with same/opposite (four trials), before/after (four trials), and more/less (four trials),

LUZ is the same as TIV Is LUZ opposite to TIV?	SAJ is different to LIR LIR is the same as VUS Is VUS the same as SAJ?
Block 1: Same/Opposite	**Block 2: Same/Different**
NOG is less than HAV HAV is less than WUQ WUQ is less than VUN Is HAV more than VUN?	LOF is after FEH WUC is after LOF LON is after WUC Is FEH before LON?
Block 3: More/Less	**Block 4: Before/After**
QUD is less than KON JOL is more than JIT Is KON to QUD the same as JIT to JOL?	REG is same as CAS WIX is opposite to DOM Is CAS to REG different from DOM to WIX?
Block 5: Analogy	

Fig. 1 Sample relational tasks for each of the 5 RAI+ modules

followed by a "same/different" relational question that probed for relationship between each relational premises specified (e.g., "is FEG to TID the same as VER to RUF?").

WASI The Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) is a widely administered, short-form assessment that gives an approximation of an individual's intellectual performance relative to his/her peers. For the current sample, all four IQ subtests (Vocabulary, Similarities, Block Design, and Matrix

Reasoning) were administered, which permitted the computation of a Full Scale IQ score, as well as both the Verbal and Performance IQ subindices.

WIAT-II-T The Wechsler Individual Achievement Test Second UK Edition for Teachers (WIAT-II-T; Pearson Education, 2006a), an adaptation of the traditional WIAT-II battery, is an individually administered standardized assessment of educational attainment and is comprised of three subtests of Reading, Spelling, and Reading Comprehension. The WIAT-II has satisfactory construct, content and criterion validity as well as test-retest reliability for an adult population (Pearson Education, 2006b).

WAIS-III: Arithmetic The Wechsler Adult Intelligence Scale (WAIS-III: UK; Wechsler, 1998) is an individually administered assessment of intellectual ability. It is one of the most popular measures in neuropsychological batteries, and is often considered a "gold standard" of intelligence testing (Butler, Retzlaff, & Vanderploeg, 1991; Ivnik et al., 1992; Strauss, Sherman, & Spreen, 2006). The Arithmetic subtest of the WAIS-III comprises one part of the Working Memory subindex of Verbal IQ, and consists of 20 arithmetic questions that successively increase in difficulty, and are subject to a time limit. Normed tables for this subtest are available for the computation of a standardized score.

Ethics

The current study was conducted in adherence to guidelines specified by Maynooth University's Social Research Ethics Committee and the Psychological Society of Ireland. As a study originally designed as an undergraduate research project, it was not required to undergo explicit committee approval but conformed to a checklist of considerations.

General Procedure

The study was conducted in a private experimental room, free from noise and other distracting stimuli. Participants were seated at a desk directly opposite the researcher and were required to provide valid consent before participation. Each participant was engaged in the task individually, on a one-to-one basis with the researcher. Participants were briefed on the general nature of the study, and signed a consent form at this first stage. Although all participants completed the RAI+, a section of the sample completed this alongside the WASI (n = 60) and another subsample completed this alongside the WIAT-T (n = 37).

Results

Mean RAI+ scores for the current sample was 59.82 out of 67 (89.3%), with scores ranging from 41 to 67. Table 1 details full descriptive statistics for individual RAI+ scores, WASI subindices, WIAT-T-II scores, and WAIS-III Arithmetic scores.

Table 1 Descriptive statistics for RAI+, WASI, WIAT-T II and WASI-III Arithmetic scores

Measure	M	SD	Range
RAI+	89.3%	9.2	61.2 - 100%
Same/Opposite	88.9%	12.3	40-100%
Same/Different	88.2%	8.8	40-100%
More/Less	80%	11.8	46.2 – 100%
Before/After	76%	14	38.5 – 100%
Analogy	65.4%	15.1	16.7 – 100%
WASI			
Full Scale IQ	109.6	9.24	89-128
Verbal IQ	109.6	10.65	88-126
Performance IQ	107.2	9.37	88-126
WIAT	108.8	6.06	96-119
Reading Comprehension	102.9	7.9	78-115
Reading	113.2	4.47	99-119
Spelling	110.3	9.35	87-125
WAIS Arithmetic	126.3	17.5	85-145

In relation to the WASI, scores for Full Scale IQ (M = 109.62, SD = 9.24), Verbal IQ (M = 109.58, SD = 10.65) and Performance IQ (M = 107.2, SD = 9.37) were all towards the upper limit of the average range. For the WIAT-II-T, scores for WIAT Overall standardized (M = 108.8, SD = 6.06), Reading (M = 102.86, SD = 7.9), Reading Comprehension (M = 113.22, SD = 4.47) and Spelling (M = 110.3, SD = 9.35) were all in the average to above average range. For the WAIS arithmetic subtest, standardized score estimates (converted from scaled scores) ranged from 85 to 145 (M = 126.32, SD = 17.5).

Mean accuracy scores were highest for the Same/Opposite (M = 88.9%, SD = 12.3) and Same/Different (M = 88.2%, SD = 8.3) modules, followed by More/Less (M = 80%, SD = 10.2) and Before/After (M = 76%, SD = 12.2). Performance on the Analogy module was significantly lower, with mean accuracy at 65.4% (SD = 9.2). Figure 2 displays the distribution of scores for total RAI+ score, as well as the distribution of each testing module.

Correlational Analyses

Each individual relational skills module demonstrated strong, significant correlations with overall RAI+ score, suggesting respectable internal consistency. Of the five frames, Same/Opposite (rho = .79, p < .001), Before/After (rho = .78, p < .001), and More/Less (rho = .75, p <.001) tasks displayed the closest relationship, followed by Analogy (rho = .67, p < .001) and Same/Different (rho = .52, p < .001). The Cronbach's alpha statistic for the RAI+ was 0.79. Results from the correlational analysis of RAI+ performance and WASI, WIAT, and WAIS Arithmetic scores are shown in Table 2.

