Assessing the Efficacy of a Relational Skills Training Intervention in Improving Intellectual Function in a Sample of High IQ Adults.

Dylan Colbert, B.A. (Hons)

Thesis submitted in fulfilment of a Master of Science (M.Sc.) degree to the Department of Psychology, Maynooth University.

October 2015

Supervised by Dr. Bryan Roche

Head of Department: Dr. Andrew Coogan

Table of Contents

List of Tables

List of Figures

Declaration

I, the undersigned, herby certify that this material, which I now submit in fulfillment of a M.Sc. degree, has not been previously submitted as an exercise for a degree at this or any other University, and is, unless otherwise stated, entirely my own work.

Signed: ____________________________

Dylan Colbert

Student ID: 10388537

Date: 29th October 2015

Acknowledgement

First of all, I would like to thank my supervisor, Dr. Bryan Roche. Your boundless energy, enthusiasm and expertise have been an inspiration. The time and energy you have graciously afforded to me has gone above and beyond what could be expected.

I would also like to express my gratitude to the wider faculty of the Department of Psychology for providing a stimulating and entertaining environment for research. To my fellow postgraduates, both past and present, thank you for sharing the struggle! I would also like to express sincere thanks to Dr. Sarah Cassidy for providing me with expert guidance, as well as invaluable opportunities to learn.

I am also extremely indebted to my extended friends and family for supporting me throughout this process. To Holly, the understanding and encouragement you have selflessly shown to me throughout this process has been more important than you know. Cian, you were kind of just around, but I suppose you helped too.

Most of all, I would to thank my parents. Not once have you ever allowed a barrier to be put in front of me. For that I am eternally and indescribably grateful. Your unwavering support, understanding and belief in me makes this yours as much as mine.

Abstract

Recent advances in Relational Frame Theory have proposed that levels of sophistication with relational concepts may be of fundamental importance to intelligent behaviour. Furthermore, previous investigations have implicated the efficacy of relational skills intervention in improving intellectual performance, as measured by traditional IQ assessments. The current thesis aimed to extend upon such research by further assessing the contribution of relational ability to intelligence, as well as the effect of relational skills training on intellectual performance. In the first study, thirty-four high IQ college students were recruited to assess the effectiveness of a relational training protocol in increasing WAIS-III IQ scores when compared to a non-intervention group. The current analysis failed to report a significant effect of relational training in increasing scores on the Full, Verbal or Performance IQ. While there was a significant effect of training on Verbal Comprehension scores, this effect was not found for the other IQ subindices.

The second study involved a correlational analysis of WAIS-III IQ scores and relational ability scores. Results indicated a high level of correlation between these measures, further proposing the importance of relational responding to intelligence. In addition, this investigation aimed to further explore possible reasons for the diminished effect of relational training in improving intellectual performance in the first study. It must be noted that while pronounced practice effects found for the control group rendered between-group differences insignificant, there were qualitatively different rises witnessed in subtest scores. Specifically, while both groups displayed significant increases on subtests that involved a timed element, only the experimental group displayed significant increases on untimed subtests. Furthermore, high IQ was not found to significantly correlate with post-intervention IQ rises. However, the current sample displayed IQ gains significantly lower than that witnessed in average IQ cohorts. While the current analysis identifies a number of

possible boundary conditions of the current intervention, much remains to be understood in terms of variables that may exert an influence on the effectiveness of relational training in improving intellectual performance.

Chapter 1:

Literature Review and General Introduction

1.1 Introduction

Intelligence is commonly believed as lying beyond the remit and descriptive powers of the behaviourist (Abramson, 2013; Block, 1981; Putnam, 1975b; Schlinger, 2003). Theoretical objections to hypothetical constructs (Skinner, 1974), its preference for functional accounts, as well as Skinner's inability to account for the generativity of language, are often employed to support the argument that a behaviouristic account of intelligence is not only difficult technically, but inappropriate conceptually. Nevertheless, interventions based upon behaviourist principles have shown considerable utility in improving the fluency, sensitivity and flexibility of behaviours assessed in IQ tests (Cohen, Amerine-Dickens & Smith, 2006; Eikeseth, Smith, Jahr & Eldevik, 2002; Lovaas, 1987; Remington et al., 2007). Most notably, advances in Relational Frame Theory (RFT; Hayes, Barnes-Holmes & Roche, 2001; see also Dymond & Roche, 2013) have led to profound new insights on intellectual behaviour (e.g., Hayes, 1994; O'Hora, Pelaez & Barnes-Holmes, 2005; O'Toole, Murphy, Barnes-Holmes, O'Connor & Barnes-Holmes, 2009; Smith, Smith, Taylor & Hobby, 2005) as well the development of intervention protocols that have shown early promise in increasing intelligence quotients (Cassidy, Roche, Colbert, Stewart, & Grey, 2016; Cassidy, Roche & Hayes, 2011; Dixon, Whiting, Rowsey & Belisly, 2014; Moran, Stewart, McElwee & Ming, 2010).

This apparent success in affecting skills and processes which are deemed "intelligent" would seem to justify a reconsideration of the nature of intelligence from a functional and behavioural point of view, and a serious examination of the methodologies designed to increase it. These recent advances in developing a behaviouristic account of intelligence provide stark contrast to much of the mainstream theory that preceded them. As such, in order to understand the context in which this new stream of research has emerged, a brief description of the history of the conceptualisation and measurement of intelligence is warranted.

1.2 Conceptualisations of Intelligence & IQ

Over the past 130 years, few research areas have attracted as much attention and interest as the study of intelligence. While there have been several concerted attempts to extract a general definition (Flynn, 2007; Legg & Hutter, 2007; Sternberg, 1996; Sternberg & Detterman, 1987; Van der Maas, Kan & Borsboom, 2014), the scientific community is yet to reach consensus over a definition of intelligence. Sternberg's (2000) assertion that there appears to be as many definitions of intelligence as experts trying to define it provides the perfect summation of the current state of intelligence theory.

A 1921 study, carried out by the editors of the Journal of Educational Psychology, may be one of the earliest attempts to formulate an operational definition of intelligence. The study surveyed the opinions of several of the eminent researchers of the time, including Thorndike, Terman and Thurstone and found relatively little agreement between each contributor's description of intelligence. A replication of this study (Sternberg & Detterman, 1986) involved contributions from esteemed intelligence researchers such as Eysenck, Jensen and Gardner. In a comparison of the two studies, Sternberg (2000) suggested that most prominent conceptualisations of intelligence involve adaptation to the environment, basic mental process and high-level thinking, for example, reasoning, problem solving and decision making. More recently, Legg and Hutter's "A Collection of Definitions of Intelligence" (2006) identified two common properties of many current definitions of intelligence. Firstly, intelligence is regarded as an adaptive mechanism, which allows the individual to interact with its environment successfully to a greater or lesser degree. Secondly, intelligence is deemed to refer to an individual's ability to succeed in achieving a desired goal or objective. Indeed it can be proposed with reasonable accuracy that most conceptualisations of intelligence implicate the ability to learn, adapt, solve problems and reason. Perhaps the closest current approximation to a definition of intelligence is provided by the American Psychological Association (Neisser et al. 1996), defining intelligence as:

> The ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning and to overcome obstacles by taking thought. (p. 1)

In the absence of a unanimous conceptualisation, the current thesis will adopt this definition as a widely-accepted description of intelligence.

Boring's (1923) infamous assertion that the only useful definition of intelligence is "what tests of intelligence test" is perhaps yet to be bettered as researchers attempt to uncover what intelligence constitutes. Due to reification of *g* (discussed in detail in Section 1.4), along with the increasingly inextricable tie between intelligence theory and IQ, intelligence is generally assumed to be a normally-distributed trait that is fixed and unmalleable (Gardner, 1993; Gottfredson, 1998). As scores on intelligence tests can predict a number of social and economic outcomes (Ali et al., 2001; Herrnstein & Murray, 1994; Neisser et al., 1996), while also identifying learning and developmental disabilities, this utility is extended to suggest that IQ score represents an essential trait that may explain individual differences. Traditional intelligence assessments also display impressive levels of predictive validity across numerous contexts (Hartlage & Steele, 1977; Juliano, Haddad & Carroll, 1988; Reschly & Rechsly, 1979; Reynolds, Gutkin, Dappen & Wright, 1979). Furthermore, the fact that IQ score appears to remain stable and unchanged across an individual's lifetime (Sattler, 1998) offers further endorsement of the conceptualisation for intelligence as an immutable personal trait.

The current section aims to highlight a number of issues regarding trait theories of intelligence. The first issue refers to the misconception that IQ provides support for the conceptualisation of intelligence as a trait. As discussed by Cassidy, Roche & O'Hora (2010), Mayr's (2001) distinction between population thinking and essentialist thinking is particularly useful in regards to intelligence theory as this dichotomy characterises the two main approaches to intelligence research. Population thinking refers to the practice of attempting to identify variation and diversity at the population level, while essentialist thinking focuses on functional relationships and properties (Cassidy, Roche & O'Hora, 2010). The fact that intelligence as a concept has become increasingly indistinguishable from IQ, a product of population thinking, has led to a relative dearth of research on what intelligence or intelligent behaviour constitutes. It has also led to the erroneous assumption that intelligence, due to its close relationship with psychometric IQ, is also a population-level concept that represents an essential property of a given individual. Psychometric measures of intelligence remain useful indicators regarding an individual's performance in a number of domains that are considered to constitute intelligence, but this measurement is always relative to his/her peers (Cassidy et al. 2010). As such, this trait concept of intelligence is decidedly indirect in measuring intelligence. This issue will be dealt with in greater detail in Section 1.4.

The finding that IQ scores generally remain stable through an individual's lifetime somewhat masks the reality that while standardised scores do not change significantly over time, raw scores do (Gottfredson 2008; Gottfredson & Saklofske 2009). Therefore, an individual's actual performance on IQ tests may vary, even if their relative performance does not. As such, the trait that IQ measures propose to measure is open to variation by a considerable extent. Improvements in raw scores are corrected statistically to keep standardised scores constant by adjusting for the natural increases in score brought about by maturation and practice. Due to this practice, only extremely large increases in raw test performance will manifest itself as an increase in standardised score. Put simply, part of the stability of IQ scores across lifespan is due to statistical manipulation, rather than the immutability of intelligence.

Furthermore, standardised tables for IQ tests are regularly adjusted in order to counteract the progressive increase in IQ scores across the past century (see the Flynn Effect, Flynn, 1998, 2007). The main objective of these modifications is to retain the normally distributed bell curve of intelligence that psychometry places such a heavy emphasis on. The utility of this practice is understandable, and indeed extremely practical in ensuring the predictive validity and descriptive power of an IQ score. However, it demonstrates that IQ is a socially constructed concept, rather than an objective and accurate measurement of a given trait or faculty. While population thinking underlies the field of psychometrics, behaviour analysis is characterised by a heavy emphasis on essential thinking. Therefore, the field of behaviour analysis, as will be further discussed, has much to offer in the domain of intelligence research.

The purpose of the current discussion is not to undermine the utility of IQ tests, but rather to reorient common conceptions of what they can provide, i.e. an estimation of an individual's relative ability in a number of domains that are generally considered to be of importance to intelligence. IQ is widely accepted to be a valuable tool in assessing an individual's general intellectual performance, and has been found to predict a number of social, economic and personal outcomes. While these assertions are not challenged by the current thesis, it is contested that IQ refers to an essential innate trait (i.e. intelligence) that constitutes an individuals intellectual capacity. The current thesis criticises both traditional theories of intelligence and misconceptions regarding IQ, but also recognises the unparalleled utility and practicality of IQ tests in giving an approximation of *intellectual performance* for a given individual in a given assessment. Due to this fact, the current thesis will avoid use of the term “intelligence”, as this implies such a trait. Rather, the terms “intelligent behaviour” or “intellectual performance” will be preferred, as such terms implicate a collection of demonstrable skills or behaviours, rather than the mentalistic construct described in many mainstream conceptualisations of intelligence.

In the absence of a general consensus regarding how intelligence is defined, the current thesis will therefore employ the psychometric concept of IQ as a benchmark for intervention success as it represents a widely accepted approximation of intellectual performance. It is hoped that by assessing the effectiveness of relational training in improving IQ scores, the importance of relational responding to intellectual performance will be investigated. While a thorough investigation of what constitutes intelligence is beyond the scope of the current thesis, it is hoped that the study of relational responding skills and their contribution to intellectual performance may pave the way for new insights in intelligence research. However, due to the prevailing assertion that behaviour analysis has relatively little to contribute to the study of intelligence, a discussion of the evolution of the intelligence concept will follow in the subsequent section. It is hoped that this discussion will illuminate some of the current issues and misconceptions present in current intelligence research, as well as underlining the relevance of behaviour analysis to this area of study.

1.3 Early Approaches to the Study of Intelligence

While a cognitive psychological paradigm now appears to dominate the study of intelligence, much of what we currently understand about intelligence has stemmed from a distinctly behaviouristic perspective. Spencer's "The Principles of Psychology" (1890), was the first psychology textbook to use the term "intelligence" and has been identified as a precursor to the unitary theory of intelligence as *g*. However, Spencer's conceptualisation of intelligence as an adaptive skill set used to ensure survival in an environment would seem to implicate a set of behaviours, operated upon by a set of selectionist contingencies, rather than a collection of extant and fixed mental constructs (e.g., Jensen, 1998). Sir Francis Galton, one of the early pioneers of differential psychology who laid the foundations for the future of psychometric testing, began the objective measurement of intelligence by assessing performance on tasks he deemed relevant to intelligence (Galton, 1883). While the simple battery employed by Galton was later found to be ill-equipped to measure the complexity of human intelligence, it is important to note that intelligence testing began by assessing demonstrable behaviours (response times & sensory discrimination tasks) rather than mental constructs, even if the former are often viewed as mere proxies for the latter.

Of great relevance to the current debate between paradigms of human intelligence are the views of Alfred Binet, the psychologist who co-formulated the first practically useful measure of intelligence in 1905. Binet was one of the first researchers to suggest that intelligence may be better conceptualised as a measure of higher psychological processes, rather than simple sensory processes (Binet & Henri, 1895), an assertion which would help shape the testing process he would later outline alongside Simon (1905). Pertinent to the current study is Binet's acknowledgement that "intellectual development progressed at variable rates ...was malleable (within limits) rather than fixed" (Siegler, 1992, p183). Indeed, the formation of a SLEPE committee (Societe Libre pour l'Etude Psychologique de l'Enfant) in 1904, led by Binet as President, tasked itself with studying individual differences between school children, with a view of implementing special educational intervention for those on the lower end of the intellectual spectrum (Binet, 1904a; 1904b; Nicolas, Andrieu, Croziet, Sanitioso & Burman, 2013). Binet's motivations were markedly humanistic, reflected in the fact that the Binet-Simon scales were intended for pragmatic purposes. Its validity in identifying those in need of additional educational support was intended to supersede the test's ability to precisely measure a single faculty of intelligence (Gregory, 1994; Vial & Hugon, 1998).

The utility of Binet's diagnostic measures were soon adopted in the United States, with Terman and his Stanford colleagues adapting Binet's original scales to form the Stanford-Binet scales (Terman, 1916). The Stanford-Binet scales were the first to multiply the scores found in the Binet-Simon scales by 100 in order to remove fractions, leading to an intelligence quotient abbreviated as IQ. While Binet had not lived to see this revision of his scale, his collaborator Simon branded the concept of IQ as a "betrayal" of their original objectives (cited by Wolf, 1974, p.203). The subsequent use of IQ assessments in support of insensitive governmental policy is well documented (Black, 2003; Reddy, 2008; Ryan, 1997), a particularly offensive example being the sterilization laws sworn in 30 US States in order to breed out the prevalence of what would now be considered learning disabilities (Buck vs. Bell, 1927; Herrnstein & Murray, 1994; State Commission in Lunacy, State Hospitals, and Care, Custody, Apprehension, Commitment of Insane and Other Incompetent Persons, 1913).

The implementation of IQ testing at Ellis Island to assess the suitability of "feeble-minded" immigrants for entry to the United States (Goddard, 1914; Kraut, 1990), overseen by Henry Goddard, one of the early translators of the Binet-Simon scales, proves a stark contrast to the original objectives Simon was concerned about. The deportation of such immigrants based on poor performance on the Binet-Simon scales provides insight into how quickly the scale was used for the demarcation of "lost causes", or worse, potential menaces to society (Gelb, 1986) rather than those who require additional intellectual support. While Goddard's (1917) assertion that these feebleminded immigrants could possibly serve as "moron labourers" if "we are wise enough to train them properly" does not entirely preclude the potential for intervention, its condescension would not seem to ally itself to Binet and Simon's original motivation for positive change. In fact, Goddard (1912), praised the pioneering efforts of Alfred Seguin, who outlined effective pedagogical strategies for the intellectually disabled in his book "Idiocy: and its Treatment by the Physiological Method" (1866). Goddard went as far as to state that due to Seguin's influence, low levels of intellectual function could be improved and that there was a possibility that people with learning difficulties "could even be cured" (1912, p.119). It appears that the worst, rather than the best of the zeitgeist of those times was retained, as many intelligence theorists continued to vouch for the stable and deterministic conceptualisation of IQ while abandoning the possibility it could be improved.

As Nicolas et al., (2013) pointed out, the Binet-Simon scale and its revisions/adaptions soon colonised the territory of intelligence research, as performance on the test no longer *reflected an approximation* of intelligence, but *equated* to it. This was an important evolution because it represents the shift from a functional, pragmatic view to a structural and mentalistic view of intelligence. The use of such a battery of tests to measure a relevant collection of intellectual abilities, and its subsequent reduction to a single number (IQ) can be seen as a major influence in single factor theories of intelligence. While the often-cited "common factor" or "positive manifold" present among various intellectual abilities has been extrapolated to support a notion of "*g*", it is important to be aware that this is a statistical tool. Factor analytic studies would indicate that there may be an underlying factor which contributes to a large number of intelligent behaviours; it does not conclusively identify the presence of a single intelligence. Due to the significance of this process, the reification of a common factor "*g*" to refer to a theory of a unitary intelligence will now be considered.

1.4 The Reification of Intelligence and its Single Factor

Since Galton first devised the correlation coefficient in 1888, its utility in application to the battery of intelligence tests and subtests that were to be developed was clear. It was this statistical procedure that led to Charles Spearman's seminal breakthrough in 1904 using factor analysis to identify a common element that underlies performance on all intelligence tests. Spearman's proposal was that, due to the positive correlations between scores on intelligence tests and subtests, it follows that there is a single factor or general trait that is fundamental to all intellectual performance (Spearman, 1904). Spearman's (1904) conceptualization of a unitary factor named "*g*" represents the general capacity to infer and apply relationships drawn from experience. Regarding Spearman's breakthrough, Guildford (1936) posited that "no single event in the history of mental testing has proved to be of such momentous importance" (p.155). According to Herrnstein & Murray (1994), Spearman's theory thus "shaped both the development and much of the methodological controversy about mental tests ever since" (p. 2). Perhaps the most significant contribution of this theory is that it has led to a consensus among classicists in intelligence testing that *"g"* is not merely an artefact of statistical manipulation, but an expression of the general mental ability that underlies intelligence (Herrnstein & Murray, 1994).

Spearman's conceptualization of "*g*" was an essentialistic attempt to reduce the wide range of intellectual ability and individual differences into some formalisation of the concept of intelligence itself. In his 1981 book, "The Mismeasure of Man", Gould suggests this endeavour was a product of "physics envy", in trying to reduce intelligence to its basic constituent parts and simplifying laws in order the validate the scientific integrity of psychology as a whole. This pursuit was of vital importance to the future of mental testing, as Spearman pledged to abandon the study of any observable phenomena or behaviour, and instead aimed to "invent underlying something which by analogy to physics, has been called mental energy" (Spearman, 1927, p.89). Spearman hoped to identify a physical property of the brain that could account for intelligence, by finding the material basis of individual differences (Fancher, 1985). Spearman posited that this may involve a factor lying "deeper than consciousness" and may consist of "something of the nature of an 'energy' or 'power' which serves in common the whole cortex" (1923, p.5). In an ironic contradiction of the type of scientific integrity he wished to commandeer for his field, Spearman boldly argued that his theory was "impervious to disproof" (Gould, 1981). Spearman tasked physiology with the identification of this mental energy, but maintained that if this was not possible, the concept must be regarded as being "purely mental" (1927, p. 408).