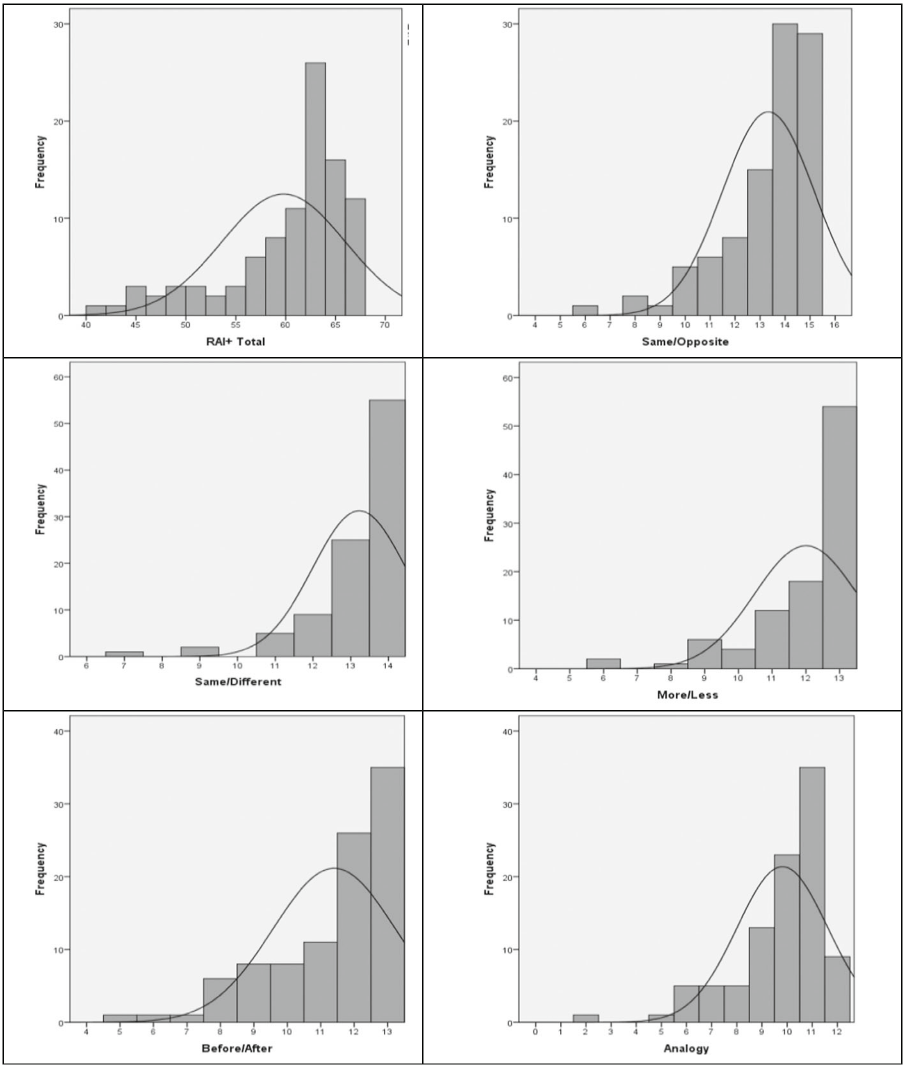

Fig. 2 Histograms displaying the distribution of scores for overall RAI+ and each of the five testing blocks. Plots were calculated using the total number of correct trials per block

WASI Overall RAI+ scores correlated significantly with Full Scale IQ (*rho* = .54, *p* <. 001), as well as both Verbal IQ (*rho* = .42, *p* <.001) and Performance IQ (*rho* = .48, *p* <.001). Performance on the RAI+ also correlated significantly with all four IQ subtests: Vocabulary (*rho* = .37, *p* =.003), Similarities (*rho* = .37, *p* =.003), Block Design (*rho* = .42, *p* =.001) and Matrix Reasoning (*rho* = .42, *p* = .001). Figure 3 represents scatterplots outlining the relationship between RAI+ scores and each of the WASI's three IQ indices.

Additional exploratory analyses revealed that out of the five relational task blocks that comprise the RAI+ More/Less tasks exhibited the strongest correlation with WASI Full Scale IQ (*rho* = .49, *p* < .001), closely followed by Same/Different (*rho* = .48, *p* <

Table 2 Correlations between RAI+ accuracy scores, and WASI IQ and its subindex scores

Measure	Correlation coefficient	Significance level
WASI		
Full Scale IQ	0.54**	<.001
Verbal IQ	0.42**	.001
Vocabulary	0.37**	.003
Similarities	0.37**	.003
Performance IQ	0.48**	<.001
Block Design	0.42**	.001
Matrix Reasoning	0.42**	.001
WIAT-T	0.27	.1
Reading	0.14	.416
Reading Comprehension	0.08	.673
Spelling	0.29	.29
WAIS-III Arithmetic	0.43**	.009

**Indicates correlation is significant at the 0.01 level (2-tailed)

.001), Same/Opposite (*rho* = .44, $p < .001$), and Before/After (*rho* = .42, $p = .001$) tasks. Each of these four relational skill sets also displayed a significant relationship with Verbal IQ (Before/After, *rho* = .43, $p = .001$; More/Less, *rho* = .42, $p = .001$; Same/Different, *rho* = .31, $p = .015$; Same/Opposite, *rho* = .26, $p = .045$). Likewise, scores for the relational skills Same/Different (*rho* = .48, $p < .001$), Same/Opposite (*rho* = .44, $p < .001$), More/Less (*rho* = .44, $p = .001$), and Before/After (*rho* = .33, $p = .01$) were correlated moderately with Performance IQ. It was surprising that Analogy tasks displayed a significant correlation with Performance IQ (*rho* = .3, $p = .02$), but not with Full Scale IQ or Verbal IQ. Removal of Analogy module scores from the aggregate RAI+ score, increased the overall RAI+ correlation with Full Scale IQ (*rho* = .55, $p < .001$) and Verbal IQ (*rho* = .46, $p < .001$), but the correlation with Performance IQ was unaffected (*rho* = .48, $p < .001$).

The relationship between RAI+ scores and Full Scale IQ for high IQ individuals (Full Scale IQ: 110+, *rho* = .44, $p = .03$) was statistically significant. In addition, significant correlations were found between RAI+ scores and both Verbal IQ (*rho* = .64, $p = .001$) and Performance IQ (*rho* = .47, $p = .02$), for this group.

WIAT-T RAI+ performance did not show a significant level of correlation with WIAT-T Standardized Score or any of the three WIAT-T subtests; Reading, Reading Comprehension, and Spelling. Further analyses revealed that performance on the Same/Opposite block showed a moderate significant relationship with WIAT-T Standardized Score (*rho* = .35, $p = .04$). However, of the four other relational modules administered, no significant correlations were found with WIAT-T Standardized Score, Reading, Reading Comprehension, or Spelling.

WAIS Arithmetic Overall scores for the RAI+ showed a moderate positive correlation with WAIS-III arithmetic scores (*rho* = 0.43, $p = .009$). WAIS-III arithmetic scores also

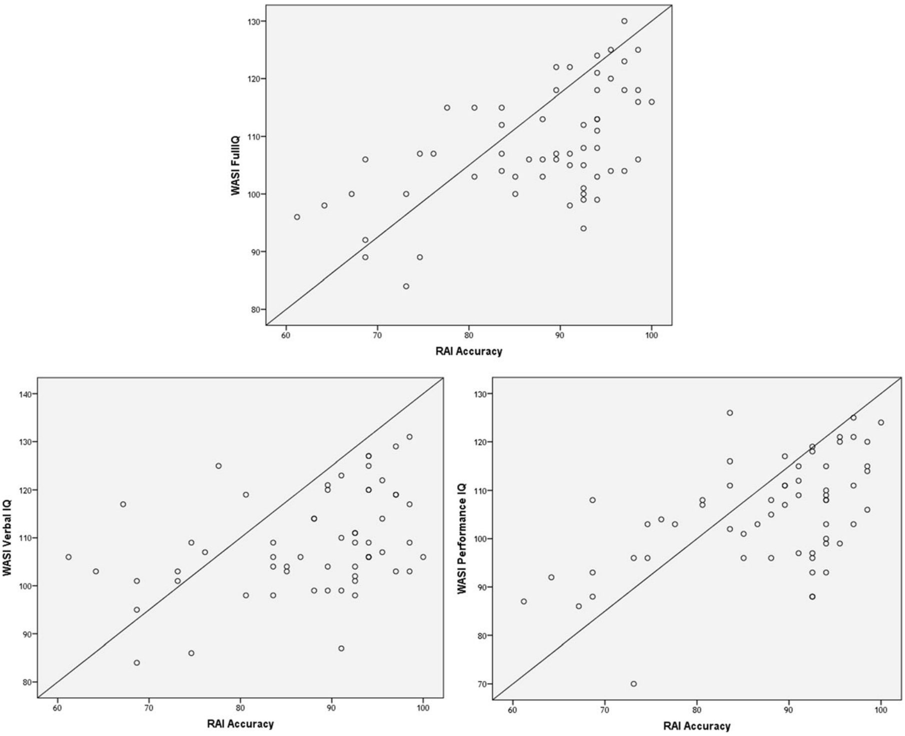

Fig. 3 Scatterplots of the relationship between RAI+ and each of the three WASI IQ indices

correlated significantly with Same/Opposite (*rho* = 0.6, p< .001) and Before/After scores (*rho* = 0.4, *p* = .013), but not with other RAI+ subtest scores.

In summary, the RAI+ aggregate score displayed significant levels of correlation with all seven WASI IQ indices and subtests, as well as WAIS-III Arithmetic. Upon investigation, it was found that scores for the Same/Opposite, Same/Different, More/Less, and Before/After test modules all correlated with the three WASI IQ indices, whereas the Analogy module only correlated with one IQ index (Performance IQ). RAI+ total and module scores generally did not correlate with any of the WIAT-T metrics, with the sole exception of the Same/Opposite module.

Discussion

The purpose of the current study was to investigate and evaluate the utility of the RAI+ as a potential proxy measure of intellectual and scholastic ability, through assessing its degree of correlation with measures of intellectual performance (WASI IQ), educational/verbal attainment (WIAT-T-III) and numeracy (WAIS-III Arithmetic). Consistent with our expectations, the results from a correlational analysis revealed the presence of a significant relationship between scores of relational responding on the RAI+ and Full Scale, Verbal, and Performance IQ on a standardized measure of intelligence, a finding that is highly consistent with previous studies (Colbert et al., 2017; Dixon et al., 2014; Gore et al., 2010; O'Hora et al., 2008; O'Toole & Barnes-Holmes, 2009). In addition, significant correlations were also observed between RAI+ scores and each of the four WASI IQ subtests. In general, these results appear to support the assertion that relational responding may play an influential role in intellectual behavior (Andrews & Halford, 1998; Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010; Gentner & Loewenstein, 2002; Halford, Wilson, & Phillips, 2010; Moran, Stewart, McElwee, & Ming, 2014; O'Hora et al., 2005; Stewart et al., 2013).

A significant correlation between RAI+ scores and WAIS-III Arithmetic scores further underline the relevance of relational skill proficiency to numeracy (Carpenter et al., 2003; Koehler, 2004; Molina, 2005; Molina, Castro, & Ambrose, 2005; Ninness et al., 2006). In terms of academic attainment, the Same/Opposite test module displayed a significant relationship with the WIAT-T index, a result predicted by the relevance of coordination relations to language acquisition (Hayes et al., 2001). It is critical to note, however, that the current study failed to identify a correlation between WIAT scores and any other RAI+ metric. This is inconsistent with the RFT perspective that relational abilities are functionally associated with academic attainment. However, there are a number of factors and limitations that may have affected this outcome. First, the WIAT-II may not have been the most appropriate assessment for the current sample, as it may be less sensitive in assessing high levels of performance (Strauss et al., 2006), and has somewhat limited utility in predicting actual school grades (Strauss et al., 2006). In effect, the question of how educational achievement relates to relational abilities perhaps cannot be solved psychometrically, but may instead require correlational analyses between relational abilities and actual academic attainment (itself a highly variable metric, bringing with it further challenges). Although this issue renders the possibility of finding significant correlations between relational ability and academic achievement less likely, this does not alter our finding that the RAI+ did not predict performance on a well-validated and widely administered academic achievement test.

Although two distinct, but closely related Wechsler measures of intelligence were administered in the Colbert et al. (2017) study and in the current study (WAIS-III and WASI, respectively), there is considerable overlap in terms of the outcomes these studies report. Although the pattern of significant relationships is similar, the strength of correlations vary between these two analyses. For example, the correlation coefficients reported for relational ability and Full Scale (.54), Verbal (.42), and Performance IQ (.48) in the current study are considerably lower than that reported in the Colbert et al. analysis (.74, .78 and .55 respectively). These studies also differ in terms of correlations between relational skills measures and the four IQ subtests shared by each IQ measure (i.e., Vocabulary; .63 and .38, respectively, Similarities; .58 and .37, respectively, Block Design; .6 and .42, respectively, and Matrix Reasoning; .48 and .42, respectively). As such, we must conclude that the addition of further relational frames into the assessment was not beneficial in improving the predictive utility of the RAI and in fact increased variance along dimensions perhaps not as strongly related to IQ as the Same, Opposite, More than, and Less than relational skills proficiencies. However, decreased sensitivity to the relevant relational skills by the WASI compared to the WAIS cannot be ruled out. Furthermore, it is still crucial to understand that the relative contributions of each relational skill to IQ and their interrelationships with each other was important to identify as part of a larger effort to elaborate a different perspective on the nature of human intelligence. In this regard, the current exercise has been informative.

In terms of what we have learned about the interrelationships between various relational skills repertoires, perhaps the most illuminating have been, first, the confirmation of More/Less, Same/Different, Same/Opposite, and Before/After as perhaps the most strongly related to IQ (Berens & Hayes, 2007; O'Hora et al., 2008; O'Toole & Barnes-Holmes, 2009; O'Toole et al., 2009; Stewart et al., 2013). Second, we have learned much from the surprising lack of correlation between analogical skills and IQ, as well as WIAT scores and WAIS arithmetic. This is highly unexpected because analogical reasoning is consistently associated with many higher cognitive skills such as abstract reasoning (Gentner, Holyoak, & Kokinov, 2001; Richland & Simms, 2015), problem solving (Gentner & Smith, 2012), creative endeavors such as writing poetry or prose (Shen & Lai, 2014) and in general is considered a ubiquitous aspect of everyday human communication (Stewart et al., 2004).