The argument that Spearman fell victim to the concept of reification has been well-expressed (Gould, 1981; Schlinger, 2003), accusing Spearman of mistakenly inferring the physical existence of *g* from a mathematical abstraction. According to Schlinger (2003), Spearman's use of statistical analysis aped the appearance of scientific credulity, yet represented another ironic artefact of Spearman's desire to retain the scientific integrity of his research. Schlinger also pointed out that Spearman's error of reification gave rise to another logical error – that of circular reasoning. The only evidence Spearman could propose for the existence of *g*, were the positive correlations that he aimed to explain at the outset. The fact that *g* is merely a "statistical abstraction" (Gottfredson, 1998), appears to have been lost on many of Spearman's disciples that have propagated such circular reasoning. As Howe (1990a) states, the term "intelligence" lacks the logical grounds necessary for conceptual status, and therefore serves merely as a descriptive or labelling construct. The descriptive function that the term provides is implicated as a poor basis for its own explanation i.e. the positive inter-correlations between test items is intelligence, and intelligence is the positive inter-correlations witnessed between test items. In discussing Howe's views on the matter, Matthews, Zeidner & Roberts (2002) term this process "linguistic sleight-of-hand" (p.88). Staats and Staats (1963) provide a very simple example to illuminate this fallacy, outlining how it is not only erroneous, but also relatively useless. Imagine a school psychologist tells Jimmy's parents that their son is constantly fighting because he is a hostile/aggressive child. If the psychologist was asked how he could attribute this trait to Jimmy, he will say it is due to the fact Jimmy is always fighting. Staats and Staats declare that this label, hostile-aggressive, adds no new knowledge, "is circular and does not explain the behaviour" (p.16). Like intelligence, even if the psychologist was to construct a measure that would predict the child's fighting behaviour, this metric would not substantiate the assumption that there is an internal disposition or process which caused the behaviour.

As *g* is a statistical construct, it does not offer any useful description of what intelligence or intelligent behaviour is. Furthermore, as a result of this, the true meaning of *g* is readily open to interpretation and manipulation. *G* remains the same as any other factor identified via factor analysis: a useful tool for identifying co-variation across multivariate data sets (Schlinger, 2003). In general, the use of factor analysis has been heavily criticized (Bernstein & Teng, 1989; De Bruin, 2004; Gorsuch, 1997; Reise, 1999; Waller, Tellegen, McDonald & Lykken, 1996). Significantly in terms of its relevance to intelligence research, a recurrent issue with factor analysis arises from the labelling of a common factor after it has been identified. As Creasy (1959) points out, naming a particular factor can often lead to the misconception that the factor has some validity in measuring or referring to the label applied to it, i.e. the error of reification. In the case of *g* being used to refer to intelligence, this label could have been easily substituted for a number of appropriate labels. The common factor itself is independent from the label applied to it, as the naming process is purely subjective. Creasy (1959) goes further to state that a label may only have absolute meaning if the tests included in factor analysis are truly representative of the skill or concept the label refers to. If the test battery is inappropriate, ineffective or limited in its ability to accurately measure the label ascribed to the common factor, then the emergence of that factor cannot be deemed valid, or even of interest, in describing that label and vice-versa. Simply put, if a battery of intelligence tests is deficient in measuring what intelligent truly is, the common factor emerging from that battery cannot be defined as the value of intelligence. At the time of Spearman's apparent breakthrough, the measurement and conceptualisation of intelligence was still in its formative stages. Therefore, the proposal that this common factor had uncovered the true meaning or quality of intelligence can only be regarded as being premature at best.

Regarding its relevance to intelligence testing, it may be more an indication of how IQ test are designed, rather than indicative of a fundamental quality of those who take them (Schlinger, 2003). Crucially, from this perspective, Spearman's object of interest was not a "thing" to be measured, but instead was a statistical artefact of the behaviours of people taking the test. Ertel (2013) argues that those general factors may represent more than one source of variance. For example, a general factor derived from a questionnaire may represent a trait disposition in addition to acquiescence or social pressure. In the same vein, a general factor of intelligence tests may be based on intelligence alongside other variables such as ambition, concentration or determination (McCaffrey, Duff & Westervelt, 2000; McCaffrey & Westerveld, 1995; Reeve & Lam, 2007; Thorndike, 1949). In effect, individual differences in motivation or focus may contribute to positive correlations among battery subtests.

A common alternative to Spearman's explanation of the shared variance argues that the common factor among IQ subtests represents processing speed, not intelligence (Jensen, 1998). Processing Speed refers to the rate at which an individual can receive and interpret information and complete basic cognitive functions (Hale, 2000). Similar to findings on intelligence, a number of studies have found that some people tend to be particularly quick on most subtests, while others are particularly slow (Hale & Jansen, 1994; Myerson, Hale, Zheng, Jenkins & Widaman, 2003; Vernon, 1983; Zheng, Myerson & Hale, 2000). Salthouse (1996) outlined two reasons why processing speed may be of vital importance to intellectual functioning. Firstly, in situations in which an individual cannot control the rate at which information is presented, a slower processing speed may lead to information being missed. Secondly, a rapid processing speed allows the individual to perform multiple tasks more efficiently, as less time is required for each. Working memory capacity has also been proposed as an explanation of this common factor (Engle, Tuholski, Laughlin & Conway, 1999). As evinced by such investigations, the validity of labelling Spearman's positive manifold as evidence for a single factor of "intelligence" is decidedly open to debate.

Spearman's insistence on identifying where in the brain that intelligence resides was closely bound to his assertion that intellectual ability was genetically inherited

(Fancher, 1985). Beginning with Galton's analysis of preeminent families, studies into the heritability of intelligence generally estimate a value between .3 and .8 (Jensen, 1969; Plomin, 1990). While such figures are supported by the gold standard twin and adoption studies (Petrill et al. 1998; Pike, Reiss, Hetherington & Plomin, 1996; Plomin, 1997; Plomin & DeFries, 1998), Schlinger (2003) asserts that many of the core assumptions underlying such studies are inherently flawed. For example, such research is heavily reliant on the assumption that fraternal twins are exposed to identical environmental factors. Beckwith (1999) and Joseph (1998) have demonstrated that this is not the case. In addition, Harvard astrophysicist David Layzer (1972) questions whether the variation of genetic and nongenetic factors contribute additively and independently to intelligence – and in the case they do, whether the heritability of IQ can be thus computed from the extant data. Layzer also argued that all estimates of the heritability of IQ are "unscientific and indeed meaningless" due to the assumption of equal environments, the assumption that intelligence represents a metric character (like height or weight) and finally, the fact that IQ tests do not do what they aim to do. Furthermore, Layzer specifically states that due to the standardisation of IQ scores as a representation of an individual's intellectual performance relative to his/her peers, IQ tests do not study whatever intelligence may be, but rather a relative ranking among peers.

As has been discussed, the reification of *g*, and its distortion from a common factor between IQ subtests into an explanation of the quality of intelligence itself has come under heavy criticism. Most crucially, Spearman's factor analysis cannot provide valid support for mentalistic conceptualisations of intelligence. Schlinger (1998) argues that the inference of unobserved constructs in other sciences is possible, given their already established base of experimentally derived relationships. This license cannot be extended to intelligence research, as it lacks such a grounding that could support the inference of an internal construct known as intelligence. Schlinger also criticized the reliance of *g* theorists on neuroimaging correlational data, which can only inform us tentatively on the biological underpinnings of intelligent behaviours, rather than offer any descriptive validity. While Spearman previously rejected such a criticism in his final book 'Human Ability', co-authored with LL. Wynn Jones (posthumously published in 1950), Spearman later admitted that his attempts to reify statistical factors and identify mental energy may have been a product of youthful inexperience and exuberance. In summary, Spearman's unitary factor theory has exerted a pronounced influence on the study of intelligence, despite serious concerns regarding its scientific validity and utility. When closely scrutinised, Spearman's seminal breakthrough comes loaded with pressing methodological and conceptual caveats. These caveats would appear to call for a re-examination, and possible reorientation of psychology's current paradigm of intelligence.

1.5 The Malleability of Intelligence

Psychologists outside the sphere of behaviour analysis have, for the most part, aimed to identify the single factor known as intelligence. Those within it have long posited that the term intelligence merely refers to the collection of behaviours that cannot be separated from their context (Schlinger, 2003). While Spearman's two factor theory of intelligence (1927) views "*g*" as an invariant trait, the behaviour analyst regards intelligence as a collection of skills that are inherently malleable, and whose fluency are the subject of mainstream IQ measures (Cassidy et al., 2011; Cassidy, Roche & O'Hora, 2010; Schlinger, 1993). Upon closer inspection of the history of psychometric testing, many of the seminal theorists in the field of intelligence have in fact highlighted the possibility that it may be open to modification. As was previously discussed, one of the forefathers of modern intelligence testing, Alfred Binet, specifically designed the earliest IQ in order to facilitate intervention. Gardner, who formulated Multiple Intelligence theory stressed that "intelligence can be learned and improved throughout life" (1983, p. 41). Importantly, unlike Spearman, Gardner warns against the fallacy of reification by stressing that he does not suggest that his multiple intelligences are entities that are materially evident, but are instead constructs that may offer practical utility. Vygotsky (1978) proposed the theory of the "zone of proximal development", arguing that every child had a potential performance level that could be reached with appropriate help. Staats and Staats (1963) argued that intelligence is, in general, learned behaviour consisting of a repertoire of skills/behaviours such as reasoning, arithmetic, discrimination, communication and so on. An individual's degree of sophistication regarding this repertoire would therefore represent his intellectual functioning at any given time. Most importantly, the conceptualisation of intelligence as a number of skills does not preclude the possibility of improving those skills, and thus, increasing intelligence. Staats and Staats argue that intelligence is not an inherited personal quality, but instead represents the stage or position in the cumulative-hierarchical learning process that an individual's life conditions and experiences have brought him/her to. In this conceptualisation, intelligence is regarded as the current level an individual has reached in their learning process. This progress is a cumulative process, which can be readily witnessed in how relative easily a child can learn the 15th, 16th or 17th letters of the alphabet compared to learning the 1st or 2nd (Staats, 1989). This process produces "better" or more accomplished learners, individuals who would be described as more intelligent by the general definition. In line with such assertions, numerous interventions have displayed tentative efficacy in improving intellectual performance, which will be discussed in the following section.

1.5.1 Working Memory Training

Sternberg's (2008) concession that "fluid intelligence is trainable to a significant and meaningful degree" was based on the pioneering work of Susanne Jaeggi and colleagues on working memory training (e.g. Buschkuehl et al., 2008; Jaeggi, Buschkuehl, Jonides & Perrig, 2008). Fluid intelligence (sometimes referred to a G*f*) is defined as an individual's ability to understand complex relationships and solve new problems independently of previously acquired knowledge (Jaeggi et al., 2008; Martinez, 2000). Gottfredson (2003) describes fluid intelligence as a measure of "mental horsepower", implicated in solving cognitive problems on the spot. Working memory refers to a multi-component system that holds and manipulates information in short-term memory (Cowan, 2009). Jaeggi et al. (2008) found that a working memory training program can lead to increases in fluid intelligence. In this study, four training groups were exposed to an *n*-back working memory training regimen for 8, 12, 17 or 19 sessions. In the training task, participants were shown two series of stimuli simultaneously for 3 seconds for each stage. One series consisted of single letters, whereas the other series consisted of spatial markers appearing on different parts of the computer screen. Participants were required to answer whether the current stimulus matched the stimulus shown a variable amount of stages ago. The letter *n* was used to denote the number of stages that had passed. For example, a participant may be asked whether the stimulus currently displayed onscreen matched the stimulus that was displayed 3, 4, 5 exposures ago etc. As the participant emitted correct responses, the value for *n* would increase incrementally. If a participant answered incorrectly, *n* would decrease. As the task adjusted its own difficulty to match the participant's performance, the task would remain challenging. There was also a control group that was administered the measures for G*f.* The results of this study received widespread attention, most notably Sternberg's aforementioned approval, and suggested that such a training program can lead to genuine benefit for intelligence. The authors themselves dubbed their findings "a landmark result" (p.4).

That being said, many of Jaeggi, Buschkuehl, Jonides & Perrig's (2008) much heralded improvements in G*f* come with major caveats. First and foremost, their results indicated very modest IQ gains. Furthermore, there are a number of serious methodological flaws with the study, including issues over insufficient time allocated to complete outcome measures, differing test batteries for experimental and control groups as well as dubious concerns over the generalisability of findings to intelligence (Moody, 2009). One such flaw with the study is the fact that the four groups did not receive the same assessment of working memory. The group which received the least amount of training (8 sessions) were administered Ravens Advanced Progressive Matrices (RAPM; Raven, 1990). RAPM is a widely used assessment tool, which requires participants to select one of five options to complete a geometric pattern. Improvements in this group were negligible. The other three groups were assessed using the Bochumer Matrices Test (BOMAT; Hossiep, Turck & Hasella, 1999). The BOMAT is an assessment of working memory that consists of 29 visual analogies. The discrepancy in the testing batteries administered to participants would therefore raise questions over the validity of Jaeggi et al.'s between-group comparisons.

To compound the issues of differing test batteries, Jaeggi et al. (2008) drastically reduced the recommended time allocated to complete the BOMAT from 45 minutes to just 10. Due to this incredibly short time frame, it was made near impossible for participants to reach the latter stages of this 29-item assessment, thereby lowering their scores irrespective of true ability. The question of whether a participant could correctly respond to the more difficult items was entirely precluded due to this unreasonable time restraint. Furthermore, Ravens Matrices are presented in a 3x3 format, while the BOMAT presents a 5x3 format. The finding that the "landmark" improvements witnessed in Jaeggi et al.'s (2008) are isolated for the three groups that completed an inappropriately administered, lesser-established measure of working memory rather that those who were administered a well established measure would appear to raise some concerns over the validity of such findings. Given the study's failure to assess intelligence using a full-scale IQ test, the generalisability of the findings to actual intellectual performance must also be questioned. In a critique of the study, Moody (2009) argues that the data presented "is not sufficient to support the authors' conclusion of any increase in their subjects' fluid intelligence" (p.1).

While criticisms of Jaeggi et al.'s (2008) original and subsequent findings may be justified, their work has added momentum to idea that intelligence may be indeed be malleable. Indeed, there is a burgeoning repertoire of research suggesting that intelligence may not be as static as once believed (Dickens & Flynn, 2001; Nisbet et al., 2012; Stankov, 1986) and may be increased via intervention (Cassidy et al., 2011; 2016; Dixon et al., 2014; Jaeggi et al., 2008; Moran et al., 2010; Stephenson & Halpern 2013).

1.5.2 Cognitive Training & Neuroplasticity

Recent neuroimaging studies have implicated the efficacy of training programs which focus on reasoning skills in altering white matter microstructure and neuroconnectivity. Mackey, Whitaker & Bunge (2012) recruited 23 participants who had enrolled in a preparation course for the Law School Admission Test (LSAT) along with 22 age and IQ matched controls that intended on sitting the LSAT in future. As the LSAT relies heavily on reasoning ability, the preparation course focused on improving reasoning skills. The LSAT preparation program involved 100 hours of training on Logic Games, Logical Reasoning and Reading Comprehension. Diffusion tensor imaging (DTI) was collected at two sessions three months apart. DTI data indicated decreased radial diffusivity in white matter connecting frontal cortices and in mean diffusivity within frontal and parietal lobe white matter in trained participants but not in controls. Participants who showed the greater improvement in LSAT performance also displayed greater decrease in mean diffusivity in the right internal capsule.

Mackey et al. (2012) propose that these findings provide evidence of experience-dependent white matter changes. To follow up on these findings, Mackey, Miller Singley & Bunge (2013) investigated whether intensive reasoning training, and the resultant shift in cognition, could affect the strength of “coupling” between regions that typically considered functionally related. As patterns of correlated activity among brain regions represent functionally relevant networks, Mackey et al. (2013) hypothesized that these networks must be based on a prior history of co-activation. Replicating the general experimental sequence of Mackey et al. (2012), the authors used FMRI to analyse activity in the fronto-parietal network, brain regions that have been previously implicated in reasoning processes (Krawczyk, McClelland & Donovan, 2010; Krawczyk, McClelland, Donovan, Tillman & Maguire, 2010). FMRI data supported the researchers’ hypothesis, demonstrating that training participants displayed strengthened fronto-parietal and parietal-striatal connections. The findings of Mackey et al. (2013) further demonstrated the “accessibility” of neuroplasticity for intensive training programs that target higher-level cognitive processes, demonstrating accruements in intellectual function at the neural level.

1.6 Unification of Approaches to the Study of Intelligence

There is a movement within cognitive psychology that has progressively leaned towards more behavioural accounts of intelligence, by focusing on observable skills and abilities, rather than mental constructs (Howe, 1989). As Anderson (2004) points out, progress in the study of intelligence may hinge on the synthesis of the study of individual differences and the study of cognition. Indeed, Cronbach (1957) led early calls of the unification of the two approaches, experimental and correlational, that dominated psychological research. Anderson (2004) argues that the current intra-disciplinary quandary present in the study of intelligence is resultant from both sides' failure to heed Cronbach's advice. While complex "cognitive" behaviours are mistakenly viewed to be beyond the reach of behaviourism, there have been numerous behaviouristic accounts of higher-level, complex behaviours that constitute much of the cognitive psychologist's research agenda. Several researchers (e.g. Dale, 2005; Galizio, Stewart & Pilgrim, 2001) have proposed that behaviour analysts and cognitive researchers share numerous areas of interest. For example, there have been comprehensive behaviourist analyses of reasoning (Maltzman, 1955), personality and psychotherapy (Dollard & Miller, 1950), problem solving (Judson, Cofer & Gelfand, 1956), concepts (Hull, 1920) and abstraction and images (Skinner, 1953). Furthermore, Schlinger (2003) and Baars (1986) point out that any study into the structure and processes of memory, a topic of defined focus to the cognitivist, relies almost exclusively on inferences gleaned from overt behaviour.

Due to the failure of many such behaviouristic accounts of complex cognitive processes to establish continuing streams of empirical or theoretical research, their influence has been diminished (Staats, 1989). However, the considerable contribution of behaviour analysis to the study of language and higher cognition, provided by

Relational Frame Theory (Hayes et al., 2001), provides the empirical basis of this study and will be later discussed in detail. Relational Frame theory posits that cognition is not a mental event, but a behavioural event, and as such, there is no justification that "a psychology of cognition cannot be a behavioural psychology" (Hayes et al., 2001, p.144).

More recently, there have been attempts to unify cognitive and functional/behavioural approaches in an attempt to construct a more comprehensive understanding of cognitive processes (De Houwer, 2011; De Houwer, Barnes-Holmes & Moors, 2013; De Houwer, Gawronski & Barnes-Holmes, 2013). Advocates of such an approach argue that the two main psychological approaches, cognitive and functional, are not adversarial but complimentary in accounting for cognitive behaviour. Such attempts to coalesce the study of intelligence reflect a more general worry over the "fragmentation" of psychology, as many researchers have called for a unification of psychology (Anastasi, 1990; Bevan, 1982, 1991; Bevan & Kessel, 1994; Fowler, 1990; Kimble, 1994; MacIntyre, 1985b; Maher, 1985; Royce, 1970; Rychlak 1988; Sternberg & Grigorenko, 2001). This shift of attention towards demonstrable behaviours offers the opportunity to improve not only our understanding of what intelligent behaviours are, but also our ability to describe what the term "intelligence" means. As mentioned earlier, there is considerable overlap within the remits of cognitive psychology and behaviour analysis. According to O'Hora, Pelaez, Barnes-Holmes, Rae, Robinson & Chaudhary (2008) this overlap is particularly salient in the study of relational responding, which serves as the focus of the current investigation.

As have been discussed, a great number of misconceptions and misunderstandings regarding behaviour analysis' relevance to intelligence continue to prevail in mainstream psychology. Such is the prevalence of these notions, the current conception of intelligence appears to sustain symbiotically with the misconception of the stability of IQ, as well as its invulnerability to intervention. However, there are a number of different streams of research that appear to provide evidence supporting the potential efficacy of training programs, based on behaviourist principles, which isolate and improve the underlying skills determining intellectual function. Pioneering work in the field of Applied Behaviour Analysis has facilitated many such research endeavours. Due to the seminal impact of ground-breaking interventions in this field, the development of applied behaviour analytic approaches will now be discussed.

1.7 Applied Behaviour Analysis

Much of the empirical validation for malleability of intelligence stems from behavioural interventions in the context of developmental disability, and in particular, autism. A wide range of applied behavioural analysis (ABA) interventions have been developed, with most sharing a strong emphasis on Skinnerian learning theory (Howlin, Magiati & Charman, 2009). Methodologies such as Pivotal Response Training (Schreibmann & Koegel, 2005), Discrete Trial Training (Maurice, Green & Luce, 1996) and Verbal Behaviour (Barbera & Rasmussen, 2007) have shown great utility in implementing positive change in autistic samples, and are at the forefront of many intervention programs currently in use (Howlin et al., 2009). ABA is now considered to be the treatment of choice for autism-spectrum disorders, receiving official commendation from the US Surgeon General (U.S. Department of Health and Human Services, 1999), as well as the state government of California (California Department of Education, 1997).

In one of the pioneering investigations of its kind, Lovaas (1987) implemented an early intensive behavioural intervention (EIBI), consisting of a systematic reinforcement procedure, in order to shape adaptive behaviours in an autistic sample of young children. The preschool children who comprised Lovaas' sample were exposed to at least 40 hours of one-to-one therapy per week, over the course of two years. The results of this controversial study appeared to indicate autism "recovery", as well as unprecedented IQ increases of 30 points, following an intensive ABA intervention. Almost half (47%) of the participants were described as presenting within the normal range of intellectual and educational functioning following the intervention. A follow-up study also indicated that this improved intellectual function was maintained through to early adolescence (McEachin, Smith & Lovaas, 1993). Baer (1993) described the study as "a triumph of behavioural science and behaviourally scientific clinical application" (p.373). Results such as these paved for the way for further studies advocating the potential increases in intelligence measures that applied behaviour analysis could provide.