We may make sense of the latter outcome in several ways. First, it may be suggested that the limited number or type of analogical reasoning trials included in the RAI+ may not sufficiently assess subtle individual differences in this skills repertoire. However, it may also be that the WASI and WIAT have a poor representation of such tasks in their battery. In the case of the WASI, whereas Matrix Reasoning can be considered an assessment of visual-spatial analogical reasoning (Carpenter, Just, & Shell, 1990; Hunt, 1974), none of the four WASI subtests directly assess verbal analogical reasoning. In addition, despite the fact that analogical reasoning is pertinent to a number of important verbal competencies, such as reading (Farrington-Flint & Wood, 2007; Farrington-Flint, Wood, Canobi, & Faulkner, 2004), vocabulary (Edwards et al., 2011; Nippold & Sullivan, 1987), grammar (Edwards et al., 2011; Hock, 1991, 2003), and spelling (Brown et al., 1996; Ehri & Robbins, 1992; Goswami, 1988), none of the WIAT-T's three subtests of Reading, Reading Comprehension, and Spelling explicitly employ analogical tasks as a means of measuring verbal attainment.

At this point it is important to note that the aim of the current research stream is not to compose a "better" measure of IQ, but to provide a functional account of intellectual performance and an accompanying assessment tool. As such, dissimilarities in the remit of measurement and/or failures to find significant correlations do not necessarily represent a psychometric failure of the RAI+, but may in fact reflect theoretical divergence in terms of what constitutes intellectual performance. The strength of the correlations reported in the current analysis suggest that although these repertoires may be related, they are not equivalent or synonymous, at least as assessed by the testing battery administered. That global issue notwithstanding, the RFT literature would propose that due to the advanced level of complexity inherent in analogical reasoning for example, its proficiency levels should predict IQ (see Carpentier et al., 2003; McHugh et al., 2004), particularly for high ability individuals. The fact that this was not the case, may point to construct validity issues for either the RAI+ or the WASI, depending on what a priori definition of intelligence one begins with.

Second, it may be the case that analogical responding is not as relevant to intellectual performance as other "core" relational skills (e.g., same/opposite, more than/less than), a suggestion supported by the reduced level of intercorrelation between analogical trials and scores on the other RAI+ modules, as well as the finding that the removal of analogy test trials actually increases the internal consistency and predictive validity of the RAI+. The apparent distinction between "core" relational skills and analogical reasoning may be related to its unfolding in the developmental process. In particular, there is some modest evidence that relational skills of coordination and comparison (more/less) emerge first and appear to be well-established prior to the development of many higher-level relational skills (Carpentier et al., 2003). In this sense, perhaps analogy comprises part of a higher-level skill set that is still unfolding in adults, insofar as it depends upon proficiency in each of the other relations and involves learning to relate relations to each other. Future research should aim to investigate to what extent some of these skills precede or functionally overlap with each other and should attempt to map out the developmental trajectory of analogical reasoning, which may extend well into adulthood.

In assessing the distribution of RAI+ test scores, the reduced variance of scores for the current sample is noteworthy. A large proportion of our sample (29%) achieved an overall RAI+ score of 95% or above. In contrast, only one participant displayed a Full Scale IQ above the 95th percentile, and none scored more than 95% on the WASI Similarities, Vocabulary, or Matrix Reasoning subtests. The skewed distribution of RAI+ scores would therefore reduce the likelihood of significant correlation with IQ metrics, and diminish its utility as a proxy measurement on intelligence in general.

One possible mechanism for enhancing the predictive validity of the RAI+ is to ensure a wider range of scores and therefore improve the sensitivity of the test. This could be perhaps most readily achieved by reducing the time limit at either a global or per trial basis. Although previous research suggests that response fluency in general may not correlate with Full Scale IQ (Binder, 1996), from an RFT and behavior-analytic perspective, the fluency with which responding occurs is an important component of intelligent behavior (Cassidy et al., 2010; Stewart, 2016; Thorndike, Bregman, Cobb, & Woodyard, 1926). Second, more difficult tasks could be included in the RAI+ in an effort to increase its sensitivity at the top end of the scale. This could involve increasing the nodal distance of relations tested, or the number of nodes that link any two stimuli in a set of conditional relations (Sidman, 1994). For instance, most tasks in the RAI+ assess two nodal (e.g., A is the same as B, B is opposite to C) or three nodal (e.g., A is the same as B, B is the same as C, C is the same as D) relational reasoning. The addition of further nodes could be integrated into the current RAI+ and would potentially allow us to ascertain a more comprehensive profile of individual strengths and weaknesses, as well as more balanced data distributions. However, it is critical to state at this point that there is no conceptual requirement that RAI+ scores be distributed normally across the population, precisely because they index a malleable skill set that is considered to be continually in flux and therefore at varying levels in various environmental contexts across various times.

Conclusion

The aim of this present study was to test a prototype extended relational abilities index that built upon the RAI in terms of the range of relational frames it assessed. Our results indicate that although the RAI+ exhibited significant correlations with a range of IQ indices and subtests, its inclusion of additional relational frames did not improve upon the predictive validity demonstrated by the original RAI. This may not be surprising given both the previously reported high correlations between the shorter RAI and Full Scale IQ (0.74; Colbert et al., 2017), and the currently reported high Cronbach's alpha for the RAI+ (0.79). In other words, any subset of relational tasks may hold the potential to function as a useful proxy of both overall relational skills and IQ.

It is interesting that performance on the RAI+ displayed a general failure to predict educational and verbal attainment as measured by the WIAT-II-T, despite a wealth of previous theoretical and empirical work that would anticipate such relationships. The work did, however, reveal important intercorrelations between relational skills repertories, and found a respectable level of internal consistency for the RAI+. Overall, the study confirms that relational skills indices may represent useful proxies of full-scale intelligence and potentially numeracy, but that such indices bear a more complex relationship to academic aptitude.

Future studies may endeavor to provide a more comprehensive examination of the relationship between academic attainment and relational skill fluency. It is interesting that the respectable correlations obtained between RAI+ scores and the standardized measures of numeracy and full scale IQ, suggest that it is may be the well standardized nature of the WASI and WAIS indices that facilitated such correlations. The WIAT, in contrast, is not a very good predictor of school grades (Strauss et al., 2006), so its construct validity may be in question, rather than that of the RAI+. In addition, the most important measure of academic attainment is actual scholastic performance, and it is more fitting for a behavioral science to validate a proxy measure for academic attainment against real school performance, than against further proxies for the same. In the meantime, the RAI+ is not ready for use as a proxy measure of academic ability but would appear to hold promise as a functionally understood, behavior-analytically acceptable proxy for assessing intellectual capacity. If this is so, we have moved some way forward in developing a progressive behavior-analytic, functionally understood assessment of that broad skill set widely referred to as "intelligence."

Appendix A

Table 3 The format and sequence of all 67 RAI+ trials

Block	No	Premise 1			Premise 2			Premise 3			Question		
1	1	a	same as	b							a	opposite to	b
	2	a	same as	b	b	same as	c				a	same as	c
	3	a	same as	b	b	same as	c				a	same as	c
	4	a	opposite to	b	b	opposite to	c				c	same as	b
	5	a	same as	b	b	same as	c				c	same as	a
	6	a	same as	b	c	same as	a				b	opposite to	c
	7	a	opposite to	b	b	opposite to	c				c	same as	a
	8	a	opposite to	b	c	opposite to	a				a	same as	c
	9	a	opposite to	b	c	opposite to	a				b	same as	a
	10	a	opposite to	b	b	same as	c				c	opposite to	a
	11	a	opposite to	b	c	same as	a				a	opposite to	c
	12	a	same as	b	b	same as	c	c	same as	d	d	opposite to	b
	13	a	opposite to	b	b	opposite to	c	c	opposite to	d	d	opposite to	a
	14	a	same as	b	b	opposite to	c	c	opposite to	d	d	same as	a
	15	a	opposite to	b	b	opposite to	c	c	same as	d	b	opposite to	c
2	16	a	different to	b							b	same as	a
	17	a	same as	b	b	same as	c				c	same as	a
	18	a	different to	b	b	different to	c				b	same as	c
	19	a	different to	b	b	different to	c				a	same as	c
	20	a	same as	b	c	same as	a				b	same as	c
	21	a	same as	b	c	same as	a				b	same as	c
	22	a	different to	b	b	different to	c				a	same as	b
	23	a	different to	b	c	different to	a				b	different to	a
	24	a	different to	b	b	same as	c				c	same as	a
	25	a	same as	b	b	different to	c				a	different to	b
	26	a	same as	b	b	same as	c	c	same as	d	b	same as	c
	27	a	different to	b	b	different to	c	c	different to	d	b	different to	a
	28	a	same as	b	b	same as	c	c	different to	d	b	same as	c
	29	a	different to	b	b	same as	c	c	same as	d	a	different to	c
3	30	a	more than	b							b	more than	a
	31	a	more than	b	b	more than	c				a	more than	b
	32	a	more than	b	b	more than	c				b	less than	c
	33	a	less than	b	b	less than	c				a	less than	c
	34	a	less than	b	b	less than	c				b	more than	c
	35	a	more than	b	c	more than	a				a	more than	c
	36	a	more than	b	c	more than	a				c	more than	b
	37	a	less than	b	c	less than	A				c	more than	a
	38	a	more than	b	b	more than	c	c	more than	d	a	more than	c
	39	a	less than	b	b	less than	c	c	less than	d	c	less than	b
	40	a	more than	b	c	more than	b	d	more than	c	d	less than	a

Table 3 (continued)

Block	No	Premise 1			Premise 2			Premise 3			Question		
	41	a	more than	b	c	more than	a	d	more than	c	d	less than	b
	42	a	less than	b	c	less than	a	d	less than	b	d	less than	b
4	43	a	after	b							a	after	b
	44	a	before	b	b	before	c				c	before	b
	45	a	before	b	b	before	c				c	before	a
	46	a	after	b	b	after	c				b	after	a
	47	a	after	b	b	after	c				c	before	a
	48	a	before	b	c	before	a				c	after	b
	49	a	after	b	c	after	a				c	after	a
	50	a	after	b	c	after	a				b	before	c
	51	a	before	b	b	before	c	c	before	d	d	before	b
	52	a	after	b	b	after	c	c	after	d	b	before	d
	53	a	before	b	c	before	b	d	before	c	b	after	d
	54	a	after	b	c	after	a	d	after	c	a	before	d
	55	a	after	b	c	after	a	d	after	c	b	before	d
5	56	a	same as	b	c	same as	d				a/b	same as	c/d
	57	a	opposite to	b	c	opposite to	d				b/a	different to	c/d
	58	a	before	b	c	before	d				a/b	same as	c/d
	59	a	before	b	c	before	d				b/a	different to	c/d
	60	a	same as	b	c	opposite to	d				a/b	same as	c/d
	61	a	opposite to	b	c	same as	d				b/a	different to	c/d
	62	a	after	b	c	before	d				a/b	same as	c/d
	63	a	after	b	c	before	d				b/a	different to	c/d
	64	a	more than	b	c	more than	d				a/b	same as	c/d
	65	a	less than	b	c	less than	d				b/a	same as	c/d
	66	a	more than	b	c	less than	d				a/b	same as	c/d
	67	a	less than	b	c	more than	d				a/b	same as	c/d

References

Andrews, G., & Halford, G. S. (1998). Children's ability to make transitive inferences: The importance of premise integration and structural complexity. *Cognitive Development, 13*, 479–513.

Barnes, D., McCullagh, P., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *Analysis of Verbal Behaviour, 8*, 1–11.

Barnes-Holmes, D., Barnes-Holmes, Y., Smeets, P. M., Cullinan, V., & Leader, G. (2004b). Relational frame theory and stimulus equivalence: Conceptual and procedural issues. *International Journal of Psychology & Psychological Therapy, 4*(2), 181–214.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *Psychological Record, 60*, 527–542.

Barnes-Holmes, D., Finn, M., McEnteggart, C., & Barnes-Holmes, Y. (2017). Derived stimulus relations and their role in a behavior-analytic account of human language and cognition. *Perspectives on Behavioral Science, 41*(1), 155–173.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., Healy, O., Lyddy, F., Cullinan, V., & Hayes, S. C. (2001). Psychological development. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of language and cognition* (pp. 157–180). New York, NY: Plenum Press.

Barnes-Holmes, Y., Barnes-Holmes, D., & Smeets, P. M. (2004a). Establishing relational responding in accordance with opposite as generalised operant behaviour in young children. *International Journal of Psychology & Psychological Therapy, 4*, 559–586.

Basak, C., Boot, W. R., Voss, M. W., & Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychological Ageing, 23*(4), 765–777.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1990). *Differential aptitude tests* (5th ed.). New York, NY: Psychological Corporation.

Berens, N., & Hayes, S. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behaviour Analysis, 40*(1), 45–71.

Binder, C. (1996). Behavioural fluency: Evolution of a new paradigm. *Behavior Analyst, 19*(2), 163–197.

Blair, C., Gamson, D., Thorne, S., & Baker, D. (2005). Rising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. *Intelligence, 33*(1), 93–106.

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

Brand, C. R. (1987). Bryter still and bryter? *Nature, 328*, 110.

Brinch, C. N., & Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences, 109*(2), 425–430.

Brown, K. J., Sinatra, G. M., & Wagstaff, J. M. (1996). Exploring the potential of analogy instruction to support children's spelling development. *Elementary School Journal, 97*, 81–99.

Butler, M., Retzlaff, P., & Vanderploeg, R. (1991). Neuropsychological test usage. *Professional Psychology: Research & Practice, 22*, 510–512.

Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J. J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecdarian Project. *Applied Developmental Science, 6*, 42–57.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*(3), 404–431.

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school*. Toronto, Canada: Pearson Education Canada.

Carpentier, F., Smeets, P. M., & Barnes-Holmes, D. (2003). Equivalence-equivalence as a model of analogy: Further analyses. *The Psychological Record, 53*, 349–372.

Cassidy, S. (2008). Relational frame theory and human intelligence: A conceptual and empirical analysis. (Unpublished PhD dissertation). Maynooth University, Ireland.

Cassidy, S., Roche, B., Colbert, D., Stewart, I., & Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learning & Individual Differences, 47*, 222–235.

Cassidy, S., Roche, B., & Hayes, S. C. (2011). A relational frame training intervention to raise intelligence quotients: A pilot study. *The Psychological Record, 61*, 173–198.

Cassidy, S., Roche, B., & O'Hora, D. (2010). Relational frame theory and intelligence. *European Journal of Behaviour Analysis, 11*, 37–52.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive improvements? A reassessment of the evidence. *Developmental Psychology, 27*(5), 359–381.

Colbert, D., Dobutowitsch, M., Roche, B., & Brophy, C. (2017). The proxy-measurement of intelligence quotients using a relational skills abilities index. *Learning & Individual Differences, 57*, 114–122.

Colbert, D., Tyndall, I., Roche, B., & Cassidy, S. (2018). Can SMART training really increase intelligence? A replication study. *Journal of Behavioural Education, 2*, 1–23.

Cowley, B. J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to teach name-face matching to adults with brain injuries. *Journal of Applied Behavior Analysis, 25*, 461–475.

Cullinan, V., Barnes, D., Hampson, P. J., & Lyddy, F. (1994). A transfer of explicitly and nonexplicitly trained sequence responses through equivalence relations: An experimental demonstration and connectionist model. *The Psychological Record, 44*, 559–585.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13–21.

de Rose, J. C. C., de Souza, D. G., Rossito, A. L., & de Rose, T. M. S. (1992). Stimulus equivalence and generalisation in reading after matching to sample by exclusion. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 69–82). Reno, NV: Context Press.

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review, 20*, 631–642.
Dickens, W., & Flynn, J. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346–369.
Dixon, M. R. (2015). *The PEAK relational training system: Equivalence module*. Carbondale, IL: Shawnee Scientific Press.
Dixon, M., Belisle, J., & Stanley, C. R. (2018). Derived relational responding and intelligence: Assessing the relationship between the PEAK-E pre-assessment and IQ with individuals with autism and related disabilities. *The Psychological Record, 68*, 1–12.
Dixon, M., Whiting, S., Rowsey, K., & Belisly, J. (2014). Assessing the relationship between intelligence and the PEAK relational training system. *Research in Autism Spectrum Disorders, 8*(9), 1208–1213.
Dugdale, N., & Lowe, C. F. (2000). Testing for symmetry in the conditional discriminations of language-trained chimpanzees. *Journal of the Experimental Analysis of Behaviour, 73*(1), 5–22.
Dutton, E., & Lynn, R. (2015). A negative Flynn Effect in France, 1999–2008/2009. *Intelligence, 51*, 67–70.
Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behaviour, 64*, 163–184.
Dymond, S., & Roche, B. (2013). *Advances in relational frame theory*. Oakland, CA: Context Press.
Edwards, L., Figueras, B., Mellanby, J., & Langdon, D. (2011). Verbal and spatial analogical reasoning in deaf and hearing children: The role of grammar and vocabulary. *Journal of Deaf Studies & Deaf Education, 16*(2), 189–197.
Ehri, L. C., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. *Reading Research Quarterly, 27*, 12–28.
Elliot, C. D. (2007). *Differential ability scales (2nd ed.): Introductory and technical handbook*. San Antonio, TX: Psychological Corporation.
Farrington-Flint, L., & Wood, C. (2007). The role of lexical analogies in beginning reading: Insights from children's self-reports. *Journal of Educational Psychology, 99*(2), 326–338.
Farrington-Flint, L., Wood, C., Canobi, K. H., & Faulkner, D. (2004). Patterns of analogical reasoning among beginning readers. *Journal of Research in Reading, 27*(3), 226–247.
Flynn, J. R. (1984). The mean IQ of Americans: Massive gains, 1932–1978. *Psychological Bulletin, 95*(1), 29–51.
Flynn, J. R. (1998). WAIS-III and WISC-III gains in the United States from 1972–1995: How to compensate for obsolete norms. *Perceptual & Motor Skills, 86*, 1231–1239.
Flynn, J. R. (2007). *What is intelligence?* Cambridge, UK: Cambridge University Press.
Flynn, J. R. (2009). Requiem for nutrition as the cause of IQ gains: Raven's gains in Britain, 1938–2008. *Economics & Human Biology, 7*, 18–27.
Gentner, D., Holyoak, K. J., & Kokinov, B. N. (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.
Gentner, D., & Loewenstein, J. (2002). Relational language and relational thought. In J. Byrnes & E. Amsel (Eds.), *Language, literacy, and cognitive development* (pp. 87–120). Hillsdale, NJ: Lawrence Erlbaum Associates.
Gentner, D., & Smith, L. (2012). Analogical reasoning. In V. S. Ramachandran (Ed.), *Encyclopedia of human behaviour* (pp. 130–136). Oxford, UK: Elsevier.
Gore, N. J., Barnes-Holmes, Y., & Murphy, G. (2010). The relationship between intellectual functioning and relatinal perspective-taking. *International Journal of Psychology & Psychological Therapy, 10*, 1–17.
Goswami, U. (1986). Children's use of analogy in learning to read: A developmental study. *Journal of Experimental Child Psychology, 42*, 73–83.
Goswami, U. (1988). Children's use of analogy in learning to spell. *British Journal of Developmental Psychology, 6*, 21–33.
Gottfredson, L. S. (1998). The general intelligence factor. *Scientific American Presents, 9*(4), 24–29.
Griffee, K., & Dougher, M. J. (2002). Contextual control of stimulus generalisation and stimulus equivalence in hierarchical categorisation. *Journal of Experimental Analysis of Behaviour, 78*, 433–447.
Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Science, 14*(11), 497–505.
Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evolution of behavior analysis. *American Psychologist, 47*, 1383–1395.
Hayes, J., & Stewart, I. (2016). Comparing the effects of derived relational training and computer coding on intellectual potential in school-age children. *British Journal of Educational Psychology, 86*(3), 397–411.

Hayes, S., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory.* New York, NY: Kluwer Academic/Plenum Publishers.

Hock, H. H. (1991). *Principles of historical linguistics* (2nd ed.). Berlin, Germany: Mouton de Gruyter.

Hock, H. H. (2003). Analogical change. In B. Josephs & R. Janda (Eds.), *The handbook of historical linguistics* (pp. 441–480). Oxford, UK: Blackwell.