Upon closer inspection of Lovaas' data, Smith, Groen & Wynn, (2000) points out that the IQ rises were far from uniform across the sample. Half of the group, those who acquired normal functioning, had a mean increase of 37 IQ points, while the remainder reported a mean increase of just 3 IQ points. Furthermore, Shea (2004) states that only one child in Lovaas's sample displayed all scores in the average category. Replications of Lovaas' (1987) results have had mixed success rates (Eikeseth et al., 2002; Jacobson, Mulick & Green, 1998; Sallows & Graupner, 2005; Sheinkopf & Siegel, 1998; Smith et al., 2000). While the reliability and validity of Lovaas' findings have been called into question, this criticism did not halt the behaviourist investigation into bringing intellectual behaviours and skills under operant control, as well as designing behavioural interventions for intellectual performance.

Despite these potential inconsistencies, a number of ABA and EIBI techniques have displayed considerable and robust utility in treating various symptoms of autism (Harris & Handleman, 2000; Myers & Johnson, 2007; Reichow & Wolery, 2009; Yamamoto & Mochizuki, 1988). Howlin et al. (2009) carried out a systematic review of EIBIs for children with autism, and included several studies which aimed to assess the approach's efficacy with regard to increasing IQ scores. Numerous studies (Cohen et al. 2006; Eikeseth et al., 2002, Lovaas, 1987, Remington et al. 2007) reported that the largest IQ gains occurred in the period between baseline and first follow-up, with IQ rises gradually decreasing following multiple follow-up assessments. Such a finding, Howlin et al. (2009) conclude, would seem to suggest that the greatest impact of EIBI is found in the primary stages of the intervention. This review also reported a mean increase of 18.3 IQ points for EIBI studies which included intelligence quotients as an outcome measure, with a mean increase of 5.4 points in comparison groups.

In another meta-review of EIBI, Reichow & Wolery (2009) analysed the results of studies based on Lovaas' program. All studies implemented long-term (over one year) and intensive (18.7 to 40 hours per week) intervention for children predominantly presenting with autism and autism spectrum disorders. While Reichow & Wolery report that the mean effect size for EIBI on IQ was large (0.69), at least one child in every study did not show significant improvement. Reichow & Wolery conclude that while EIBI may be an effective treatment in general, it may not lead to improvements in all children. This caveat can also be concluded upon reflection of Howlin et al.'s (2001) synthesis, as well as Lovaas' (1987) seminal breakthrough. Nevertheless, despite a lack of complete uniformity in IQ rises, early behavioural interventions have repeatedly and reliably shown a capacity to catalyse positive IQ changes in children with autism.

In addition to ABA interventions proposing the malleability of intelligence, Staats and Burns (1981) have produced a number of studies which empirically tested the hypothesis that intelligence is nothing but learned behaviour, and therefore, is inherently subject to change. The first of these studies tested the hypothesis that intelligence tests measure several independent intelligent factors. Two subtests of the WPPSI (Wechsler, 1967), Geometric Design and Mazes, were employed as they are suggested to represent two different intelligences (abstract-conceptual-mathematical & problem-solving respectively). A sample of 22 preschool children was recruited to be divided equally into a training group and a control group. The training involved 6 hours of teaching children to read and write the letters of the alphabet, as the researchers viewed these skills as central to performance on the WPPSI. Results found that the intervention group reported increases of 14 IQ points. In a second study, it was found that training preschool children simple discrimination skills led to increased performance on intelligence tests. Training labelling language repertoires (e.g. class to member hierarchies) was also found to positivity affect performance on various intelligence tests (Staats & Burns, 1981). While it may be contested that such findings may reflect a "training-to-the-test" effect, the fact that such skills can be trained may be as much an indictment of the validity of intelligence testing as a support for any sort of intelligence training program.

While behavioural interventions have a well-established role in improving functioning in samples with learning and developmental difficulties, there has been a shift in focus towards understanding the underlying processes that underpin a wide range of intellectual abilities, in both normally developing and disabled samples. We will first outline these developments and then go on to show how they have led to intervention and new ways of conducting intellectual ability assessments from a functional point of view. The integration of these interventions into basic and traditional behavioural methodologies will also be outlined. These new conceptualisations address some of the most complex forms of behaviour that psychologists assess and, in so doing, challenge the stereotype that behaviour analysis is only suited for animal research or for analysing the behaviour of children with learning difficulties.

1.8 Relational Frame Theory

The impetus for the recent surge of interest in behavioural interventions to increase general intellectual performance in normal populations comes from advances in a behaviour-analytic theory of language and cognition, known as Relational Frame Theory (RFT; Hayes, et al., 2001; see also Dymond & Roche, 2013). Relational Frame Theory (Hayes, 1991) is an account of human language and cognition which provides a bottom-up explanation of the increasingly complex and interwoven interactions between behaviour and environment that can account for many of the higher order processes, such as language and intelligence, that may previously have been beyond the reach of the behaviour analyst (Stewart & Roche, 2013).

RFT can be viewed as a "reorientation" of the major behaviour analytic conceptualisation of language that preceded it – outlined in Skinner's Verbal Behaviour (1957), by focusing its attention upon the listener, rather than the speaker, in verbal interaction (Cassidy et al., 2010). The major criticism of Skinners account of language (and by extension, behaviourism's) was that it placed the burden of learning on an expansive history of direct reinforcement, necessitating the explicit training of countless numbers of word-object relations (and vice-versa) in order to account for the rich vocabulary acquired by young children in such a short period of time (O'Toole et al., 2009). Such a stringent adherence to direct reinforcement may also lead to stereotypy, as reinforcement would lead to decreased behavioural variability (Vogel & Annau, 1973). This result would strongly contrast with the diversity and heterogeneity of language. Indeed, such a behavioural rigidity has been cited as a common criticism of ABA (Lovaas & Wright, 2006). Behavioural rigidity has also been associated with reduced levels of intellectual performance and problem solving (Fattu, Mech & Kapos, 1954; Schaie, 1958). Furthermore, once the variety of responses occasioned by the synonymy of common words, as well as the numerous context in which they may be presented, is considered, the feasibility of direct training appears less realistic still.

While Skinner's account was not entirely based upon such a history (he also refers to the "generic extension" of tacts and the creation of mands on the "analogy" of others, 1957), the bulk of his work would appear to advocate the prevalence of direct, rather than "derived" learning of language. RFT draws upon a repertoire of research that would highlight the inherently derived nature of human language (Sidman, 1994; Hayes et al., 2001), and undermine the centrality of direct stimulus control in language acquisition. In its place, it was proposed that derived relational responding was required to understand the listener's understanding in verbal interaction. Derived relational responding refers to the predictable untrained responses that are facilitated by a framework of relationships between known and novel stimuli regulated by arbitrary contextual cues (Cassidy et al., 2010). Hayes et al. (2001) proposed that the capacity to identify relations between and among stimuli is the fundamental aspect of human cognition.

A number of species have demonstrated a capacity for non-arbitrary relational responding (i.e. the ability to identify relationships based on the formal properties of the stimuli). Such animals include monkeys (Harmon, Strong & Pasak, 1982), pigeons

(Wright & Delius, 1994), fish (Perkins, 1931) and bees (Giurfa, Zhang, Jenett, Menzel & Srinivasan, 2001). However, due to the contingencies provided by the complexity of our socio-verbal environment, humans show a unique ability for arbitrarily applicable relational responding (AARR; Barnes-Holmes & Barnes-Holmes, 2000). Such responding is not reliant on the formal physical relations between stimuli, but instead on abstraction of relations. This higher-level relational responding removes the centrality of the physical stimulus, and allows the application of such relations to be applied across a wide array of stimuli and circumstances (Stewart & McElwee, 2009). For example, if a child is taught verbally (i.e. without visual stimuli) the simple relation that "House A is bigger than House B, and House B is bigger than House C", the child will typically show the ability to then further derive relations between stimuli that have not been explicitly trained. In this instance, the child will be able to derive that House A is bigger than House C, and House C is smaller than House A, even though the child has not been given any specific information regarding the actual size of any of the houses. It is in this way that humans appear to form networks of relations that allow the derivation of further relations that have not been explicitly defined.

RFT would suggest that the phrase "bigger than" in this case represents a contextual cue that has previously been established in a child's learning history as controlling a particular pattern of generalised relational responding. Once this relation is established, that response pattern can be applied arbitrarily to a collection of stimuli regardless of their formal properties, which in turn helps to establish a coherent network between those stimuli. Of key importance is the definition of AARR as a generalised, functionally defined operant (Hayes & Barnes-Holmes, 1997). Such operant classes are defined in terms of their functional context, as opposed to their topography – which, in a sense, liberates such responding from relying on formal properties of stimuli, allowing a much wider range of applications. While members of these classes may have topographical features in common, in many cases, this is not the case, such as generalised imitation (Baer, Peterson & Sherman, 1967) and novel responding (Pryor, Haag, & O'Reilly, 1969), and are thus grouped together due to a functional similarity.

While Skinner's account heavily emphasised a "content-delivery" interaction between speaker and listener, RFT focuses on learning rather than content, as evinced by the functional definition of AARR. The derivation of an increasingly complex and interwoven set of relations can provide a behavioural account of language acquisition that according to some (e.g. Chomsky, 1959), was previously impossible. A simple example comes in the form of a typical interaction between adults and young children, in which the adult teaches the complex arbitrary relationship between a physical object and its lexical and verbal association. By showing an object (e.g. apple) and pairing it with the verbal utterance of the word "apple", the adult is training the child to establish a relationship of symmetry (sameness) between the physical object and the spoken word. This symmetry is reinforced by then holding up the apple and asking the child "what is this?" or "where is the apple?". This explicitly trains the symmetrical relation in both directions i.e. object to word & word to object. The complexity of this relational frame is made more complex once the spoken word is paired with the written word, and vice versa. Another example of derived relational responding can be commonly perceived when the child then matches the written word to the matching object. The symmetrical responding between words and objects is a generalised response class formed by a history of reinforcement across multiple exemplars, and once established, the child will then show the ability to derive untrained symmetrical relation regardless of the physical properties of the stimuli involved. This example of generalised contextually controlled relational responding is the first children establish and most importantly, underpins the development of language (Stewart & Roche, 2013).

The relation of "sameness" outlined above is just one of a number of "relational frames" that underpin linguistic and cognitive performance. Other relational frames include such relations as comparison ("A is bigger than B)", opposition ("big is opposite to small"), distinction ("this is not the same as that"), hierarchy ("an apple is a type of fruit"), analogy ("foot is to sock, as hand is to glove"), deixis ("I am here and you are there") and temporality ("morning comes before afternoon"). It is likely that the "language explosion" coincides with the child developing a reasonably coherent network of relations for the first time – allowing him/her to derive a multitude of novel relations, causing a sudden expansion of vocabulary (Stewart & Roche, 2013). Relational Frame Theory has demonstrated that people respond in accordance with these frames and that such frames can be established and/or strengthened (Berens & Hayes, 2007; Carpentier, Smeets & Barnes-Holmes, 2003; Roche & Barnes, 1997).

1.8.1 Relational Frame Theory & Intelligence

Relational Frame Theory invokes generalised operant classes with great utility in explaining a number of higher-level complex behaviours (Barnes, 1994; Barnes, Browne, Smeets & Roche, 1995; Catania, 1996, 1998; Hayes, 1991; Hayes & Barnes, 1997; Healy, Barnes & Smeets, 1998; Horne & Lowe, 1996; Lowekron, 1998; Roche & Barnes, 1997; Smeets, Barnes, & Roche, 1997; Roche, Barnes- Holmes, Smeets, Barnes-Holmes, & McGeady, 2000). Therefore, RFT posits that it may be of considerable practical utility to train a core set of these higher-order operant response classes (i.e. generic cognitive skills) with the view of subsequently applying these response classes arbitrarily to stimulus content (O'Toole et al. 2009). Thus, training a collection of versatile and overarching response classes can bypass the issue of direct training, as the responder will show greater degrees of flexibility in a variety of contexts. O'Toole at al. (2009) argued that by focusing on derived performances and cognitive processes, the issue of direct training and reinforcement can be avoided. As it can be considered a generalised operant class, derived relational responding is by definition, flexible (Hayes, 1994), facilitating the possibility of modifying such responding by means of manipulating the environmental contingencies that act upon it. Contextual cues have been consistently found to alter patterns of derived relational responding (Dymond & Barnes, 1994; Roche et al., 2000; Sidman, 1971). Consequential control has also been demonstrated to change such responses by manipulating reinforcement patterns (Wilson & Hayes, 1996).

Numerous studies have suggested that this flexibility may be of fundamental importance to higher-level processes, such as creativity, problem solving and intelligence in general (Barnes, Hegarty & Smeets, 1997; Hayes, 1994; Healy, Barnes-Holmes & Smeets, 2000). Cognitive flexibility is regarded by many mainstream theories as one of the most important features of intelligence (Cattell, 1971; Kyllonen, Lohman & Wolta, 1984). To complement this finding, its antithesis, cognitive rigidity, has been found to exert a deleterious effect on intelligence, as well as underlying many psychological disorders (Lovecky, 2004; Turner, 1999; Wulfert, Greenway, Frakas, Hayes & Dougher, 1994). It would therefore follow that any intervention that displays a potential to increase cognitive flexibility may have a positive effect on intelligence.

While Spearman's influential conceptualisation of *g* may appear oppositional to such a viewpoint, a striking parallel can be drawn between the centrality of relational ability to intelligence, propagated by RFT, and Spearman's operational definition of *g*. Spearman (1904, 1927) in fact concluded that *g* is most strongly reflected in tasks that call for the "eduction of relations and correlates". The eduction of relations and correlates, the second and third of Spearman's noegenetic laws, refer to tasks that involve grasping relationships, induction & deductive reasoning, inferring rules, generalising, recognising similarity and difference, and decontextualising a problem (Jensen, 1998), tasks which would also implicate the RFT theorist's concept of relational framing. Therefore, despite disagreement over the malleability of intelligence, there is considerable theoretical overlap between Spearman's influential work and the RFT perspective on the skills that underlie intelligence.

Relational Frame Theory studies have provided robust empirical support for the argument that relational ability is correlated with overall intellectual functioning. O'Hora et al. (2005) divided 26 monolingual and 46 bilingual college students into two groups, depending on their performance on a complex relational task. The relational task involved training and acquiring mastery on three levels of responding complexity – before/after, same/different and finally, a test of instructional control involving complex networks of same and before/after relations. It was found that performance on the relational task predicted subsequent performance on two of the three subtests of the WAIS-III included in the analysis (Vocabulary & Arithmetic). In a further study (O'Hora et al., 2008), 81 undergraduate students were asked to complete a temporal relations task, followed by the full battery of the WAIS-III. The temporal relations task required the participant to learn the temporal relational function ("before" and "after") of two abstract symbols ("()()" and "::::") within 12 blocks of 16 trials. Each trial consisted of the presentation of a "statement" comprised of one of the abstract symbols between two simple geometric shapes (e.g. square ()() circle) at the bottom of a computer screen. Participants were then presented with the same geometric shapes at the top of the screen and two statements at the bottom of the screen, which stated a temporal relation between them (e.g. “square before circle” or “square after circle”). In order to complete the trial, participants would be required to choose which statement matched the initial statement. Those who passed the temporal relations task were found to have significantly higher Full Scale and Verbal IQs. Completion of the task was also associated with significantly higher scores for two of the four WAIS subscales, Verbal Comprehension and Perceptual Organisation. This association was not found for the Working Memory or Processing Speed subscales. The finding of a close relationship between relational responding and intellectual performance, as assessed by a traditional IQ measures, would therefore imply the relevance of relational ability to intelligence.

These results are complimented by a more recent study by Gore, Barnes-Holmes & Murphy (2010) who identified strong correlation between deictic relational responding and scores for the Full Scale, Verbal and Performance IQ scales of the WASI. The sample used in the study consisted of 24 adults presenting with mild to moderate intellectual disabilities. The group was administered a simplified 34-item version of RFT Perspective Taking Protocol (Barnes-Holmes et al. 2004) which involved the relational frames of I/You, Here/There and Then/Now. Each item consisted of responding correctly to two questions, e.g. “where am I sitting?” and “where are you sitting?” Each relational task involved three levels of complexity: single relation, reversed relation and double reversed relation. Single relation trials asked questions such as “I have a red brick and you have a green brick. What colour brick do you have? What colour brick do I have?” The reversed relation trial consisted of similar propositions, but required the participants to answer as if he/she had switch roles with the experimenter. Finally, the double reversed trials asked the participants to not only switch roles with the experimenter but to switch the meanings of another of the relations, i.e. “if I was you and you were me, and here was there and there was here.”

As well as the correlations with IQ mentioned above, Gore et al. (2010) reported a strong correlation between verbal mental age and performance on the perspective-taking task. The results reported would appear to replicate previous findings regarding the connection between complexity and flexibility of relational responding and intellectual functioning (Andrews & Halford, 1998; Cattell, 1971; Gentner & Loewenstein, 2002), as well as the ability to predict IQ given levels of relational ability (O'Hora et al., 2005; O'Hora et al., 2008; O'Toole & Barnes-Holmes, 2009). The strong connection between relational ability and traditional measures of intelligence would suggest that it may be possible that AARR fluency underlies intelligence levels as assessed by standardised tests. If this is so, it would follow that it should be possible to increase intelligence scores through AARR interventions. Before research designed to address this precise question is outlined, it may be helpful to expand upon how RFT might conceptualise the main components of intelligence as assessed by standardised tests.

Various forms of relational responding have been shown to be of fundamental importance to a wide range of skills and behaviours that are constituent of intellectual functioning (Stewart, Tarbox, Roche & O'Hora, 2013). Furthermore, there is now a burgeoning research repertoire that links specific relational frames to specific areas of intellectual performance.

Coordination/sameness.

Perhaps the most clearly evident of the links before relational responding and intelligence can be witnessed for the relational frame of coordination/sameness. As evinced by the earlier example of symmetrical word-object relations, the relational frame of coordination serves as the basis for linguistic reference (Stewart et al., 2013). Therefore, levels of sophistication in responding to, and deriving, frames of coordination are of vital importance to developing a depth and breadth of vocabulary. Such is the contribution of verbal ability to levels of intelligence, vocabulary has been well established as a predictor, as well as being predicted by, general intellectual function (Smith et al., 2005; Marchman & Fernald, 2008; Vetterli & Furedy, 1997). A repertoire of correlational research suggest that relational responding is strongly linked to linguistic as well as cognitive performance (Cassidy et al., 2011; O'Hora et al., 2008; O'Toole & Barnes-Holmes, 2009). Barnes-Holmes et al. (2005a, 2005b) found that derived equivalence showed semantic priming effects as well as produced event-related potentials in a similar vein to language processing. In cohorts of children diagnosed with autism spectrum disorders (ASDs), derived equivalence relations were also correlated with verbal competence (O'Connor, Rafferty, Barnes-Holmes & Barnes-Holmes, 2009) and scores on the Vineland Assessment of Behaviour Scales (Moran et al., 2010). O'Connor, Rafferty, Barnes-Holmes & Barnes-Holmes (2009) also found that children with higher levels of verbal ability demonstrated improved performance regarding arbitrarily applicable sameness relations.

As will be discussed further in more detail, relational responding training procedures have previously demonstrated an efficacy for establishing many academically and intellectually relevant skills – such as fraction-decimal equivalence (Leader & Barnes-Holmes, 2001). More importantly perhaps, such a training regimen displays a potential for the generalisation of such abilities, such as the transfer of manding to novel contexts (Ehfelt & Root, 2005; Murphy & Barnes-Holmes, 2009a, 2009b; Murphy, Barnes-Holmes & Barnes-Holmes, 2005).

Comparison & Temporal Relations.

The relational frame of comparison is of defined importance in mathematics as well as our everyday language (Stewart et al., 2013). Comparison relations are employed “whenever one event is responded to in terms of a quantitative or qualitative relation along a specified dimension with another event (Hayes, Fox, Gifford, Wilson, Barnes-Holmes & Healy, 2001).

The importance of temporal relations (such as before/after), considered a subset of comparative relations, is clearly of importance in a social and occupational context due the anchoring role time and time keeping plays in society. Temporal relations may also have a strong link to intelligence, as documented by O’Hora et al. (2005). While temporal responding proficiency was not isolated in this study, it was part of a complex relation task involving sameness, distinction and temporal relations, that was then correlated with performance on 4 subtests of the WAIS-III. Results indicated that those who successfully completed the relational task received significantly higher scores for the vocabulary and arithmetic subscales. In an extension of this study, O’Hora et al. (2008) compared performance on the same relational task and the full-scale WAIS-III. Those who successfully completed the relational task received significantly higher scores on the Verbal Comprehension and Perceptual Organisation indices.

Furthermore, O’Toole & Barnes-Holmes (2009) employed the implicit Relational Assessment Procedure (IRAP) to assess relational responding involving the same three relational framing and examined its correlation with scores on the Kaufman Brief Intelligence Test (K-BIT). The IRAP is a methodology that assesses the speed of relational responding in accordance to rule-consistent and rule-inconsistent test batteries. Before the testing procedure, a collection of verbal relations was pre-established for each participant. Each participant would then be asked to either respond in accordance or defiance of that pre-existing rule in alternating sequence. Response latencies for the consistent trials were then subtracted from the inconsistent trials in order to gain a metric of relational flexibility. It was found that a lower score (resulting from lower levels of divergence for both conditions) predicted higher IQ. As noted previously, it appears that not only relational skills, but perhaps more crucially, relational flexibility is of key importance to intelligence.

Analogical Relations.

Analogy is a higher-level form of relational framing that usually requires proficiency in understanding relations between multiple items and across multiple networks. Due to its relative complexity, analogical reasoning is often cited as being of fundamental importance to intelligence (Esher, Raven & Earl, 1942; Sternberg, 1927, 1977). The first relational frame model of analogy was proposed by Barnes, Hegarty & Smeets (1997), describing it as the derivation of a relation of sameness between sameness relations. This model proposed that analogy is essentially equivalence-equivalence responding, as the individual must demonstrate stimulus equivalence within each relation as well as equivalence across the two relations. Barnes-Holmes et al. (2005a & 2005b) have demonstrated that patterns of neural activity witnessed during equivalence-equivalence relations parallels activation seen during natural language analogical reasoning. According to Stewart et al. (2013), performance on this model of analogical reasoning correlates strongly with traditional measures of analogical reasoning.

While this model has been supported using samples of adults (Barnes et al. 1997), Carpentier et al. (2003) reported that although nine-year old children demonstrated equivalence-equivalence responding, five-year-old children only demonstrated this capability after a specific training program is implemented. Stewart et al. (2013) cite such findings as support to the claim that due to the relational sophistication required, there may be a “developmental divide” between early and late childhood, similar to findings on analogical reasoning in mainstream developmental research (e.g. Sternberg & Rifkin, 1979). While such an argument has been contested (e.g. Goswami, 1991), more recent RFT research suggest that successful analogical reasoning that occurs before the age of 5 may be indicative of a process distinct from the type of equivalence-equivalence responding that characterises analogy (Capentier, Smeets, Barnes-Holmes & Stewart, 2004).

Perspective/Deictic reasoning.

According to RFT, perspective-taking is underlined by the process of arbitrarily applicable relational responding under the control of deictic contextual cues, such as I/you, here/then and now/then. Interestingly, McHugh, Barnes-Holmes & Barnes-Holmes (2004) found that accuracy in deictic framing appeared to increase as a function of age, which, as the researchers argue, would seem to identify such framing as being of an operant nature. It was also found that children appear to master spatial relations before temporal relations, which parallels many findings in mainstream developmental and theory of mind research (Stewart et al., 2013). Gore et al. (2010) extend this research by applying the deictic framing procedure employed by McHugh et al. (2004), along with IQ and language ability metrics to 24 adults with varying levels of intellectual disability. Results demonstrated that perspective taking correlated with verbal ability, Performance IQ and Full-scale IQ.

1.8.2 Enhancing Relational Skills

Central to RFT's contribution to increasing intelligence is multiple exemplar training (MET). As relational framing can be considered an operant process, MET is the most appropriate means of verbal relations and enhancing their sensitivity to contextual control. Such training regimen teaches the individual to derive relations across a set of relational frames (same/opposite, before/after etc.) across a large number of exemplars. The topography of the stimuli engaged in such exemplar should vary, as this facilitates the abstraction of the derived relation across numerous contexts, while also indicating the relative unimportance of the stimuli, as opposed to the relation. This training aims to facilitate the consolidation of complex relational skills that can thus be applied to a virtually infinite number of similar relational tasks (Cassidy et al., 2010). Barnes-Holmes, Barnes-Holmes, Roche & Smeets (2001) have demonstrated that MET is a reliable means of establishing generalisation of the relational skill of symmetry. In this study, 16 children between the ages of 4 and 5 were recruited, and then divided into four groups for four different experimental conditions. In each experiment, the children were trained to name two actions and two objects by demonstrating listening, echoic and tacting behaviours. Each different group was exposed to a different training procedure, and were tested for derived object-action symmetry relations. Explicit symmetry training was conducted by means of multiple exemplars, and proved to be far more effective than the other training methods. In fact, 13 of the 16 participants failed to show the required derived object-action or action-object symmetry until they received MET. The findings of this study have subsequently been replicated on numerous occasions (Barnes-Holmes et al., 2001; Gomez, Lopez, Martin, Barnes-Holmes & Barnes-Holmes, 2007; Luciano, Becerra & Valverde, 2007). In a further study, the relational frame of opposition was found to be established effectively by MET (Barnes-Holmes, Barnes-Holmes & Smeets, 2004). Berens and Hayes (2007) found that reinforced MET was successful in facilitating the development of arbitrary comparative relations, as well as their subsequent generalisation across stimuli and trial type. Due to its empirically validated efficacy across numerous relational frames, multiple exemplar training protocols have become a widely implemented means of establishing and improving relational responding.

One of the most promising of these protocols, emerging from the ABA literature, comes in the form of an autism evaluation and treatment program known as the PEAK relational training system (Promoting the Emergence of Advanced Knowledge; Dixon, 2014a 2014b). The PEAK training system provides a comprehensive tool for developing verbal, social and learning skills from basic to advanced by utilising a behaviour analytic perspective. The PEAK comprises of four modules – direct training, generalisation, stimulus equivalence and transformation of functions. Ranging from the training of fundamental learning skills such as eye contact and object permanence, to more advanced verbal skills such as understanding sarcasm and metaphor, PEAK has shown early promise with tentative correlational analysis with various IQ measures (Dixon et al., 2014). If the PEAK's behavioural assessments can be reliably demonstrated as showing a high degree of correlation with traditional IQ measures, this would seem to further underline the potential benefit of implementing behaviour analytic interventions in developing these "intelligent" behaviours.

The TARPA (Training & Assessment of Relational Precursors & Abilities, Moran et al., 2010) is a computer-based protocol designed to track the emergence and development of arbitrarily applicable relational responding. Such is the importance of AARR to generative language in particular, the developers of this system propose that a standardised methodology of monitoring the precursors and properties of AARR as they emerge in young children can be of great utility to basic researchers and practitioners in the field. The TARPA assesses a hierarchy of skills such as non-arbitrary conditional discrimination, arbitrary condition discrimination, mutually entailed relational responding, combinatorially entailed relational responding and transformation of functions (Moran et al. 2010).

The TARPA protocol comprises of six stages. Stage 1 involves a test of the participants' ability to make simple discriminations (two visual and two auditory). For example, participants would be exposed to an abstract picture and a blank box, appearing in random corners of a computer screen, and were then asked to choose the picture as opposed to the box. The second stage is an assessment of non-arbitrary conditional discrimination in which the matches are physically similar. For example, an item in this task would expose the patient to an auditory recording of a nonsense word. The participant would then be shown two visual stimuli on the screen successively, one being accompanied by the same nonsense word and the other accompanied by a different nonsense word. Participants were instructed to choose the stimuli with the matching nonsense word. Stage 3 mimicked the protocol for the previous stage, although on this occasion, the matches were not physically similar.

Stage 4 assessed mutually entailed relational responding across 3 levels (word to picture, picture to word & sound to word). This stage utilised a conditional discrimination format akin to the previous 2 levels, but also involved a training task. In the training task, the participant is required to select a particular auditory nonsense word from an array when there are exposed to a particular sound. In the testing task, they must then display mutually entailed responding by selecting the sound from an array after hearing the nonsense word again. Stage 5 trains and tests for combinatorial entailment, and involves one level comprising of three sections – word to picture, sound to picture and word to sound. The procedure is similar to preceding stages, but requires participants to derive the final conditional discriminative response based the two conditional discriminative responses already trained. To complete this stage, the participant must therefore, in independent trials, choose a particular nonsense word (A) and, in the next trail, a particular sound (C) in the presence of a particular abstract picture (B). To complete the stage, the participant must then derive the bi-directional relation combinatorially entailed between stimuli A and C based on their common relation to stimuli B. Finally, stage 6 is an adaptation of the Murphy et al. (2010) procedure, assessing the transformation of function, involving one level comprised of three sections – (i) mand training, (ii) conditional discrimination training & derived testing and (iii) derived mand testing. This stage begins by teaching the participant to mand using a particular token by selecting an onscreen button, which depicts that token. The token thus represents a particular mand function. In the second section of this task, the participant repeats the combinatorial entailment procedure seen in the previous stage, although this time using the token as the “A” stimulus and a new novel arbitrary stimulus as the “C” stimulus. In the third and final section, the participant must mand using the C stimulus, thus showing transformation of function of this stimulus in accordance with the derived combinatorial entailed relation.

In their preliminary analysis of the system, Moran et al. (2010), the TARPA was administered to five children diagnosed with autism. Their TARPA scores were then correlated with their scores on the Vineland Adaptive Behaviour Scales (VABS). The results of this correlational analysis suggested a high level of correlation between the TARPA and the VABS, which just fell outside the threshold for statistical significance. These findings would appear to implicate the utility of such an assessment and training system in improving the core relational abilities that underlie verbal behaviour. Subsequent studies (Kishita, Ohtsuki & Stewart, 2013) further underline the benefits that the TARPA may exert in improving language skill.

Due to the strong links between relational ability and intelligence, as well as the methods by which relational responding can be established and improved, it follows that a relational skills intervention may provide benefits for intellectual performance. In a study specifically focused on IQ measures, Cassidy et al. (2011) demonstrated the efficacy of two relational frame MET interventions in substantially increasing intelligence quotients in a sample of young children. In the first experiment, eight normally developing children between the ages of 8 and 12 were recruited to participate in a relational skills training regimen. These 8 participants were then divided evenly into an experimental group and a control group. Baseline WISC assessments were administered to all participants before commencing the relational training. There were five relational training phases in total: (1) stimulus equivalence training and testing, (2) MET for stimulus equivalence (3) MET to establish the relational frame of "sameness", (4) MET to establish the relational frame of "oppositeness" and (5) MET to establish the relational frames of "more than/less than". Experimental participants were exposed to all five stages, whereas control participants completed only the first stage. The training procedure took place two weeks after the initial IQ test administration. The training was delivered during ten 90-minute sessions over the course of 5-6 weeks. The follow-up WISC assessment was then administered 12 weeks after the baseline assessment.

All participants completed the first stage, which consisted of conditional discrimination training to criterion, followed by testing for symmetrical relations. Participants were then exposed to the same conditional discrimination training, followed by testing for derived transitive relations. Conditional stimulus relations were trained using a standard one-to-many matching-to-sample training protocol using nonsense syllables (e.g. A1 to B1, A2 to B2, A1 to C1, A2 to C2). All training was completed on a computer, in which participants selected one of the two options using a mouse. Blocks of 16 training trials (four for each relational frame) were administered until 100% correct responding was achieved. On-screen feedback was provided throughout. In the testing section, symmetry was assessed the following relations (B1 to A1, C1 to A1, B2 to A2, C2 to A2). The transitivity test then assessed the mutually entailed relations (B1 to C1, C1 to B1, B2 to C2, C2 to B2). The number of trials, block length and success criterion mimicked that of the training procedure.

In the second phase, experimental participants were exposed to MET and symmetry and transitivity testing for stimulus equivalence. Feedback was provided and then withdrawn on alternate blocks until stimulus equivalence performance became generalised i.e. participants displayed symmetry and transitivity with unseen stimuli without feedback. Five novel stimulus sets were employed for the intervention. For each of these, the training and testing cycle was presented with feedback once and once again without feedback. If symmetry or transitivity was not demonstrated during the first block, the procedure was repeated with a new stimulus set. When feedback was provided during testing, participants were exposed repeatedly to the symmetry and transitivity tests until they achieved the success criterion. The final training and testing cycle did not provide feedback during testing (i.e. MET).

In the third phase, participants were exposed to a series of contextually controlled conditional discrimination training and test blocks in order to establish the contextual functions of SAME and OPPOSITE for two arbitrary stimuli (see Steele & Hayes, 1991). This training required participants to discriminate between three comparison stimuli that were physically similar after being presented with a non-arbitrarily related sample in the presence of one or two contextual cues. Each task item began by exposing the participant to one of the contextual cues (same or opposite), and then presenting a sample stimulus in the middle of the screen. A second later, the sample stimulus was replaced onscreen by three physically similar stimuli. Each trio contained a stimulus that was identical, a stimulus that was slightly different and a stimulus that was opposite to the sample. In order to successfully complete the trial, the participant would have to select the stimulus that was in accordance with the contextual cue (same/opposite) that was presented at the outset. Each stimulus set was preceded by four pretraining tasks. In each 16-trial block of pretraining, each task was presented four times in a quasi-random order. Blocks were recycled until 100% correct responding was produced by the participant. If this criterion was not reached within four blocks, a new stimulus set was presented. The participant was re-exposed to a new training block until success criterion on a single block of 16 trials was achieved, or until four blocks had been administered. This procedure was continued until success criterion was met. Once this was achieved, the test for contextual control by the arbitrary cues was then administered. This assessment mirrored the procedure of the training stage, but omitted any feedback for responses. Novel stimuli were employed during the test.

In stage three, multiple exemplar training was administered to participants employing a procedure that was a combination of a Relational Evaluation Procedure (see Cullinan, Barnes-Holmes, & Smeets, 2001) and a Yes-No Procedure (see Fields, Adams, Verve & Newman, 1990). This training aimed to establish arbitrary SAME relations, which formed a coordination network during the testing (A= B, B = A, B = C, C = D). Each pairing was presented onscreen, with a contextual cue presented in between two stimuli. The words Yes and No were presented in counterbalanced positions on the bottom corners of the screen. Participants' choice of response was guided by corrective feedback. Participants were also trained to respond to the novel stimulus N1 was not the same as N2. This was included to safeguard against the possibly that the contextual cue SAME would directly control responding i.e. Same = yes. In order to complete training, a participant was required to produce 100% correct responding across a block of 20 trials. The testing stage assessed the establishment of the following relations without feedback: D SAME A (Yes), D OPPOSITE A (No), C SAME A (Yes) and C OPPOSITE A (No). The success criterion for this stage mirrored that of the training stage. If this criterion was not met, the training and testing cycle was repeated using a novel stimulus set. If this was necessary, feedback was provided during testing until criterion was met. The participant was then exposed to another training and testing cycle, which did not provide feedback. This cycle continued until success criterion was met.

In stage 5, multiple exemplar training for the relation frame of OPPOSITE was administered. Contextual control by the arbitrary SAME and OPPOSITE cues had already been established in the previous stage. The training and testing procedure was identical to that of the previous stage, with the contextual cues for SAME and OPPOSITE alternating positions in each task item. Stage 6 also followed the MET procedure utilised in the previous two stages, while training for MORE THAN and LESS THAN relations. For example, for a MORE THAN item, the participant was exposed to an image of two balls as a sample, and would then be asked to choose the image of three balls, rather than the image of one ball. This training was conducted in a single protocol. The training utilised three separate frames that established that A > B > C > D. Three additional frames precluded direct control by contextual cues by including the following: A < B (No), N1 > N2 (No) and N1 < N2 (Yes). Passing criterion was 100% correct responding across 30 trials. The testing task items were as follows: D > A (No), D < A (Yes), C > A (No), C < A (Yes), A > D (Yes), A < D (No). Passing criterion for the testing phase was also 100% accuracy across 30 trials.

Results of the study indicate a significant effect for phases and their interaction. As there were no significant differences between experimental and control participants at baseline, this interaction is explained by the more complex training received by the experimental group. It was found that the multiple exemplar training exerted a significant effect on Full Scale IQ, with experimental participants exhibiting a mean increase of 27.25 points. Control participants displayed a mean decrease of 2.25 points. Furthermore, Cassidy et al. (2001) also reported statistically significant differences in IQ score changes for both Verbal and Performance IQ, with experimental participants demonstrating remarkable rises (17.75 and 32.5 respectively) compared to controls (0.25 and -4 respectively).

In Cassidy et al.'s (2011) second study, eight 11- and 12-year-old children with educational difficulties took part in a multiple exemplar training program to train SAME/OPPOSITE and MORE THAN/LESS THAN relations. The training was conducted over a 6 to 14 week period and was both preceded and succeeded by the administration of a full scale WISC-IV assessment. The training schedule generally adhered to twice-weekly training sessions with duration of 90 minutes. Due to the limited access and availability of the schoolchildren recruited, training was spread out over a 9 month period. The training and testing protocol itself followed the same format as the latter three stages of Study 1. Following research conducted by Berens and Hayes (2007) remedial training protocols were also utilised to aid in the generalisation of relational operants. In the instance that a participant did not pass the SAME relational testing phases within seven cycles of training, testing and MET testing, remedial training and testing was employed. This remedial training was identical to the standard training and testing cycles, but replaced the nonsense syllables used with non-arbitrarily related stimulus sets (e.g. lines, circles, boxes etc.). Once 100% correct responding was displayed on this remedial cycle, the participant returned to the standard cycle of training and testing.

An additional metric, a Relational Abilities Index (RAI) was also devised and computed for this study. The RAI, consisting of three stages of successive blocks testing for SAME/OPPOSITE & MORE THAN/LESS THAN relations, was used to assess baseline fluency of relational ability of all participants. Each relation type was assessed across 60 test trials, 20 per stage. Each test block was administered once and resulted in an overall RAI score out of 60. The RAI took the form of a collection of syllogistic relational networks requiring the participant to respond either a Yes or No response based on that network. For example, participants may be exposed to the following logical puzzle: A > B, B > C and then asked is A > C? No feedback was provided during this assessment. Each trial consisted of a novel stimulus set. SAME and OPPOSITE were assessed in separate test blocks, while MORE THAN & LESS THAN were tested and trained in the third block. In Stage 1, two statements (e.g. A OPPOSITE B, B OPPOSITE C) were presented, along with a question (e.g. is A OPPOSITE to C?). Stage 2 followed the same format, but reversed the order of presentation for the two statements. Stage 3 followed the same format as Stage 2, but was presented with a 5 second time limit.

Results indicated that the training program resulted in a significant increase in overall relational performance, with mean correct responding rising from 58.5% (just above the 50% chance level) to 92.4%. This increase was significant across all four relations. Significant improvements were also found for Full Scale IQ, with a mean increase of 13.1 points. There were also significant rises in three of the four IQ subscales (Verbal Comprehension, Perceptual Reasoning and Processing Speed). The findings of Cassidy et al. (2011) lend further support to the suggestion that RFT-based relational training programs can result in improvements in intellectual functioning as assessed by traditional IQ tests. It also provides additional empirical evidence for the argument that derived relational responding is of fundamental importance to intelligence. The IQ rises demonstrated in their study are particular robust, given the control measures implement to preclude the possibility of practice effects or educational or developmental processes accounting for these improvements.

Cassidy et al. (2016) expands upon these findings by implementing the same relational training in two separate experiments. The RAI implemented in this analysis extended upon the assessment previously implemented in Cassidy et al. (2010) by assessing a more expansive repertoire of relational responding. The first of these experiments analysed the effect of relational skills training for SAME/OPPOSITE and MORE/LESS in a group of fifteen primary school children aged between 10 and 12. Following the intervention, analyses of WISC-IV scores found there was a significant increase from baseline to follow-up, with the mean IQ of the group rising 23 points from 97 to 120. Furthermore, every child included in the study displayed increases, with the lowest of these (14 points) falling marginally short of one standard deviation. All other participants showed increases above one standard deviation, with three participants displaying rises approaching two standard deviations. Mean RAI scores for this group also rose significantly from 33.8 (out of 55) to 48.5, implicating the efficacy of the training in improving relational skills.

In the second study, the same relational skills procedure was administered to a sample of 30 secondary school students aged between 15 and 17. In this experiment, the effects of relational training on a widely used standardised measure of scholastic ability (Differential Aptitude Test; DAT, Bennett, Seashore & Wesman, 1990). The DAT is a group-administered assessment of a number of skills and abilities that are deemed relevant to academic performance. The DAT consists of eight subtests: Verbal Reasoning, Numerical Ability, Clerical Speed and Accuracy, Abstract Reasoning, Mechanical Reasoning, Space Relations, Spelling and Language Usage. The verbal and numerical ability domains of the DAT are the domains of most relevance theoretically to the relational framing skill intervention, and therefore were the only subtests administered. An Educational Aptitude score can be computed from these subtests, and provided an overall reflection of an individual's scholastic ability. Results found that there were significant increases in both Verbal Reasoning and Numerical Ability following the intervention. There was also a significant increase found for the Educational Aptitude composite score. Finally, significant increases were also found for RAI scores. Such findings would support the argument that such a relational skills training program can provide practical and demonstrable improvements in intellectual performance that extend beyond IQ and into the academic domain.

As discussed in previous sections, relational ability has shown a high level of correlation with various traditional measures of IQ (Andrews & Halford, 1998; Barnes-Holmes, Barnes-Holmes, Stewart & Boles, 2010; Cattell, 1971; Dixon et al., 2014; Gentner & Loewenstein, 2002; Gore et al., 2010; Moran et al., 2010; 2014; O'Hora et al., 2005; O'Hora et al. 2008; O'Toole & Barnes-Holmes, 2009; Stewart et al., 2013). Experimental evidence, taken in conjunction with the apparent theoretical overlap between relational ability and intellectual activity (Hayes et al., 2001; Jensen, 1998; Spearman, 1904, 1927) implicate the relevance of relational ability to intellectual development.