Howe, M. J. A. (1990). Does intelligence exist? *The Psychologist, 3*, 490–493.

Hunt, E. (1974). Quote the raven? Nevermore! In L. W. Gregg (Ed.), *Knowledge and cognition* (pp. 129–158). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., Peterson, R. C., Kokmen, E., & Kurland, L. T. (1992). Mayo's Older American Normative Studies: WAIS-R norms for ages 56 to 97. *The Clinical Neuropsychologist, 6*(Supplement), 1–30.

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

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108*(25), 10081–10086.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory, 18*(4), 394–412.

Jencks, C. (1972). *Inequality.* New York, NY: Basic Books.

Jensen, A. (1980). Chronometric analysis of intelligence. *Journal of Social & Biological Structures, 3*, 103–122.

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

Juliano, J., Haddad, F., & Carroll, J. (1988). Three-year stability of WISC-R factor scores for black and white, female and male children classified as learning-disabled. *Journal of School Psychology, 26*(4), 317–325.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence, 40*(2), 123–138.

Koehler, J. L. (2004). *Learning to think relationally: Thinking relationally to learn*. Dissertation research proposal. University of Wisconsin-Madison.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.

Locurto, C. (1991). *Sense and nonsense about IQ: The case for uniqueness*. New York, NY: Praeger.

Lynn, R. (1990). Differential rates of secular increase of five major primary abilities. *Social Biology, 37*(1–2), 137–141.

Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence, 37*(1), 16–24.

Lynn, R., & Meisenberg, G. (2010). The average IQ of sub-Saharan Africans: Comments on Wicherts, Dolan, and Van der Maas. *Intelligence, 38*, 21–29.

Mackay, H. A. (1985). Stimulus equivalence in rudimentary reading and spelling. *Analysis & Intervention in Developmental Disabilities, 5*(4), 373–387.

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science, 14*(3), 582–590.

Mackey, A. P., Miller-Singley, A. T., & Bunge, S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *Journal of Neuroscience, 33*, 4796–4803.

Mackey, A. P., Whittaker, K. J., & Bunge, S. A. (2012). Experience-dependent plasticity in white matter microstructure: Reasoning training alters structural connectivity. *Frontiers in Neuroanatomy, 6*, 32.

Mackintosh, M. J. (1998). *IQ and human intelligence.* Oxford, UK: Oxford University Press.

Matos, M. A., & d'Oliveira, M. M. H. (1992). Equivalence relations in reading. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 83–94). Reno NV: Context Press.

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *The Psychological Record, 54*, 115–144.

Mingroni, M. A. (2007). Resolving the IQ paradox: heterosis as a cause of the Flynn effect and other trends. *Psychological Review, 114*(3), 806–829.

Molina, M. (2005). Resolución de Igualdades Numéricas por Estudiantes de Tercer Grado. Un Esudio sobre la compression del signo igual y el desarrollo de pensamiento 9 relacional [Resolution of Numerical Equality in Third Grade students. A study on the compression of the equal sign and the development of relational thinking] (Guided Research Work). Granada University, Spain.

Molina, M., Castro, E., & Ambrose, R. (2005). Enriching arithmetic learning by promoting relational thinking. *International Journal of Learning, 12*(5), 265.

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2010). Brief Report: The training and Assessment of Relational Precursors and Abilities (TARPA): A preliminary analysis. *Journal of Autism & Developmental Disorders, 40*(9), 1149–1153.

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2014). Relational ability and language performance in children with autism spectrum disorders and typically developing children: A further test of the TARPA protocol. *The Psychological Record, 64*(2), 233–251.

Murphy, C., & Barnes-Holmes, D. (2010a). Establishing complex derived manding withchildren with and without a diagnosis of autism. *The Psychological Record, 60*, 489–504.

Murphy, C., & Barnes-Holmes, D. (2010b). Establishing five derived mands in three adolescent boys with autism. *Journal of Applied Behavior Analysis, 43*, 537–541.

Murphy, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2003). Manding and derived transformation of function in children with autism spectrum disorder. Paper presented at the annual convention of the Association for Behavior Analysis, San Francisco, USA.

Naglieri, J. A. (2008). Traditional IQ: 100 years of misconception and its relationship to minority representation in gifted programs. In J. VanTassel-Baska (Ed.), *Critical issues in equity and excellence in gifted education series, alternative assessment of gifted learners* (pp. 67–88). Waco, TX: Prufrock Press.

Naglieri, J. A., & Bornstein, B. T. (2003). Intelligence and achievement: Just how correlated are they? *Journal of Psychoeducational Assessment, 21*(3), 244–260.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77–101.

Nelson, H. E. (1982). *National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual*. Windsor, UK: NFER-Nelson.

Ninness, C., Rumph, R., McCuller, G., Vasquez, E., Harrison, C., Ford, A. M., et al. (2005a). A relational frame and artificial neural network approach to computer-interactive mathematics. *The Psychological Record, 55*(1), 135–153.

Ninness, C., Rumph, R., McCuller, G., Harrison, C., Ford, A. M., & Ninness, S. K. (2005b). A functional analytic approach to computer-interactive mathematics. *Journal of Applied Behavior Analysis, 38*(1), 1–22.

Ninness, C., Barnes-Holmes, D., Rumph, R., McCuller, G., Ford, A. M., Payne, R., et al. (2006). Transformations of mathematical and stimulus functions. *Journal of Applied Behavior Analysis, 39*(3), 299–321.

Ninness, C., Dixon, M., Barnes-Holmes, D., Rehfeldt, R. A., Rumph, R., McCuller, G., et al. (2009). Constructing and deriving reciprocal trigonometric relations: A functional analytic approach. *Journal of Applied Behavior Analysis, 42*(2), 191–208.

Nippold, M. A., & Sullivan, M. P. (1987). Verbal and perceptual analogical reasoning and proportional metaphor comprehension in young children. *Journal of Speech & Hearing Research, 30*(3), 367–376.

Nisbett, R., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D., & Turkheimer, E. (2012). Group differences in IQ are best understood as environmental in origin. *American Psychologist, 67*(6), 503–504.

O'Hora, D., Pelaez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-III. *The Psychological Record, 55*, 155–175.

O'Hora, D., Pelaez, M., Barnes-Holmes, D., Rae, G., Robinson, T., & Chaudhary, T. (2008). Temporal relations and intelligence: Correlating relational performance with performance on the WAIS-III. *The Psychological Record, 58*, 569–583.

O'Hora, D., Roche, B., Barnes-Holmes, D., & Smeets, P. M. (2002). Response latencies to multiple derived stimulus relations: Testing two predictions of relational frame theory. *The Psychological Record, 52*, 51–76.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training on working memory. *Nature Neuroscience, 7*(1), 75–79.

O'Toole, C., & Barnes-Holmes, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record, 59*, 119–132.