As the Relational Abilities Index employed by Cassidy et al. (2016) was specifically designed to target and assess an individual's sophistication in relational responding, it is expected that a strong correlation would be found between this index and a mainstream measure of intellectual performance, the WAIS-III. Previous analyses reported significant correlations between measures of relational responding fluency and a number of subtests of the WAIS-III, such as Vocabulary, Arithmetic, Digit-symbol Coding, and also two of the four subindices, Verbal Comprehension and Perceptual Organisation (O'Hora et al., 2005; 2008). However, Cassidy et al. (2011, 2016) did not report upon correlations between RAI scores and the various metrics produced by a full-scale IQ assessment. Therefore, a study of the relationship between RAI and IQ indices, subindices and subtests may produce a more comprehensive understanding of the nature of relational responding's importance to intellectual performance. While high levels of correlation between relational ability and IQ measures can be expected, it is important to note that such correlations should not be perfect. Relational ability may display sufficient utility in describing intellectual performance as to be considered somewhat of a proxy for intelligence. However, it is not to be regarded as being a pure measure of intelligence as its remit of measurement is not identical to traditional IQ assessments. That is, it is not yet known precisely what aspects of traditionally defined intelligence are synonymous with generalized relational responding repertories. Furthermore, the objective of an RFT programme of research into intelligence is not to merely understand "intelligence" but to provide a new and radically different conceptualisation and definition of this term. In that regard, the aim is not to assess the RAI in terms of its convergence with traditional IQ measures, but to use it as a jumping off point to begin to explore how relational skills measures and IQ measures overlap.

1.9 Potential Boundary Conditions for Intellectual Improvements

Due to the efficacy relational training has previously displayed in increasing IQ (Cassidy et al. 2011, 2016), it is of interest to investigate a number of potential boundary conditions that may impact of the effectiveness of such training.

1.9.1 Age

IQ and standardised intellectual performance has long been proposed to remain relatively stable across an individual's lifetime (Ghisletta & Lindenberger, 2004; Horn & Cattell, 1976; Jones & Bayley, 1941; Mackay, Connor, Albert & Obler, 2002; Schaie, 1983, 1994; Schaie, Labouvie & Buech, 1973; Weinert & Hany, 2003; Weinert & Schneider, 1999). Short-term (< 6 months) test-retest correlations coefficients tend to be .8 and above for Full, Verbal and Performance IQ, as assessed by the WAIS (Covin, 1977; Irwin, 1966; Quereshi, 1968; Wechsler, 1974) while similar results have been found for the WISC (Stavrou & Flanagan, 1996; Watkins & Smith, 2013). Deary, Whalley and Crawford (2004) found a significant correlation coefficient of .74 between participants' cognitive performance scores as measured by the Moray House test aged 11 and the National Adult Reading Test at the age of 77. In a follow-up study, Deary, Whiteman, Starr, Whalley & Fox (2004) reported that a correlation coefficient of .61 was found between a cohort's scores on the Moray House Test at ages 11 and 80. Schwartzmann, Gold, Andres, Arbuckle & Chaikelson (1987) found a correlation of .82 in IQ scores for individuals in their twenties and then again in their mid-sixties. Mortensen & Kleven (1993) found correlations coefficients of .94 and .9 for 50-year-old participants retested with the WAIS after 10- and 20- year interval respectively. Owens (1996) reported a coefficient of .78 on the Army Alpha assessment after an interval of 42 years. In a study of individual IQ scores taken from the age of 2-3 up to 40, McCall (1977) found that IQ stability tends to increase with age. For example, correlations between IQ from age 3 to 4 was as slightly lower than .4 for males, rising to .8 between scores at ages 12 and 14. Such a research repertoire would suggest a degree of robustness to adult IQ scores.

There is however, evidence that IQ in childhood can vary significantly in a way unlike adulthood (Bayley, 1940, Bradway, 1944; Sontag, Baker & Nelson, 1955; McCall, Appelbaum & Hegarty, 1973; Ramsden et al. 2013). While adult IQ is propose to remain relatively stable, it is also proposed to decrease in old age. Lower raw scores found for the older groups of Wechsler's adult scales when compared to younger groups has been used to assert a diminishment of intellectual performance through ageing (Baxendale, 2011; Ryan, Sattler & Lopez, 2000; Wechsler, 1958). While comparing the IQ of different age groups may be at the mercy of a number social confounds, Dickinson & Hiscock (2010) determined the group difference in norms for 20- and 70-year olds as well as a Flynn effect difference between the subtests of the WAIS and the WAIS-R and also the WAIS-R and WAIS-III. This then allowed the estimation that the true ageing effect for each subtest (i.e. the decline in performance due to natural ageing) stood at approximately 15%. Flynn (2009) subsequently recommended a correction to the WAIS-III normative tables, it was then concluded that the Flynn effect accounted for 100% of the difference in performance between the 20- and 70-year old groups. This finding proposed that the lower scores displayed by the older sample were due to social factors and actual poorer performance, rather than a decline in intellectual function i.e. supporting the stability of IQ through adulthood. This finding provides stark contrast to Wechsler's (1958, p.142) assertion that "the abilities by which intelligence is measured do in fact decline with age." Furthermore, after controlling for educational level, Birren & Morrison (1961) & Kaufman, Reynolds & McLean (1989) found that the age difference in normative tables for Verbal IQ disappeared.

Differences in Performance IQ, however, persisted supporting research which proposed that fluid intelligence (major contributor to Performance scales), but not crystallised intelligence declines with age (Horn & Cattell, 1967; Wang & Kaufman, 1993; Kaufman, Johnson, & Xin Liu, 2008; Cunningham, Clayton & Overton, 1975; Bugg, Zook, DeLosh, Devalos & Davis, 2006; Manard, Carabin, Jaspar & Collette, 2014).

Age has been shown to be an influential factor regarding relational responding (Andrews & Halford, 1998, 2002; Gentner & Rattermann, 1991; Sugarman, 1982; Younger, 1993). Children appear to be increasingly proficient in responding to complex relational items as they mature. For example, unary relations (relations involving only one argument, e.g., Ben is a boy) appear to be understood at the age of 1 (Sugarman, 1982; Younger, 1993), while binary relations (involving two arguments, e.g., Ben is smaller than Jack) appear to be established by the age of 2 (Halford, 1982). In a study of ternary relations, Andrews and Halford (2002) demonstrated that 15.5% of 3- and 4-year olds, 48.3% of 5 year olds, 70.2% of 6 year-year-olds and 77.8% of 7- and 8-year olds could respond correctly in accordance to ternary (three argument) relations. In a comparison of relational ability across numerous age ranges, McHugh et al. (2004) assessed performance on a perspective taking relational task across 5 age ranges. The age groups were segregated as adulthood (18-30 years), adolescence (12-14 years), late childhood (9-11 years), middle childhood (6-8 years) and early childhood (3-5 years). Results indicated that errors in responding to a range of relational frames decreased as a function of age.

Due to the stability of adult IQ, as well as the finding that relational sophistication appears to increase with age, the effectiveness of a relational training program may be diminished when employed in adult populations. As an advanced level of relational responding may be expected of an adult sample of college students, it is of interest to ascertain the relative benefits a relational training program may provide. Due to the age-related confounds that may present themselves in an adult populations, it is of defined importance to ascertain the whether the IQ gains found in younger populations as a result of relational training (Cassidy et al. 2011, 2016) are consistent with those found in the current adult sample.

1.9.2 Baseline Intellectual and Relational ability

While a very small number of previous studies have investigated the effect of relational training on child populations (Cassidy et al., 2011, 2016), the potential impact of such training on adult populations has not been investigated at all. As stated previously, IQ scores for children can vary as they mature, while there may be more stability in IQ scores in adulthood. Therefore, the possible delimiting effect of such stability on increases in intellectual functioning following relational training warrants closer study. Furthermore, as relational skill typically improves through childhood and into adulthood (McHugh et al., 2004), the higher level of performance expected for an adult population may significantly limit the improvements in ability that a relational skills training programme could offer. In effect, a relational skills training system of the kind used by Cassidy et al. (2011) may limited by a ceiling effect when applied to adults.

In addition to developmental age, baseline IQ score may also have a deleterious effect on the benefit an individual may glean from relational skills training. As of yet, there has not been an investigation into the efficacy of relational skills training in producing IQ increases in high IQ participants. Due to the close relationship between relational responding and intellectual performance, the issues of both baseline IQ and baseline RAI scores inhibiting the level of improvement seen post-intervention are closely intertwined. In Cassidy et al. (2016), the younger group, consisting of primary school children with comparatively lower levels of relational skill, showed mean increases in RAI scores that were double that of the teenager group. It must be noted that while Cassidy et al. (2016) reported that there was no significant correlation between baseline RAI and IQ rise, the current sample is expected to have even higher baseline RAI scores which may result in a significant ceiling effect. The current study thus aims to address this issue and attempt to extrapolate the effects of age and starting levels from IQ increases, using correlational analysis.

From a purely statistical standpoint, high IQ participants by definition have fewer IQ point gains to make, due to the natural limits of the testing tools. This fact, allied to the high level of relational sophistication predicted by their IQ, may limit the utility of relational training in improving their performance, resulting in significantly lower IQ increases when compared to child cohorts with average IQ's. The current study is the first to assess the impact of this relational skills training protocol on an adult sample. As college students typically display above-average IQ scores (Herrnstein & Murray, 1994; Plant & Richardson, 1958; Educational Testing Services, 2015; Wolfle & Oxtoby, 1952), they represented a convenient sampling frame for above average IQ adults as participants for this study.

1.10 The Current Thesis

The current thesis involves two research investigations that aim to further contribute to the growing research repertoire pertaining to relational skills and relational skills training. The first of these investigations assessed the efficacy of a relational skills training program in improving intellectual performance (as assessed by WAIS-III) in a sample of high IQ adults. As such, this investigation assessed high IQ and age as possible delimiting factors for the potential efficacy of a relational skills intervention. The study recruited a sample of 34 young adults who were currently attending or had recently graduated from third level education. Previous analyses have investigated the impact of the training program using primary and secondary school children within the average IQ range, but as of yet there has not been a randomised controlled trial using high IQ adults. It was hoped that this analysis may shed light on possible boundary conditions for the relational training program, specifically age and baseline IQ. Relational Ability Index scores were also assessed at baseline and post intervention as a potentially interesting covariate of intellectual performance increases.

The second study focused on the degree of correlation between the Relational Abilities Index (RAI) and traditional IQ as measured by the WAIS-III. IQ and RAI data collected from Study 1 was repurposed for this analysis. Correlational analyses of the relationship between relational responding, as measured by the RAI, and the three main IQ indices (Full, Verbal & Performance IQ) were of primary focus. Correlations between RAI scores and the four IQ subindices (Verbal Comprehension, Working Memory, Perceptual Organisation & Processing Speed) were also investigated, alongside standardised scores on all 13 subtests administered. It is hoped that such analyses would provide further insight into the relationship between relational skills and general intellectual functioning, as well as producing specific information on the types of intellectual tasks it contributes towards.

Chapter 2

Study 1: Assessing the efficacy of a relational training intervention in improving intellectual performance in a sample of above average IQ adults

2.1 Study 1

The current study aimed to build upon previous findings by analysing the efficacy of a relational skills training program in improving IQ scores, but using a control sample that received no intervention. As relational responding has been shown to be of considerable importance to intellectual performance (O'Hora et al. 2005; 2008; O'Toole & Barnes-Holmes, 2009; O'Toole et al., 2009; Stewart et al., 2013), a number of intervention programs have been devised to isolate and improve relational skills (Cassidy et al., 2008; Dixon et al., 2014; Moran et al., 2010). The current study implemented the relational training program employed previously by Cassidy et al. (2016) based on the protocol outlined by Cassidy et al. (2008). This relational skills program aims to increase intellectual performance, and therefore intelligence quotient, by targeting the relational frames of same/opposite and more than/less than. The impact of relational training on IQ indices, subindices and subtests will be analysed in order to ascertain the effect the current relational skills training program can exert on the intellectual performance of high IQ individuals.

2.2 Method

2.2.1 Participants

In total, thirty-four participants (19 females) were recruited to participate in a study to assess the efficacy of a relational skills training program versus a no intervention group in raising IQ scores, as assessed by the WAIS-III. All participants were currently attending third level education or had recently graduated (<1 year). Participants were informed at the outset that they should not volunteer if they had attended special education outside the mainstream schooling system, or have been diagnosed with a learning difficulty.

All recruited participants were administered the full battery of the Wechsler Adult Intelligence Scale (3rd edition) as a measure of general intellectual performance. Eighteen participants were required to complete a relational skills training program in order to assess its efficacy in improving intellectual performance. Participants allocated to the intervention group were also required to complete a specially-devised online assessment of relational ability, the Relational Abilities Index (RAI). Following completion of the training program, participants were administered repeat assessments of the WAIS-III and the RAI. A further 16 control participants were recruited, who did not participate in the training program but would be administered the same WAIS-III assessment at baseline and follow-up. The testing intervals for this group were approximately 3 months, in line with the average completion time for the training intervention.

Participants ranged from 18 to 44 years old (M = 22.2 years, SD = 5.13). In terms of educational level, 21 participants were currently completing a Bachelors degree (16 Bachelor of Arts, 5 Bachelor of Science). A further seven participants had recently graduated with a Bachelors degree (5 Bachelor of Arts, 2 Bachelor of Science), while six participants were currently completing a Masters degree. Of the 34 participants, three experimental participants failed to complete the training or follow-up assessment.

2.2.2 Settings and materials

All assessments took place in a private experimental room 4m x 8m approx. that contained two chairs, a desk, as well as an observation window.

2.2.2.1 Wechsler Adult Intelligence Scale III

Each participant was administered the Wechsler Adult Intelligence Scale (WAIS-III: UK; Wechsler, 1997), an individually administered clinical instrument for assessing overall intellectual ability. The full test battery comprises of 13 subtests, along with two other optional procedures that are not necessary for the calculation of IQ indices. For the purpose of this study, only the 13 core subtests were administered (i.e., Picture Completion, Vocabulary, Digit-Symbol Coding, Similarities, Block Design, Arithmetic, Matrix Reasoning, Digit Span, Information, Picture Arrangement, Comprehension, Symbol Search & Letter-Number Sequencing). Three composite scores for performance intelligence quotient (PIQ), verbal intelligence quotient (VIQ), and full-scale intelligence quotient (FSIQ) can be calculated from these subtest scores. Full-scale IQ was considered to be the primary outcome variable, with its entailed subscales and subtests used to further inform main effects. The Object Assembly and Digit-Symbol Copy measures were not employed, as these are supplementary rather than core procedures, and as such are not required to compute any of the relevant composite scores (i.e. Performance, Verbal and Full IQ and their respective subindices). Administration time was approximately 90 minutes with breaks allowed as appropriate.

2.2.2.2 Relational Abilities Index

Participants were instructed to complete a relational abilities assessment in order to compute their current level of sophistication in identifying complex relations between items. The Relational Abilities Index (RAI) assessment is administered via the website RaiseYourIQ.com. The RAI took the form of a series of syllogistic relational network problems requiring the participant to respond with either a Yes or No response based on a posed relational question. The RAI consists of a battery of 55 relational puzzles delivered using three letter nonsense words as relata. For example: “If CUG is opposite to BEH, and BEH is opposite to MER, is MER opposite to CUG?” The user answers each question by clicking on a YES or NO button located on the computer screen. The positions of these response buttons were randomly switched repeatedly throughout the assessment to control for positional responding. The questions remained on screen until the user has provided an answer. Block 1 consisted of 29 SAME/OPPOSITE tasks while Block 2 consists of 26 MORE/LESS tasks. The Same/Opposite and More/Less task blocks are presented in sequence with no break or additional instructions between the blocks. Trial stimuli were comprised of three-letter nonsense words (e.g., BEF, DIL, FAS) in the formulation *vowel-consonant-vowel*, which were generated randomly by the test software. English language words and known slang words were omitted. A total of 248 stimuli were used in the RAI test with no stimulus appearing more than once.

As the RAI assessment proceeds, task complexity increases. Complexity is controlled in terms of; 1) the number of sample relational statements presented; 2) the order in which these statements are presented (i.e., in a sequential or random order); 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) whether or not the relational statement or question utilised more than one relation type (e.g., presentation of only same relations, or combination of same and opposite); 5) whether or not the relational term presented in the question was present in any of the premises (e.g., premises A>B, B>C followed by the question: is A<C?). Each task had a 30 second time limit. Failure to respond within this limit was treated as incorrect responding. The average time for completion of this Relational Skills assessment is ten to fifteen minutes. Upon completion of this assessment, a number of

valuable metrics (correct responses, total completion time, average item completion time etc.) are computed, which offer a greater understanding of baseline relational ability.

Figure 1 outlines 6 sample relational tasks representing the 3 general degrees of difficulty across two relational types (SAME/OPPOSITE & MORE THAN/LESS THAN). All relational frames are assessed by a relational evaluation procedure in which the participant is exposed to a network of 1-3 relational statements and then asked a question relevant to those relational statements. Tile 1 describes a task type that assessed the relational skill of identity matching. Identity matching is referred to as non-arbitrary relational responding (NARR; Reese, 1968), as participants respond based on the formal properties (e.g. size, shape, colour etc.) of the stimuli being presented. As such, the participant is required to decipher whether the sample relation matches the relational question being posed, as is therefore representative of a simple matching-to-sample task. At this most basic level, the relational question topographically matches the relational statement presented (e.g. A>B, is A>B?). In effect, the participant virtually bypasses the linguistic or communicative value of these presentations, and responds due to the identical forms of the statement and question rather than what they actually communicate. The relational statement and relational question therefore serve in a similar capacity to matching shapes or objects in this case. Assessments of reflexivity are also present even in the earliest stages of training as the presentation of sample relations are altered and reversed. Tile 2 describes task types that assessed the same type of relational responding but uses the relational frame of MORE and LESS. In this case, the question posed does not match the sample relation presented, increasing the level of difficulty. By reversing the relation between stimuli, this task assessed the relational property of mutual entailment. Mutual entailment refers to the derivation of a bidirectional relation from a unidirectional relational premise. For example, if an individual is taught that A is greater than B, the relation of B is less than A is mutually entailed.

Tiles 3 and 4 illustrate the assessment of combinatorial entailment across the two relational frames. Combinatorial entailment refers to the process by which the presentation of two stimulus relations allows the derivation of an implied third relation (i.e. A > B and B > C, allows the derivation of A > C). Tiles 5 and 6 illustrate the assessment of complex combinatorial entailment in which three stimulus relations allow the derivation of a fourth relation (i.e. A > B, B > C, C > D allows the derivation of A > D or D < A). Through these varied presentations of relational tasks during the training stage, the individual practices and learns a comprehensive repertoire of relational responding from basic stimulus equivalence responding to complex forms of combinatorial entailment. Most importantly, due to the multiple exemplar format used, these forms of relational responding become generalised and applicable to novel relational frames and networks.

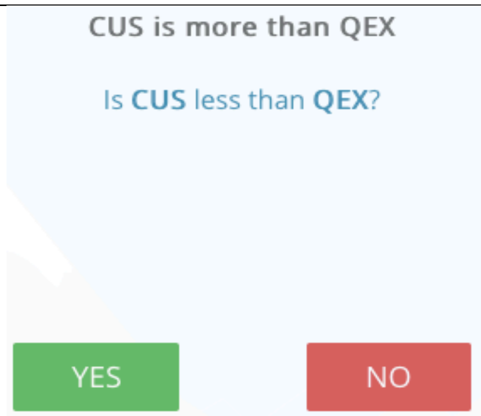
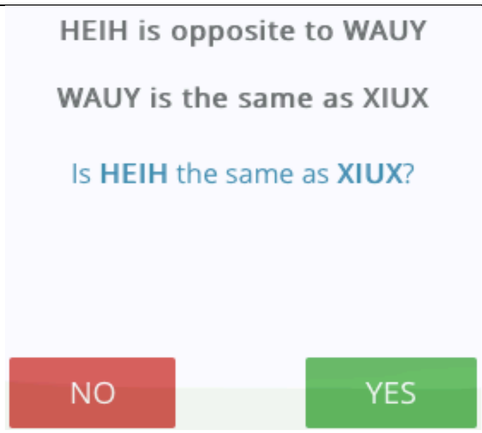
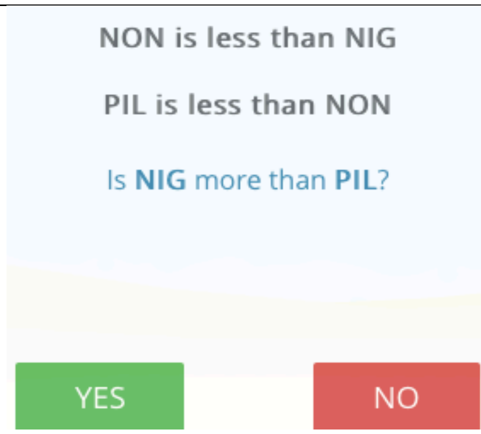
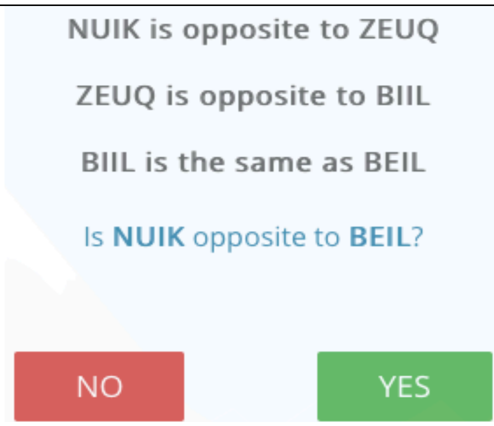
Same/Opposite	More/Less
1 	2 
3 	4 
5 	6 

*Figure 1.*Examples of the types of relational task involved in the RAI assessment

For the first two levels assessing each relational frame (SAME/OPPOSITE & MORE THAN/LESS THAN), a single relational premise (e.g. "is A the same as B") is followed by a relational question (e.g. "is A the same as B?"). Task complexity for these questions increased by varying the content and directionality of the relational question posed (e.g. "is B the same as A?", "is A opposite to B?", "is B opposite to A?"). Task difficulty was also increased by modifying whether the relational question contained the relational type as presented in the premise. Furthermore, tasks were made more difficult by controlling the directionality of the relational question. This format is then repeated for subsequent tasks which include additional relational premises. For the SAME/OPPOSITE relational frame, questions 1 and 2 pose a single relational premise, questions 3 to 21 contain two relational premises, followed by three premise tasks from questions from 22 to 29. Similarly, for the MORE THAN/LESS THAN relational frame, questions 30 & 31 include one relational premise, question 32 to 44 include two relational premises and questions 45 to 55 include 3 relational premises.

2.2.2.3 Relational Training Protocol

The relational training program mirrored the format of the relational abilities assessment previously outlined by extending the 55 individual task items into 55 multitask training levels across two blocks (SAME/OPPOSITE & MORE THAN/LESS THAN). More specifically, during the training stage, each of the 55 assessment items were individually trained across multiple exemplars in a training stage of potentially infinite trial number and duration. Each of the 55 task types present in the RAI assessment were thus targeted and extended upon during training (using multiple exemplars comprising of novel stimuli), with the aim of establishing correct responding to each. During each training stage, participants were required to produce 16 consecutive correct responses, and as such, each stage presented a potentially infinite sequence of relational framing tasks (i.e., of the same form but using different nonsense words). The 30-second time limit also applied during training tasks, and failure to respond within this window was treated as an incorrect response.

During the training stage, responses were succeeded by corrective feedback to guide responding (i.e., the words Correct or Wrong appearing on the screen for one second alongside a short tone for correct responses, but not incorrect responses). The training program employed a randomization of stimuli that ensured participants were not exposed to any nonsense word in more than one trial across all 55 levels of training. If participants produced 100% correct responding on their first 16 training tasks, the testing stage for that level could be skipped. This process was only relevant to the first ten training stages. After ten stages, participants were required to complete the testing stage regardless of training stage performance. If an incorrect response was produced, participants were required to produce 16 consecutive correct responses before proceeding to the testing stage.

A testing stage was delivered after each training level was mastered. The testing stage matched the relational complexity of the training stage that preceded it. The testing stage consisted of 16 novel tasks that were presented without corrective feedback. Errorless responding on all 16 tasks was required to pass the test. In effect, testing stages were identical to training stages, with the difference that no feedback was provided, and only 16 trials were presented in a finite block. The 30s response time limit applied on each trial of testing. If this test was failed, the participant was directed to complete the training stage once again. This process would repeat until the participant satisfied the success criterion for both the training and testing stages for every level. Once this was achieved, the participant progressed to the next level of training. Participants could not progress to the next level without completing the previous level. This process continued until participants completed 55 levels and were administered the follow-up RAI assessment.

Each level mirrored the progressively increasing complexity of the RAI assessment across the two blocks. In this way, task difficulty progressively increased from level 1 to level 29 for the SAME/OPPOSITE trial types, before this difficulty level was reset for the presentation of the MORE THAN/LESS THAN trial types in stages 30-55. In keeping with the RAI format, the first two levels of each relational frame posed a single relational premise followed by a relational question. Once again the complexity for these levels increased by altering the relational question posed (e.g. "is A the same as B", "is B the same as A?", "is A opposite to B?", "is B opposite to A?"). Level of complexity and difficulty was progressively increased by controlling the number of relational premises, the order in which these premises were presented, the directionality of the relational question, the number of relational types presented in the relational premise or question (e.g. only same relations, combination of same & opposite etc.) and whether of the relational type in the question was present in the premise(s). The schedule for the introduction of additional relational premises also mirrored that of the RAI assessment. As such, for the SAME/OPPOSITE relational frame, levels 1 and 2 pose a single relational premise, questions 3 to 21 pose two relational premises and questions 22 to 29 contained three premises. For the MORE THAN/LESS THAN relational frame, levels 30 & 31 include one relational premise, levels 32 to 44 include two relational premises and questions 45 to 55 include three relational premises.

Participants were encouraged to train at their leisure, at home or in college, 2-3 times per week always ensuring a consistently quiet and non-distracting environment. The online training program capped the maximum possible level completed per day at 5 levels, ensuring that training was not condensed into a small number of sessions. Due to this level cap, participants required at least 11 sessions in order to fully complete the 55 levels of training. The experimenter could access an administrative dashboard which allowed the monitoring of each participant's progress, as well as tracking metrics such as total items completed, total correct responses, number of log-ins and information on the RAI assessment. As training took place at the participant's leisure, the online program also allowed the experimenter to prompt participants if they had stopped training regularly.

2.2.3 Procedure

The procedure implemented in the current study was approved by the Maynooth University's Social Research Ethics Committee. All participants were provided with information sheets, as well as consent forms at the outset. The study as reported here and as described to participants in the information sheet was approved by the Maynooth University research ethics committee. Study participation began by asking participants to read an information sheet detailing the aims of the study, their degree of participation and the general experimental sequence. This sheet also gave participants contact details for the primary researcher, the research supervisor and also an educational psychologist who was available to address any concerns with their subsequent IQ report. It must be noted that participants did not receive an IQ report until their participation was complete. Participants were then asked to sign a consent form. The initial baseline WAIS-III assessment took place in a private experimental room in the Department of Psychology, Maynooth University. Participants sat across a desk from the experimenter during the assessment. Participants were informed that they could take breaks if they wished during the assessment. The IQ assessment procedure was explained at the outset. Participants were also informed that they are welcome to ask any questions throughout the testing procedure. The testing procedure for follow-up and baseline assessments was identical.

Once experimental participants completed the baseline RAI, they could begin training. After 55 levels of training were completed, participants were directed to complete a second RAI assessment. All participant progress was tracked remotely by the experimenter. The follow-up WAIS assessment was arranged at the soonest time of convenience following the completion of the second RAI assessment, after a minimum testing interval of 3 months. As the testing interval for experimental participants was contingent on completion of the training, the testing interval for the control group was based on the average time it took to fulfil this criterion (3 months). The second RAI assessment was identical in format and difficulty to the baseline assessment, but different stimulus sets were employed. In the event that a participant completed training in advance of this minimum retest interval, the online program allowed them to revisit previous stages and continue training. Control participants received no intervention during the testing interval, and were not given access to the online training program during their participation. In interest of fairness however, a training account was offered to such participants following the conclusion of the study. For both groups, the follow-up assessment was identical to the baseline assessment, as participants completed a second WAIS-III assessment. Following participation, a report was provided to participants outlining their baseline and follow-up IQ scores.

2.3 Results

Study 1 aimed to analyse the efficacy of a relational skills training program in improving IQ scores in comparison to no intervention, using a sample of high IQ adults. A series of mixed between-within analyses of variance were first computed to investigate differences between the intervention and non-intervention groups regarding pre to post-intervention changes in IQ scales. A series of paired sample *t*-tests were then computed to further analyse the nature of changes in IQ scores across groups. The impact of relational training on Relational Ability Index scores was also studied.

2.3.1 Baseline IQ and RAI scores

Table 1 shows that mean scores for the three main IQ indices (Full, Verbal & Performance) were in the high average to superior categories of performance for the current sample. Mean scores for this sample on Verbal Comprehension, Working Memory and Perceptual Organisation were also in the high average category, while mean Processing Speed scores were in the average range. The mean Relational Ability Index score was 48.94 out of a maximum score of 55.

Table 1

Study 1 Baseline Descriptive Statistics

Variable	Mean	SD	Range
Age	21	2.58	18-26
RAI	48.94	5.79	36-54
Full IQ	118.11	13.04	95-137
Verbal IQ	120.22	13.83	99-143
Verbal Comprehension	118.61	13.25	100-145
Working Memory	112.72	18.17	80-141
Performance IQ	111.5	12.2	87-136
Perceptual Organisation	113.17	12.77	88-133
Processing Speed	102.67	11.8	86-128

2.3.1 Analysis of IQ Change

2.3.1.1 IQ indices

An independent-samples *t*-test was conducted to compare baseline Full IQ scores for control and training groups. There was no significant difference in scores for the control group (M = 111.8, SD = 11.16) and the training group (M = 118.9, SD = 12.76; $t(-1.58)$, $p = .13$, two-tailed). In effect, there was no difference in starting IQ across the two groups.

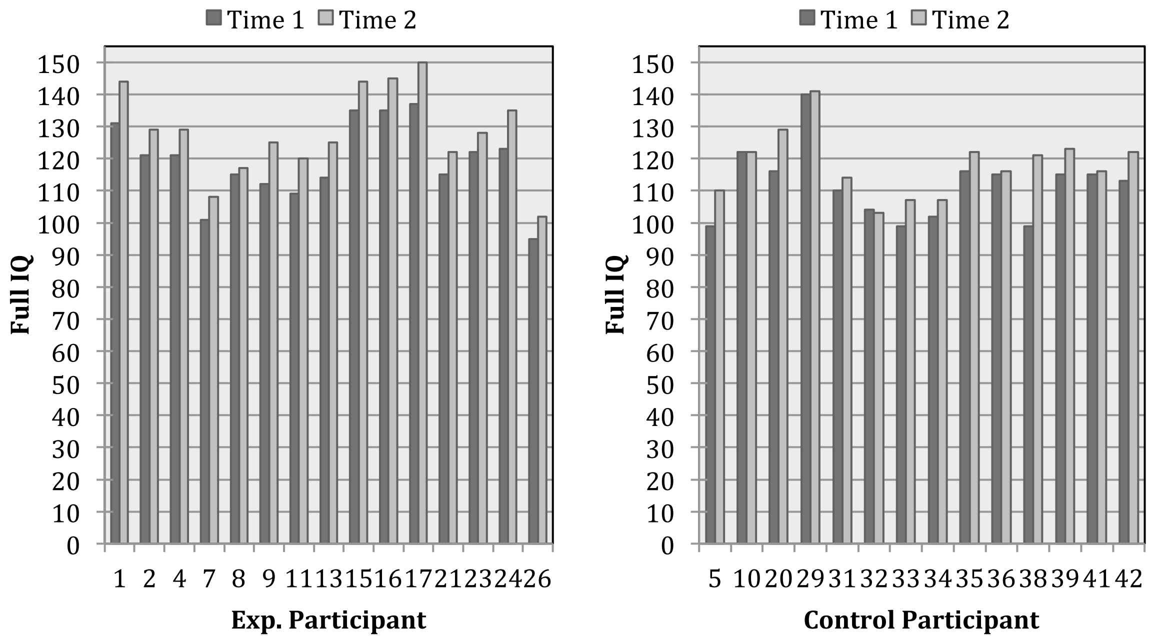


Figure 2. Bar charts comparing Full Scale IQ scores at baseline and follow-up for Experimental (left) and Control (right) participants. Experimental participants experienced a mean rise of 9.2 points, while Control participants had a mean rise of 6.3 points.

A mixed between-within ANOVA was conducted to assess the impact of relational skills training versus no intervention on participants' Full Scale IQ increase from baseline to follow-up as assessed by the WAIS-III. There was no significant interaction effect between intervention type and time, Wilks' Lambda = .92, $F(1,26)$ =

2.41, $p = .13$, partial eta squared = .09. There was a large main effect for time, Wilks' Lambda = .28, $F(1,26) = 67.56$, $p < .001$, partial eta squared = .72, with both groups showing an increase in Full IQ scores across the two time periods. The main effect comparing the two types of intervention was not significant, $F(1,26)= 3.67$, $p = .07$, partial eta squared = .12. Figure 2 outlines baseline and follow up Full Scale IQ scores for both experimental and control participants.

There was no significant interaction effect between intervention type and time for Verbal IQ, Wilks' Lambda = .94, $F(1,26) = 1.63$, $p = .21$, partial eta squared = .06. There was a large main effect for time, Wilks' Lambda = .59, $F(1,26) = 17.9$, $p < .001$, partial eta squared = .41, with both groups showing an increase in Verbal IQ scores across the two time periods. The main effect comparing the two types of intervention reached significance $F(1,26)= 5.12$, $p = .03$, partial eta squared = .17.

For Performance IQ, there was no significant interaction effect between intervention type and time, Wilks' Lambda = 1, $F(1,26) = .09$, $p = .76$, partial eta squared = .004. There was a large main effect for time, Wilks' Lambda = .34, $F(1,26) = 51.7$, $p < .001$, partial eta squared = .67, with both groups showing significant rises in Performance IQ scores. The main effect comparing the two types of intervention was not significant, $F(1,26)= .54$, $p = .47$, partial eta squared = .02, indicating that the effect of the training program on Performance IQ was not significantly greater than that seen in the no intervention group.

2.3.1.2 IQ subindices

For Verbal Comprehension scores, there was a significant interaction effect between intervention type and time, Wilks' Lambda = .83, $F(1,26) = 5.19$, $p = .03$, partial eta squared = .17. There was a large main effect for time, Wilks' Lambda = .62,

$F(1,26) = 15.98$, $p < .001$, partial eta squared = .38, with both groups showing an increase in Verbal Comprehension scores across the two time periods. The main effect comparing the two types of intervention was approaching significance, $F(1,26) = 3.78$, $p = .06$, partial eta squared = .13. The presence of an interaction effect for this IQ subscales would suggest that the relational skills training program exerted a significantly greater effect on Verbal Comprehension scores increases compared to no intervention.

There was no significant interaction effect between intervention type and time for Working Memory, Wilks' Lambda = .99, $F(1,26) = .34$, $p = .57$, partial eta squared = .01. There was no main effect for time, Wilks' Lambda = 1, $F(1,26) = .01$, $p = .91$, partial eta squared = .001, as neither group demonstrated a significant change in scores at follow-up. The main effect comparing the two types of intervention was not significant, $F(1,26) = .84$, $p = .37$, partial eta squared = .03.

For Perceptual Organisation scores, there was no significant interaction effect between intervention type and time, Wilks' Lambda = 1, $F(1,26) = .03$, $p = .87$, partial eta squared = .001. There was a large main effect for time, Wilks' Lambda = .3, $F(1,26) = 62.12$, $p < .001$, partial eta squared = .71, with both groups showing an increase in Perceptual Organisation scores across the two time periods. The main effect comparing the two types of intervention was not significant, $F(1,26) = .59$, $p = .45$, partial eta squared = .02.

Regarding Processing Speed scores, There was no significant interaction effect between intervention type and time, Wilks' Lambda = 1, $F(1,26) = .06$, $p = .8$, partial eta squared = .002. There was a large main effect for time, Wilks' Lambda = .53, $F(1,26) = 22.7$, $p < .001$, partial eta squared = .47, with both groups showing an increase in Processing Speed scores across the two time periods. The main effect comparing the two types of intervention was not significant, $F(1,26) = .2$, $p = .66$, partial eta squared = .01.

2.3.2 Post-hoc analyses of IQ change

A series of post-hoc paired samples *t*-tests were then computed to further investigate the change in WAIS-III scores and subscores separately for both groups. This was done simply to ask the question, did the IQ rise within each group rise significantly from pre- to post-intervention, considered on its own and in isolation from other known variables? Answering this question may help to clarify the source of the lack of interaction between IQ rises across time and treatment condition. That is, while the IQs of both groups rose significantly across time as a whole, the effect of the intervention was not significant enough to increase the IQs of the experimental group by a further significant degree. While the relevant ANOVA outcome stands, it is not yet clear if considered alone, both groups experienced significant IQ rises. Paired samples *t*-tests were computed for all seven IQ indices (Full, Verbal & Performance) and subindices (Verbal Comprehension, Working Memory, Perceptual Organisation & Processing Speed). For these 14 pre- to post-intervention analyses, alpha level of $p < .004$ (two-tailed) was set in line with Bonferroni procedures.

2.3.2.1 IQ indices

For the control group, there was a statistically significant increase in Full IQ scores from Time 1 (M = 111.8, SD = 11.16) to Time 2 (M = 118.1, SD = 10), $t(13) = -3.75$, $p = .002$ (one-tailed). Scores for Verbal IQ did not change significantly from Time 1 (M= 111.8, SD = 9.33) and Time 2 (M = 114.9, SD = 11.41), $t(13) = -1.74$, $p = .11$ (one-tailed). Scores for Performance IQ from also rose significantly from Time 1 (M = 109.8, SD = 14.36) to Time 2 (M = 120.3, SD = 14.25), $t(13) = -3.89$, $p = .002$ (one-tailed).

For the experimental group, Full IQ scores rose significantly from Time 1 (M = 118.9, SD = 12.76) to Time 2 (M = 128.1, SD = 14.3), $t(13) = -10.65$, $p < .001$ (one-tailed). There was a statistically significant increase in Verbal IQ scores from Time 1 (M = 120.7, SD = 13.29) and Time 2 (M = 126.6, SD = 14.71), $t(13) = -5.21$, $p <$ .001(one-tailed). Scores for Performance IQ from also rose significantly from Time 1 (M = 112.9, SD = 11.96) to Time 2 (M = 124.3, SD = 12.91), $t(13) = -8.09$, $p < .001$ (one-tailed).

2.3.2.2 IQ subindices

For the control group, changes for Verbal Comprehension between Time 1 (M = 114.1, SD = 7.78) and Time 2 (M = 115.8, SD = 10.35) $t(13) = -1.18$, $p = .26$ (one-tailed) were not statistically significant. Scores for Working Memory did not change significantly from Time 1 (M = 109.1, SD = 12.63) to Time 2 (M = 108.2, SD = 8.4), $t(13) = .36$, $p = .73$ (one-tailed). Scores for Perceptual Organisation increased significantly from Time 1 (M =110.5, SD = 12.53) to Time 2 (M = 122.4, SD = 14.73), $t(13) = -4.78$, $p <.001$. Processing Speed scores also rose significantly from Time 1 (M = 105.9, SD = 12.74) and Time 2 (M = 114.4, SD = 13.64), $t(13) = -3.46$, $p = .004$ (one-tailed).

For the experimental group, there was a significant increase in Verbal Comprehension scores between Time 1 (M = 120.1, SD = 13.14) and Time 2 (M = 126.1, SD = 13.11) $t(13) = -4.58$, $p = .001$ (one-tailed). Working Memory scores did not change significantly from Time 1 (M = 112.7, SD = 16.39) to Time 2 (M = 114, SD = 18.49), $t(13) = -.46$, $p = .66$ (one-tailed). Scores for Perceptual Organisation increased significantly from Time 1 (M =114.4, SD = 12.45) to Time 2 (M = 125.7, SD = 18.49), $t(13) = -7.17$, $p <.001$. Processing Speed scores did not rise significantly from

Time 1 (M = 104.1, SD = 13.07) and Time 2 (M = 111.8, SD = 15.91), $t(13) = -3.28$, $p = .006$ (one-tailed).

These results combined suggest that the relational skills intervention exerted a significantly greater impact than no intervention only on the Verbal Comprehension subscale of IQ. While the experimental group showed significant rises on this subscale, the control group demonstrated only marginal rises. The interaction effect found in the relevant ANOVA analyses would further support this conclusion. Analyses of interaction effects and post-intervention score changes suggest that in the current sample, the training program was not significantly more efficacious than no intervention in improving Full IQ, Verbal IQ or Performance scores or the three other IQ subindices. However, it must be noted that for Verbal IQ, mean rises seen in the experimental group almost doubled those found for the control group.

The lack of interaction between the effects of time and treatment condition may be due to the large practice or test-retest effects clearly apparent for the control group. The control group displayed an unexpectedly high and consistent level of increase on test scores, particularly for Performance IQ and Performance subscales. For example, control participants demonstrated greater rises in Perceptual Organisation and Processing Speed scores than the experimental group. There are a number of possible reasons for this, including participants starting IQ level and the nature of Performance subtests. Another issue worth noting relates to the timed aspects of these tests, which may be related to the size of practice effects. These issues will be further investigated in the second study, and discussed later in detail in the General Discussion section.

2.3.3 Analysis of RAI Score Change

A paired samples *t*-test was conducted to assess the efficacy of relational skills training in improving relational ability, as measured by the Relational Abilities Index.

There was statistically significant increase in RAI scores from Time 1 (M = 50.4, SD = 4.7) to Time 2 (M = 53.2, SD = 1.73), $t(12) = -2.64$, $p = .02$. The mean increase in RAI scores was 2.77 with a 95% confidence interval ranging from .48 to 5.06. The eta-squared statistic (.36) indicated a large effect size. Such significant and robust improvements in relational ability as measured by the RAI further support the effectiveness of the relational training program in targeting and improving relational sophistication. As Figure 3 shows, every participant who completed the training finished with an RAI score of at least 50 out of 55, indicating extremely high levels of relational skills post-intervention.