O'Toole, C., Barnes-Holmes, D., Murphy, C., O'Connor, J., & Barnes-Holmes, Y. (2009). Relational flexibility and intelligence: Extending the remit of Skinner's Verbal Behaviour. *International Journal of Psychology & Psychological Therapy, 9*, 1–17.

Pearson Education. (2006a). *Wechsler individual achievement test second UK edition for teachers*. London, UK: Pearson Assessment.

Pearson Education. (2006b). *Wechsler individual achievement test second UK edition for teachers: Examiner's manual*. London, UK: Pearson Assessment.

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn Effect (1909–2013). *Perspectives on Psychological Science, 10*(3), 282–306.

Ramsden, S., Richardson, F. M., Josse, G., Thomas, M. S. C., Ellis, C., Shakeshaft, C., et al. (2011). Verbal and non-verbal intelligence changes in the teenage brain. *Nature, 479*, 113–116.

Rehfeldt, R. A., & Barnes-Holmes, Y. (2009). *Derived relational responding: Applications for learners with autism and other developmental disabilities*. Oakland, CA: New Harbinger Publications.

Rehfeldt, R. A., & Root, S. L. (2005). Establishing derived requesting skills in adults with severe developmental disabilities. *Journal of Applied Behavior Analysis, 38*(1), 101–105.

Rey, A. (1958). *L'examen Clinique en psychologie [The Clinical examination in psychology]*. Paris, FR: Presses Universitaires de France.

Reynolds, C. R., Gutkin, T. B., Dappen, L., & Wright, D. (1979). Differential validity of the WISC-R for boys and girls referred for psychological services. *Perceptual & Motor Skills, 48*(3), 868–870.

Richardson, K. (2002). What IQ tests test. *Theory & Psychology, 12*(3), 283–314.

Richardson, K., & Norgate, S. H. (2015). Does IQ really predict job performance? *Applied Developmental Science, 19*(3), 153–169.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *WIREs Cognitive Science, 6*(2), 177–192.

Rindermann, H. (2007). The *g*-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21*(5), 667–706.

Roche, B., & Barnes, D. (1997). A transformation of respondently conditioned stimulus function in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behaviour, 67*, 275–300.

Roid, G. H. (2003). *Stanford Binet intelligence scales*. Itasca, IL: Riverside Publishing.

Rosales, R., Rehfeldt, R. A., & Lovett, S. (2011). Effects of multiple exemplar training on the emergence of derived relations in preschool children learning a second language. *Analysis of Verbal Behaviour, 27*(1), 61–74.

Schlinger, H. D. (1993). Separating discriminative and function-altering effects of verbal stimuli. *Association for Behaviour Analysis International, 16*(1), 9–23.

Schlinger, H. D. (2003). The myth of intelligence. *The Psychological Record, 53*(1), 15.

Schrank, F. A., McGrew, K. S., Mather, N., & Woodcock, R. W. (2014). *Woodcock-Johnson IV.* Rolling Meadows, IL: Riverside Publishing.

Shayer, M., & Ginsburg, D. (2009). Thirty years on: A large anti-Flynn effect? (II): 13- and 14-year-olds. Piagetian tests of formal operations norms 1976–2006/7. *British Journal of Educational Psychology, 79*(Pt. 3), 409–418.

Shayer, M., Ginsburg, D., & Coe, R. (2007). Thirty years on: A large anti-Flynn effect? The Piagetian test *Volume & Heaviness* norms 1975–2003. *British Journal of Educational Psychology, 77*, 25–41.

Shen, T., & Lai, J. C. (2014). Exploring the relationship between creative tests of ATTA and the Thinking of Creative Works. *Procedia: Social and Behavioural Sciences, 112*, 557–566.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech, Language, & Hearing Research, 14*(1), 5–13.

Sidman, M. (1994). *Equivalence relations and behaviour*. Boston, MA: Authors Cooperative.

Skinner, B. F. (1974). *About behaviourism*. New York, NY: Alfred A. Knopf.

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

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

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms and commentary* (3rd ed.). Oxford, UK: Oxford University Press.

Stephenson, C. L., & Halpern, D. F. (2013). Improved matrix reasoning is limited to tasks with a visuospatial component. *Intelligence, 41*, 341–357.

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences, 105*(19), 6791–6792.

Stewart, I. (2016). The fruits of a functional approach for psychological science. *International Journal of Psychology, 51*, 15–27.

Stewart, I., Barnes-Holmes, D., & Roche, B. (2004). A functional-analytic model of analogy using the relational evaluation procedure. *The Psychological Record, 54*, 531–552.

Stewart, I., & McElwee, J. (2009). Relational responding and conditional discrimination procedures: An apparent inconsistency and clarification. *The Behavior Analyst, 32*, 309–317.

Stewart, I., Tarbox, J., Roche, B., & O'Hora, D. (2013). Education, intellectual development, and Relational Frame Theory. In S. Dymond & B. Roche (Eds.), Advances in relational frame theory-research and application (pp. 177–197). Oakland, CA: New Harbinger Publications.

Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence, 32*, 349–362.

Taylor, E. M. (1959). *Psychological appraisal of children with cerebral defects*. Cambridge, MA: Harvard University Press.

Teasdale, T. W., & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence, 36*(2), 121–126.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale*. Boston, MA: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence*. Boston, MA: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1960). *Stanford-Binet intelligence scale: Manual for the third revision form L-M. Stanford-Binet intelligence scale: Manual for the third revision, Form L-M*. Boston, MA: Houghton Mifflin.

Thorndike, E. L., Bregman, E. O., Cobb, M. V., & Woodyard, E. (1926). *The measurement of intelligence*. New York, NY: Teachers College Bureau of Publications.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale* (4th ed.). Itasca, IL: Riverside Publishing.

van der Maas, H. L. J., Kan, K. J., & Borsboom, D. (2014). Intelligence is what the intelligence test measures. *Seriously. Journal of Intelligence, 2*(1), 12–15.

Wai, J., & Putallaz, M. (2011). The Flynn effect puzzle: A 30-year examination from the right tail of the ability distribution provides some missing pieces. *Intelligence, 39*, 443–455.

Wechsler, D. (1955). *Wechsler adult intelligence scale manual*. New York, NY: Psychological Corporation.

Wechsler, D. (1981). *Wechsler adult intelligence scale (rev. manual)*. New York, NY: Psychological Corporation.

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: Psychological Corporation. https://doi.org/10.1177/1073191102009001003.

Wechsler, D. (1998). *Wechsler adult intelligence scale (3rd ed)* (UK ed.). London, UK: Psychological Corporation.

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

Wechsler, D. (2008). *Wechsler adult intelligence scale* (4th ed.). Statistics Solutions.

Woodley, M. A., & Meisenberg, G. (2013). In the Netherlands the anti-Flynn effect is a Jensen effect. *Personality & Individual Differences, 54*(8), 871–876.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*, 125–144.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.