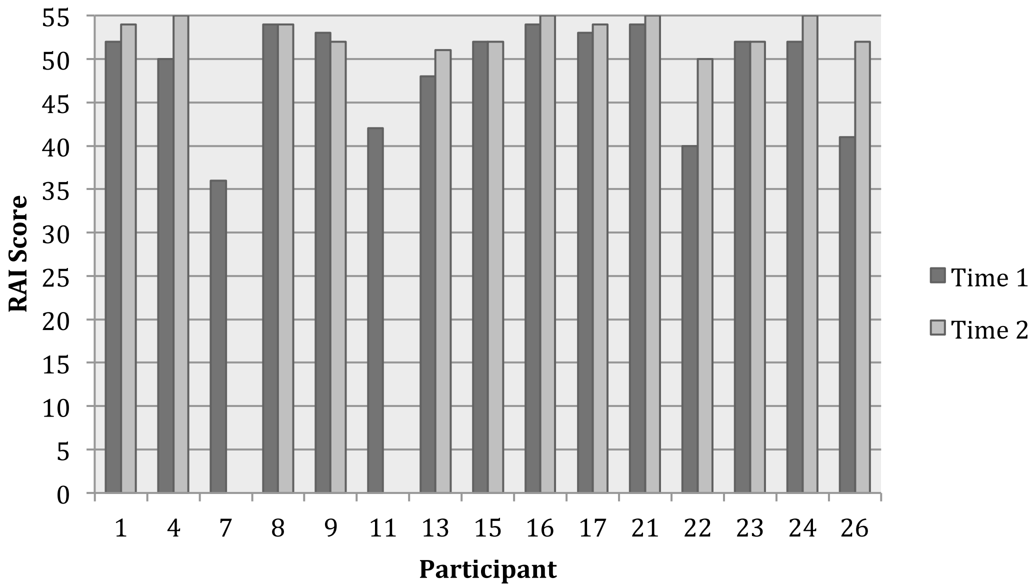


Figure 3. RAI Scores across both testing stages for the Experimental group. Bar chart shows that all completed participants displayed an RAI score of at least 50 out of 55 at follow up, with a mean increase of 2.8 points.

2.3.4 Summary

In summary, this study provides a number of interesting findings, as well as uncovering a number of important issues to be addressed in subsequent studies and future research. These relate to delineating in greater detail the precise functional relationship between IQ and relational skills and the limits of IQ increases among high IQ samples, as well as the increasing role played by practice effects at the higher end of the IQ spectrum. While the mean increase in RAI scores across the intervention period was smaller than those seen in previous analyses (Cassidy et al., 2016), it was none the less statistically significant. This can be viewed as a considerable finding given the fact that participants presented with an extremely high level of relational sophistication at baseline (mean RAI = 48.9). Due to this fact, the potential for increase may have been limited by the upper limit of scores, while participants also possibly stood less to "learn" from the training. This issue will be further investigated in Study 2.

In comparison to the control group, the training only exerted a significantly greater impact on Verbal Comprehension scores. Therefore, it can be concluded that the current results support the argument that the relational training procedure significantly improves performance on Verbal Comprehension test items but not on others. There was no interaction effect between experimental condition and time for the three main indices of IQ (Full, Verbal, & Performance) or the other three IQ subindices (Working Memory, Perceptual Organisation & Processing Speed). No doubt this is related to the unexpected and pronounced practice effects observed for the control group. The diminished effect of relational training, as witnessed in the current study, will be further investigated in Study 2. Furthermore, the findings of the current study will be further discussed in the General Discussion Chapter.

Chapter 3

Study 2: An Analysis of the relationship between IQ and scores on a Relational Abilities Index

3.1 Study 2

The current study aims to contribute to the burgeoning research repertoire that has previously outlined the relevance of relational responding to intellectual performance (Hayes, 1994; O'Hora et al., 2005, O'Toole et al., 2009; Smith et al., 2005). In addition, following the results of Study 1, Study 2 aims to further investigate possible reasons for the diminished effect found for relational training in the current sample. As such, Study 2 involved a correlational analysis of RAI scores and WAIS IQ scores, subscores and subtests in order to gain a more comprehensive understanding of relational ability's contribution to intellectual functioning. In a previous correlational analysis of relational responding and IQ scores, O'Hora et al. (2005) investigated the relationship between IQ scores as assessed by the WAIS-III and performance on a complex relational task using a sample of 72 college students. Following the relational task (see Section 1.7.1), participants were administered the vocabulary, arithmetic and digit-symbol coding subtests of the WAIS-III. Results indicated that successful completion of the relational task was positively correlated with vocabulary and arithmetic scores, but not digit-symbol coding. In a related analysis, O'Hora et al. (2008) found that completing a temporal relations task predicted performance on the Verbal Comprehension and Perceptual Organisation subindices of the WAIS-III. Completion of the temporal relations task was also associated with higher scores on the Vocabulary, Similarities, Information, Block Design and Symbol Search WAIS subtests. In a study of deictic relational responding, Gore et al. (2010) found that perspective-taking correlated strongly with verbal ability, Full IQ and Performance IQ.

Due to the close relationship between relational responding and IQ, correlation coefficients for the Relational Abilities Index employed by Cassidy et al. (2016) and the main indices, subindices and subscales of IQ will be investigated. As the Relational

Abilities Index probes for sophistication regarding Same/Opposite and More Than/Less Than relations, the current analysis will therefore assess the relevance of the relational frames of coordination, opposition, and comparison to intellectual performance. Furthermore, it is hoped that by investigating this relationship, the mechanisms underlying the IQ gains previously witnessed following relational skills training (Cassidy et al., 2008; 2016), may be better understood. Finally, it is hoped that the current correlational analysis will allow a closer and more rigorous inspection of the factors that may have contributed to the diminished effect of relational training found in Study 1, when compared to the findings of these previous studies.

3.2 Method

3.2.1 Participants

All participant WAIS and RAI data collected as part of Study 1 was repurposed in the current study to allow a correlational analysis of IQ, RAI and a collection of related statistics (age, number of incorrect responses, training duration etc.). All participants were currently attending third level education or had recently graduated (<1 year). Participants were informed at the outset that they should not volunteer if they had attended special education outside the mainstream schooling system, or have been diagnosed with a learning difficulty. Participants ranged in age from 18-44, with a mean age of 22 years. All were native English speakers, who, as expected, presented in the High Average to Superior range of intellectual function for Full IQ ($M = 118.1$, $SD = 13.04$), Verbal IQ ($M = 120.2$, $SD = 13.83$) and Performance IQ ($M = 111.5$, $SD = 12.2$). Mean Relational Ability Index score was 48.94 out of a possible 55.

3.3 Results

3.3.1 Correlational Analysis of RAI and IQ scores

In order to assess the importance of relational responding to intellectual performance, a correlational analysis was undertaken to investigate the relationship between baseline Relational Abilities Index scores and WAIS IQ indices, subindices and subtests at both baseline and follow-up. In addition, the relationship between pre-intervention and post-intervention scores RAI and IQ will be investigated. Finally, correlations between changes on these two metrics will be explored. For the purpose of clarity, the terms Time 1/pre-intervention and Time 2/post-intervention will be used to refer to baseline and follow-up assessment respectively.

3.3.1.1 RAI scores & IQ indices at baseline

Correlations between RAI scores and the three major IQ scales (Full, Verbal & Performance IQ) were computed. The relationship between RAI scores and the four IQ subscales, as well as the 13 IQ subtests will also be investigated subsequently. The alpha level of .05 was set for all correlational analyses for this study. At Time 1, Relational Ability Index scores were found to correlate strongly with Full IQ (r = .71, p = .002). Figure 4 represents a scatterplot outlining the distribution of these scores.

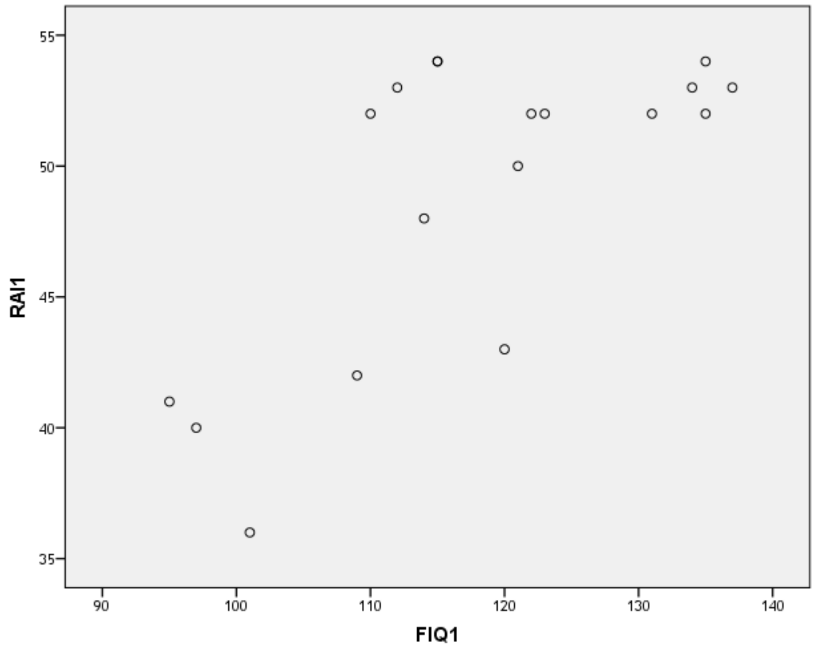


Figure 4. Scatterplot outlining the distribution of baseline RAI and Full IQ scores.

Relational Ability Index scores also showed strong correlations with Time 1 scores for Verbal IQ ($r = .76$, $p = .001$) and Performance IQ ($r = .53$, $p = .02$). Furthermore, baseline RAI scores also displayed high levels of correlation with baseline Full IQ percentile ($r =. 78$, $p < .001$), Verbal IQ percentile ($r = .87$, $p < .001$) and Performance IQ percentile ($r = .58$, $p = .01$) at baseline. Figure 5 displays scatterplots presenting a graphical representation of correlations between pre-intervention RAI scores and Verbal and Performance IQ respectively. RAI scores showed strong correlations with Total Raw IQ Scores ($r = .79$, $p = .001$) and Total Standardised IQ Scores ($r = .69$, $p = .004$).

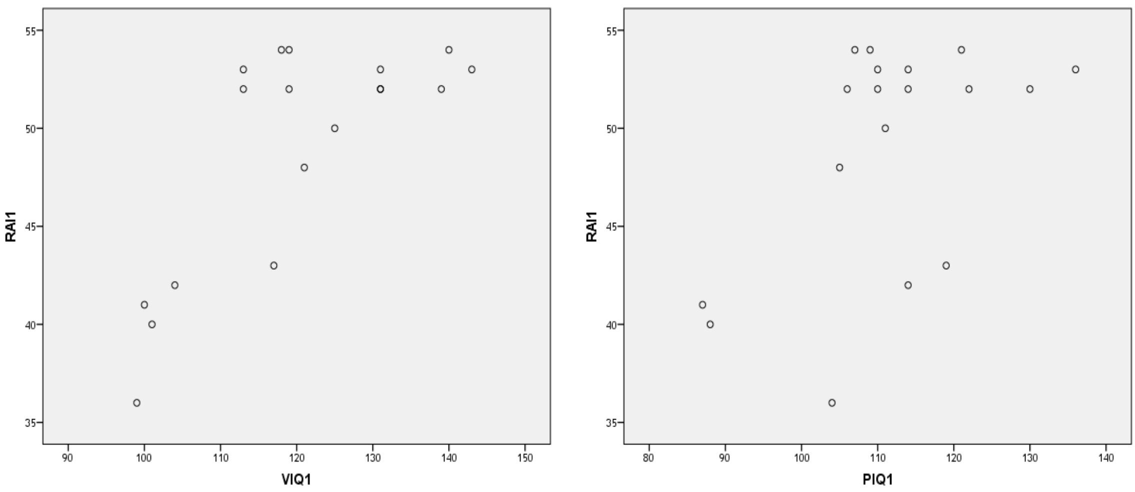


Figure 5. Scatterplot outlining the distribution of RAI and Verbal IQ scores (left) and Performance IQ scores (right).

3.3.1.2 RAI scores & IQ indices at follow-up

Following the results of Study 1, the current study analysed correlations between post-intervention RAI and IQ scores. This was carried out to assess the relationship between relational skill and IQ following intervention. In addition, correlations between baseline RAI and follow-up IQ scores were also investigated. It was hoped that this investigation may lead to greater understanding of the impact baseline relational ability may exert on IQ scores. For the purpose of clarity, Time 1 hereafter refers to baseline and Time 2 refers to follow-up assessment.

Relational Abilities Index scores at Time 1 correlated strongly with RAI scores at Time 2 (r = .73, p =.002). Relational Abilities Index scores at Time 1 were also found to correlate strongly with Full IQ at Time 2 (r =.72, p =.002), with Verbal IQ at Time 2 (r =.7, p = .002) and with Performance IQ scores at Time 2 (r =.54, p = .02). RAI scores at Time 1 correlated strongly with Full IQ percentile at Time 2 (r = .79, p < .001), Verbal IQ percentile at Time 2 (r = .83, p < .001) and Performance IQ at Time 2

($r = .67$, $p = .005$). RAI at Time 1 also bore strong correlations with Total Raw IQ Scores at Time 2 ($r = .7$, $p = .004$) and Total Standardised IQ Scores at Time 2 ($r = .66$, $p = .008$). This finding indicates that, despite the IQ increases seen following intervention, post-intervention IQ scores were still functionally related to baseline RAI scores.

The level of correlation found between follow-up RAI scores and follow-up IQ scores were not as strong as those witnessed between these measures at baseline. RAI scores at Time 2 correlated significantly with Verbal IQ Time 2 ($r = .56$, $p = .04$). Follow-up RAI scores showed weak to moderate relationships, but did not correlate significantly, with scores for Full IQ at Time 2 ($r = .48$, $p = .09$) and Performance IQ at Time 2 ($r = .26$, $p = .37$). In terms of percentile scores, RAI scores at Time 2 correlated significantly with Verbal IQ percentile at Time 2 ($r = .58$, $p = .02$) only. There was no significant correlation between post-intervention RAI scores and post-intervention percentile scores for Full IQ ($r = .43$, $p = .12$), Verbal IQ ($r = .44$, $p = .12$) or Performance IQ ($r = .33$, $p = .25$). However, correlation coefficients once again showed moderate relationships between follow-up RAI scores and each of these IQ indices. Furthermore, RAI scores at Time 2 did not correlate significantly with Total Raw Score at Time 2 ($r = .23$, $p = .45$) or Total Standardised Score at Time 2 ($r = .44$, $p = .13$). These findings would suggest that there is a weaker relationship between RAI scores and IQ scores after training than seen between these measures at baseline. Therefore, it may be the case that these two metrics are less dependent on each other following training. These results propose that post-intervention relational ability may be "unhinged" from intellectual performance to a certain extent. It may be the case that at higher levels of performance, the interdependent relationship between relational ability (as measured by the RAI) and intellectual performance is less pronounced.

3.3.1.3 Changes in RAI scores & IQ indices

Correlations between changes in RAI and IQ scores were also studied. In addition, correlations were also sought to delineate the relationship between baseline RAI and IQ scores and subsequent changes on these metrics. Finally, the relationship between RAI and IQ changes and a number of statistics related to training were also computed. Changes in RAI scores correlated negatively with RAI scores at Time 1 (r = -.9, $p < .001$) and percentage of correct responses across the entire training program (r = -.77, $p = .006$). Changes in RAI scores did not correlate significantly with Full IQ change ($r = .2$, $p = .58$), RAI at Time 2 ($r = -.3$, $p = .37$), days required to complete training ($r = -.19$, $p = .57$) or total trials completed ($r = .44$, $p = .2$). Correlation coefficients show weak to moderate relationships between RAI and these variables, which may reach significance in a larger sample. This trend in correlation, allied with the high level of relational skill observed across the board at baseline, may suggest that due to tight concentration of RAI scores towards the "ceiling" of possible scores, there was relatively little room for improvement for such participants. While the upper IQ "limit" afforded participants space to increase scores (therefore resulting in greater variance in scores), the RAI scores did not (resulting in lower variance in scores). Table 2 outlines correlation coefficients for post-intervention IQ scores and subindex scores with RAI scores at both Time 1 and Time 2.

As indicated by the current findings, the nature of the relationship between intellectual performance and relational ability is complex, particularly for this sample. This topic will be expanded upon subsequently. Alongside analyses of the main IQ scales, correlations between RAI scores and IQ subscales and subtests were also studied to gain a more specific understanding of what aspects of intellectual functioning relational ability may contribute to.

Table 2

Correlations for post-intervention IQ scores with baseline and follow-up RAI scores

IQ measure	Baseline RAI	Follow-up RAI
Full IQ	.72**	.48
Verbal IQ	.7**	.56*
Verbal Comprehension	.66**	.56*
Working Memory	.6*	.53
Performance IQ	.54*	.26
Perceptual Organisation	.44	.37
Processing Speed	.6	.53

* indicates correlation is significant at the 0.05 level (2-tailed)
** indicates correlation is significant at the 0.01 level (2-tailed)

3.3.1.4 RAI scores & IQ subindices

Correlations between RAI scores and the two verbal subscales, Verbal Comprehension and Working Memory, and the two Performance subtests, Perceptual Organisation and Processing Speed were also analysed. The relationship between RAI scores and IQ subindex scores were computed in order to gain a more specific understanding of the relevance of relational ability to intellectual performance. Correlations between baseline RAI scores and follow-up IQ subindex scores were also studied. This was done in order to assess the impact of relational training on the relationship between RAI scores and IQ subscores. Finally, correlation coefficients were investigated between follow-up RAI and IQ subindex scores to ascertain the nature of this relationship following relational skills training.

At Time 1, RAI scores showed a high degree of correlation with baseline Verbal Comprehension scores ($r = .55$, $p = .02$) and percentile scores ($r = .55$, $p = .02$), as well as with baseline Working Memory scores ($r = .66$, $p = .003$) and percentile scores ($r = .73$, $p = .001$). Baseline RAI scores correlated well with baseline Perceptual Organisation scores ($r = .48$, $p = .05$) and percentile scores ($r = .52$, $p = .03$). Correlations between RAI scores and Processing Speed scores ($r = .42$, $p = .08$) and percentile ranks ($r = .33$, $p = .07$) at Time 1 failed to reach significance despite correlation coefficients that may indicate a moderate relationship between these variables.

RAI scores at Time 1 showed a high degree of correlation with Verbal Comprehension scores at Time 2 ($r = .66$, $p = .005$), as well as with Working Memory at Time 2 ($r = .6$, $p = .01$). Correlations between RAI scores at Time 1 showed a moderate relationship but did not reach statistical significance with Perceptual Organisation at Time 2 ($r = .44$, $p = .09$), or with Processing Speed at Time 2 ($r = .35$, $p = .18$). In terms of follow-up scores, RAI scores at Time 2 displayed a significant correlation with Verbal Comprehension at Time 2 ($r = .56$, $p = .005$). Correlations between RAI at Time 2 did not reach statistical significance for post-intervention scores for Working Memory ($r = .53$, $p = .051$), Perceptual Organisation ($r = .37$, $p = .19$) or Processing Speed ($r = -.006$, $p = .98$). It can be stated however, that a significant correlation may be found for RAI and Working Memory scores at Time 2 in a larger sample, as results this relationship was approaching significance.

3.3.1.5 RAI scores & IQ subtests

Correlations between RAI scores and standardised scores on each of the 13 individual subtests were also analysed. These comprised seven Verbal subtests (Vocabulary, Similarities, Arithmetic, Digit Span, Information, Comprehension & Letter-Number

Sequencing) and six Performance subtests (Picture Completion, Digit-Symbol Coding, Block Design, Matrix Reasoning, Picture Arrangement & Symbol Search). Regarding Verbal subtests at Time 1, strong significant correlations were found between baseline RAI and Vocabulary scores ($r = .68$, $p = .004$), Similarities ($r = .56$, $p = .02$), Arithmetic ($r = .7$, $p = .003$), Digit Span ($r = .55$, $p = .03$) Information ($r = .65$, $p = .03$), and Comprehension ($r = .68$, $p = .004$). RAI scores did not correlate significantly with Letter-Number Sequencing ($r = 45$, $p = .08$). However once again the correlation coefficient in this case indicates a moderate relationship between these two variables. Table 3 displays the correlation coefficients and respective significance levels for RAI correlations with each of these subtests.

Table 3

RAI Correlations with Verbal subtest scores at Time 1

Subtest	Correlation coefficient	Significance level
Vocabulary	.68	.004**
Similarities	.56	.02*
Arithmetic	.7	.03*
Digit Span	.55	.03*
Information	.65	.03*
Comprehension	.68	.004**
Letter Number Sequencing	.45	.08

* indicates correlation is significant at the 0.05 level (2-tailed)
** indicates correlation is significant at the 0.01 level (2-tailed)

In terms of Performance subtests at Time 1, baseline RAI scores displayed a significant correlation with standardised scores for Block Design ($r = .5$, $p = .048$). RAI scores were not significantly correlated with scores for Picture Completion ($r = -.1$, $p = .72$), Digit-Symbol Coding ($r = .38$, $p = .14$), Matrix Reasoning ($r = .37$, $p = .16$), Picture Arrangement ($r = .3$, $p = .26$) or Symbol Search ($r = .39$, $p = .13$). While these correlations did not reach statistical significance, scores for each subtest (with the exception of Picture Completion) indicated a moderate relationship with RAI scores. These correlation coefficients and relevant significance levels are outlined in Table 4.

Table 4

RAI Correlations with Performance subtest scores at Time 1

Subtest	Correlation coefficient	Significance level.
Picture Completion	-.1	.72
Digit Symbol Coding	.38	.14
Block Design	.5	.048*
Matrix Reasoning	.37	.16
Picture Arrangement	.3	.26
Symbol Search	.39	.13

* indicates correlation is significant at the 0.05 level (2-tailed)

Strong significant correlations were found for RAI at Time 1 and follow-up scores for Vocabulary ($r = .66$, $p = .01$), Similarities ($r = .59$, $p = .02$), Arithmetic ($r = .7$, $p = .004$) and Digit Span ($r = .52$, $p = .05$). RAI scores at Time 1 showed moderate relationships, but did not correlate significantly with post-intervention scores for Information ($r = .49$, $p = .06$) and Letter-Number Sequencing ($r = .39$, $p = .15$). Baseline RAI scores did not correlate significantly with follow-up Comprehension scores ($r = .12$, $p = .67$). For RAI at Time 2, there was a significant correlation for Vocabulary at Time 2 ($r = .58$, $p = .04$) only. Correlations did not reach significance for RAI at Time 2 and follow-up scores for Similarities ($r = .41$, $p = .16$), Arithmetic ($r = .24$, $p =.44$), Digit Span ($r = .42$, $p = .15$), Information ($r = .52$, $p =.06$), Comprehension

(r = .33, p = .27) or Letter Number Sequencing (r = .13, p = .68). However, while these correlations did not reach significance, their respective correlation coefficients indicated weak to moderate relationships.

RAI scores at Time 1 did not correlate significantly with post-intervention scores for Picture Completion (r = .1, p = .73), Digit-Symbol Coding (r = .3, p =.27), Block Design (r =.35, p =.21), Matrix Reasoning (r = .13, p =.65), Picture Arrangement (r = .34, p =.21) or Symbol Search (r = .36, p =.19). Regarding RAI at Time 2, correlations failed to reach significance for all Performance subtests at follow-up: Picture Completion (r = 24, p = .44), Digit-Symbol Coding (r = -.23, p =.46), Block Design (r =.17, p =.58), Matrix Reasoning (r = .16, p =.6), Picture Arrangement (r = .25, p =.41) or Symbol Search (r = .07, p =.82).

As found in Study 1, RAI scores at Time 2 displayed a similar trend towards higher levels of correlations with Verbal scales and subtests. This relationship however, was weaker than those found for RAI scores at baseline. Nevertheless, despite this lower degree of correlation, there remained a considerable tendency for the RAI to correlate best with the verbal aspects of the WAIS-III. Furthermore, despite failing to reach statistical significance, the current correlational analysis indicated the presence of numerous moderate relationships shared between RAI scores and Performance subtest scores.

3.3.2 Potential Boundary Conditions affecting IQ rises

As the relational skills intervention employed in Study 1 was less effective in raising IQ when compared to previous analyses (see Cassidy et al. 2010, 2016), potential ceiling effects and covariates were investigated in order to gain an understanding of the factors which may have contributed to this reduced effectiveness.

3.3.2.1 Starting IQ

Correlational analyses indicated that baseline Full IQ did not predict changes in scores for Full IQ (r = -.04, p = .82), Verbal IQ (r = .19, p = .32), Performance IQ (r = -.23, p = .24) or RAI change following intervention (r = -55, p = .08). However, correlations between baseline Full IQ and RAI change were approaching statistical significance with correlation coefficients indicating a strong inverse relationship. Baseline Verbal IQ scores did not predict changes in Verbal IQ or at follow-up (r = .23, p = .25). Similarly, Performance IQ scores at baseline did not predict change in Performance IQ at follow-up (r = -.25, p = .2). For the experimental group, baseline Full IQ did not predict post-intervention RAI change (r = -.04, p = .82). These findings therefore, suggest that starting IQ appears not have a delimiting impact on improvements post-intervention, as there appears to be very little relation between baseline IQ measures and the changes in both IQ and RAI scores for the most part. That said, it must be noted that the inverse relation between starting Full IQ and RAI change warrants further investigation.

3.3.2.2 Age

No significant correlations were found for age and Full IQ change from pre- to post-intervention (r = .04, p = .84) or RAI change following intervention (r = -.37, p = .26). These results would indicate that age did not significantly affect changes in Full IQ or RAI following the training. However, it must be noted that the low level of variability in age (M = 21.8, SD = 3.5) lessen the likelihood of a significant correlation. As such, the current findings offer only limited utility to the study of age as a confound of relational skills intervention. These findings will be placed into context with previous results gleaned from child samples in the General Discussion chapter.

3.3.3 Changes in timed vs. untimed subtest scores

Previous investigations (Basso, Bornstein & Lang, 1999; Dodrill & Troupin, 1975; Rapport, Brines, Thieson & Axelrod, 1997; Sattler, 2001) on the impact of multiple administrations of IQ assessment have indicated that IQ subtests that include a timed element may be more susceptible to practice effects that those that do not. It has been suggested that timed subtests are more vulnerable to memorisation and the development of more effective problem-solving strategies (Rapport et al, 1997). Due to the pronounced IQ increases found for the control group, increases in total standardised scores for timed and untimed subtests were investigated for both groups. The thirteen subtests were therefore divided into timed subtests which rewarded speed of response (Block Design, Arithmetic), and/or involved a time limit (Digit-Symbol Coding, Arithmetic, Picture Arrangement & Symbol Search) and those that did not have a timed element (Vocabulary, Similarities, Matrix Reasoning, Digit Span, Information, Comprehension & Letter-Number Sequencing). While there are exceptions (Arithmetic & Matrix Reasoning), this segregation generally adheres to the Performance (timed) and Verbal (untimed) separation.

A mixed between-within ANOVA was computed to assess the impact of the training intervention versus no intervention on total score on timed subtests. There was no significant interaction effect between intervention type and time, Wilks' Lambda = .98, $F(1,26) = .53$, $p = .47$, partial eta squared = .02. There was a large main effect for time, Wilks' Lambda = .2, $F(1,26) = 107.36$, $p < .001$, partial eta squared = .81, with both groups showing an increase in timed subtest scores across the two time periods.

The main effect comparing the two types of intervention was not significant, $F(1,26)=$ .11, $p = .75$, partial eta squared = .004.

A mixed between-within ANOVA was run to assess the impact of the training intervention versus no intervention on total untimed subtest scores. There was no significant interaction effect between intervention type and time, Wilks' Lambda = .89, $F(1,26) = 3.38$, $p = .08$, partial eta squared = .12, although this effect was approaching significance. There was a large main effect for time, Wilks' Lambda = .78, $F(1,26) =$ 7.4, $p = .01$, partial eta squared = .22. The main effect comparing the two types of intervention was significant, $F(1,26) = 5.45$, $p = .03$, partial eta squared = .17.

For baseline timed scores, there was no significant difference between the control (M = 67, SD = 10.2) and experimental groups (M = 64.2, SD = 20.2, $t(.47)$, $p =$ .65). The control group showed a statistically significant increase in timed subtest scores from Time 1 (M = 67, SD = 10.2) to Time 2 (M = 76.8, SD = 9.1), $t(13) = -7.03$, $p < .001$ (two-tailed). The experimental group also showed significant rises in timed subtests scores from Time 1 (M = 68.8, SD = 9.98) to Time 2 (M = 76.8, SD = 9.85), $t(13) = -7.03$, $p < .001$. For follow-up timed scores, there was no significant difference between the control (M = 76.2, SD = 9.1) and experimental groups (M = 76.8, SD = 9.9), $t(-.16)$, $p = .88$.

For untimed scores at baseline, the picture was quite different. Differences between the control (M = 86.1, SD = 11) and experimental groups (M = 94.2, SD = 11.64, $t(-1.88)$, $p = .07$, at baseline were not significant. There was no significant rise in scores for the control group from Time 1 (M = 86.1, SD = 11) to Time 2 (M = 87, SD = 11.5), $t(13) = -.49$, $p = .63$ (two-tailed). On the other hand, the experimental group showed significant rises in untimed subtests scores from Time 1 (M = 94.2, SD = 11.64) to Time 2 (M = 98.6, SD = 11.7), $t(13) = -5.1$, $p < .001$. For follow-up timed scores, there was a significant different between the control (M = 87, SD = 11.5) and experimental groups (M = 98.6, SD = 911.7, *t*(2.66), *p* = 01.

As Figure 6 illustrates, there is a discrepancy between mean increases in timed subtest scores (9.8 points) versus untimed subtest scores (0.9 points) for control participants. This discrepancy may provide a tentative explanation for the large increases in IQ scores over time for control participants. In effect, the largest contributor to increases in control participant scores was undoubtedly improvement in timed subtest performance. Given the mean Full IQ of the control group, an increase of 10 standardised points equates to roughly 7 points, which is consistent with the mean increase found for this group. As scores on timed subtests remained virtually unchanged (mean increase – 0.9 points), it could be suggested that the overwhelming majority of the practice effect witnessed for control participants may be attributable to improved performance on timed aspects of the IQ assessment. In comparison, the experimental group showed significant rises on both timed and untimed subtest scores.

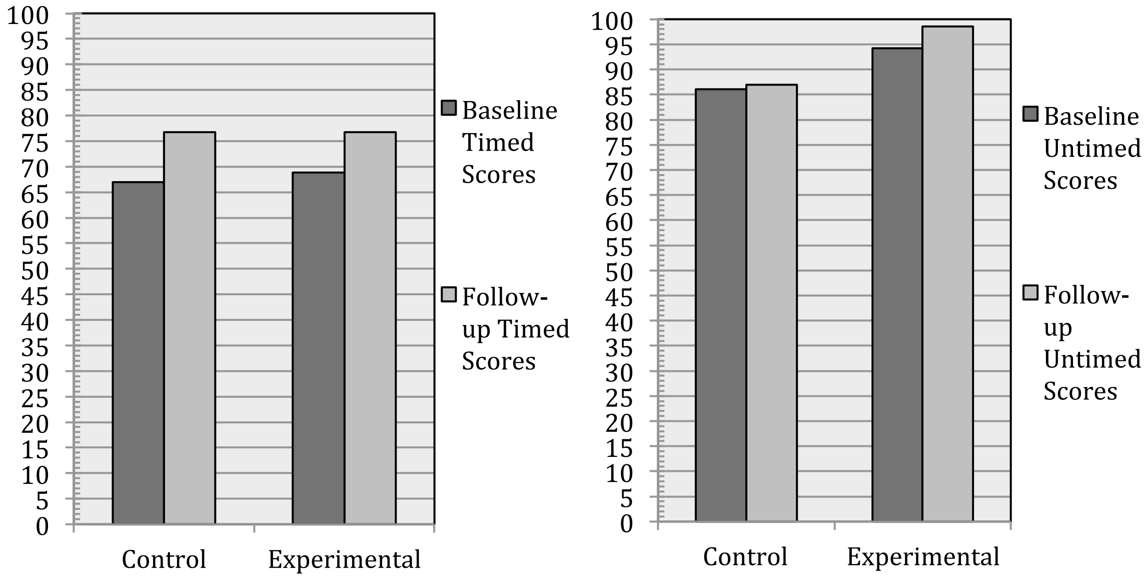


Figure 6. Baseline and follow-up scores on timed (left) and untimed (right) subtests for both groups. Experimental group rises in both timed (M = 8 points) and untimed (M = 4.4 points) subtest scores reached statistical significance. While control participant displayed large significant increases in timed subtests scores (M = 9.8 points), increases in scores for untimed subtests (M = 0.9 points) were not significant.

3.3.4 Summary

In summary, Relational Ability Index scores show strong correlations with the three main indices of IQ: Full, Verbal and Performance IQ. Such a finding would suggest that there is considerable overlap between these IQ constructs and the relational skills repertoire. The RAI also demonstrated strong correlations with three of the four IQ subindices: Verbal Comprehension, Working Memory and Perceptual Organisation. Due to this consistently high level of correlation with IQ measures, it is reasonable to suggest that the RAI shows a substantial validity in assessing many of the skills that are deemed intelligent. More significantly, due to the nature of the tasks involved in the RAI, the current findings further support the fundamental importance, or at least the significant contribution, of relational skills to intelligence. Of course, the functional direction of causality between IQ and relational skills is not implied by these correlations, but the presence of these correlations supports the existence of the functional relationship as hypothesised in the Introduction section.

The nature of the functional relationship between IQ and relational skills notwithstanding, these results would support the use of the RAI as a more or less valid proxy measure of intelligence. Indeed, correlational analyses are commonly used to implicate the potential efficacy of a novel measure of intelligence and in this regard, the RAI has performed well. This degree of correlation between the RAI and Full IQ (r = .7) is comparable to many other traditionally accepted proxy measures of intelligence. The correlation coefficients between RAI and IQ scores found in this study outrank many currently validated proxy measures and short form tests of IQ, and would appear to support the tentative suggestion that IQ scores overlap functionally with relational skill fluency test scores. This will be discussed in detail in the General Discussion chapter.

Finally, the notably more frequent and significant correlations between RAI scores and the Verbal scales and subtests of the WAIS suggest that relational responding, at least as assessed by the RAI, are germane to linguistic performance and possibly underlie language processes. As outlined previously, the importance of relational responding for language performance is well documented. The current findings would suggest that this significant contribution extends to even high-level verbal ability, as the RAI correlated significantly with all but one Verbal subtest. However, for Performance subtests, only one correlation reached significance (Block Design), indicating that the RAI may not accurately assess the skills relevant to non-verbal IQ items. While only one Performance subtest appeared to be significantly correlated with relational ability scores, combined scores of these subtests (i.e. Performance IQ) correlated strongly with RAI scores. To complement this finding, correlation coefficients indicate weak to moderate relationships between RAI scores and Performance subtest performance. This would indicate that relational ability displays a relationship with Performance IQ, but that the relationship is not significant at the individual subtest level for the most part, at least in a sample of this size. Thus in general terms, the current findings would appear to further support the proposal that there exists a close relationship between relational responding and intellectual performance but it has not emerged that this correlation applies at every level of the IQ test procedure.

Correlational analyses further supported the close relationship proposed between relational skill and IQ, with the RAI scores continuing to display strong correlations with the Verbal aspects of the WAIS in particular. However, the correlations for follow-up RAI and IQ scores were less consistent, which may possibly be a statistical artefact of a discrepancy in variances for RAI and IQ measures that was larger than seen in baseline assessments. These issues will be expanded upon in the General Discussion chapter. In addition, age did not correlate with changes in IQ or RAI scores, indicating that age may be unrelated to the rise in IQ scores following relational skills intervention. However, the low degree of variance in age across the participants may well limit this conclusion derived from weak correlational coefficients. Furthermore, it must also be noted that the IQ increases reported in the current study are much lower than those reported previously in child populations (Cassidy et al. 2011, 2016), indicating that age may play a role in the magnitude of IQ increases.

Full IQ scores at baseline did not correlate significantly with rises in any of the three main IQ indices, or with changes in RAI scores, indicating that a high IQ does not preclude the possibility of improvement, as evinced by previous findings. However, it is important to note once again that previous analyses (Cassidy et al., 2011; 2016), found considerably larger IQ increases among average IQ cohorts. Therefore, as witnessed in the current study, baseline RAI may be more indicative of subsequent improvements gained from training. Correlational analysis on this effect indicated that higher baseline RAI scores were associated with lower levels of increase. The relative influence of baseline IQ and relational ability on IQ improvements will be discussed in further detail in the General Discussion.

A key determinant of the failure to find a significant effect of relational training on most IQ measures in Study 1 was due to the pronounced practice effect displayed by control participants. Analyses of these practice effects indicate that the dominant contribution to increases was improvement in performance on IQ subtests that involved a timed element. While experimental participants showed significant rises in scores for both timed and untimed subtests, control participants only displayed notable improvements for timed subtests. Untimed subtest scores for control participants remained virtually unchanged, suggesting that the intervention may have a unique ability to improve performance on untimed subtests above and beyond practice effects. This is the first finding of such an effect, and therefore represents an important contribution to research into relational skills intervention. Results from the current study will be further expanded upon discussed an placed into the wider research context in the General Discussion chapter

Chapter 4

General Discussion

4.1 Introduction

The current thesis aimed to investigate (1), the efficacy of a relational skills training program in improving adult intellectual performance as assessed by the WAIS-III and (2) the relationship between relational ability and WAIS-III IQ scales and subscales. The first study failed to find a significant effect for the training program for Full, Verbal or Performance IQ. Further analyses indicate that there was a significant effect of the intervention on Verbal Comprehension scores. It appeared that the intervention resulted in qualitatively different increases in subtests scores, as only the experimental group experienced significant rises in untimed IQ subtest scores. The second of these studies found a high degree of correlation between relational skill and intelligence, with significant correlations found for Full, Verbal and Performance IQ. There were also significant correlations between RAI scores and Verbal Comprehension, Working Memory and Perceptual Organisation. Age and starting IQ did not significantly correlate with RAI or Full IQ increases. These results will be now be placed into the wider research context and discussed in further detail.

4.2 Relational Ability Index

4.2.1 Correlations with Baseline IQ and Relational Ability Index scores

Relational Ability Index scores displayed impressive levels of correlation with the three main IQ indices: Full, Verbal with Performance IQ scores. RAI scores also displayed significant correlations with percentiles for each of these indices. This positive trend continued into the subscale scores, with high levels of correlation found between RAI scores and both Verbal subscales, with strong degrees of linearity with Verbal Comprehension scores and percentile scores and Working Memory scores and percentile scores. Finally, RAI scores correlated strongly with Perceptual Organisation scores and percentile scores, but not with scores or percentile ranks for Processing Speed.

In terms of Verbal subtests, the RAI correlated significantly for scores for 6 of the 7 subtests: Vocabulary, Similarities, Arithmetic, Digit Span, Information and Comprehension. RAI scores did not, however, correlate significantly with Letter-Number Sequencing. Regarding Performance subtests, the RAI did not show such a linear relationship, only correlating significantly with standardised scores for Block Design. RAI scores were not significantly correlated with scores for Picture Completion, Digit-Symbol Coding, Matrix Reasoning, Picture Arrangement or Symbol Search.

These findings would therefore suggest a high level of covariance between relational ability and intellectual performance, supporting previous assertions of the importance of relational ability in determining intelligence (Andrews & Halford, 1998; Barnes-Holmes et al., 2010; Cattell, 1971; Dixon et al., 2014; Gentner & Loewenstein, 2002; Gore et al., 2010; Moran et al., 2010, 2014; O'Hora et al., 2005; 2008; O'Toole & Barnes-Holmes, 2009; Stewart et al., 2013).

4.2.2 Comparison with other IQ measures and proxy measures

Relational Ability Index Scores displayed a very strong positive correlation with Full IQ as measured by the WAIS-III. Due to the strength of this relationship, it follows that there is considerable overlap in performance on relational tasks and measures of intelligence. The correlation found between RAI scores and Full IQ is at a level comparable to proxy measures of intelligence. For example, the coefficient reported in the current study for Full IQ was at a similar level to those cited between the WAIS-III and Ravens Standard Progressive Matrices (.65, Fletcher & Hattie, 2011; .64, Wechsler, 1997; .49 - .79, Silva, 2008), Wechsler Individual Achievement Test (.53 - .81, Silva, 2008), Woodcock Johnson Test of Cognitive Ability (.71, Cheramie, Stafford, Boysen, Moore & Prade, 2012; .82, Metz, 2005) and the General Ability Measure for Adults (.8, Martin, Donders & Thompson, 2000; .75, Naglieiri & Bardos,1997). The correlation coefficient found for the RAI in the current study is also comparable to figures reporting the correlation between WASI Verbal IQ and WAIS Full IQ (.75, Axelrod, 2002). To put these figures into perspective, correlation coefficients between the WAIS and the Stanford-Binet Intelligence Scale (Fourth Edition, Thorndike, Hagan & Sattler, 1986a), perhaps the most widely accepted IQ measure beside Wechsler assessments, generally fall somewhere between .77 and .89 (Groth-Mamat, 2003; Kamphaus, 2005; Silva, 2008; Zimmerman & Woo-Sam, 1973). The Kaufman Brief Intelligence Test (Kaufman & Kaufman, 2004) has also produced a correlation coefficient with the WAIS-III above .8 (Cohen, Cohen, West & Aiken, 2003). The strength of the correlation between RAI and IQ scores, when compared to those between various other cognitive assessments, would therefore suggest that the current RAI assessment represents considerable predictive validity regarding general intellectual performance.

In addition, the correlation coefficient between RAI and IQ scores in the current study is on par with those reported in support of many short form and proxy assessments of intelligence. In a comprehensive study of proxy and short-form IQ measures, Spinks et al. (2009), compared scores on the WAIS-III with widely used proxy IQ measures, demographic formulae and various abbreviated IQ assessments such as the Oklahoma Premorbid Intelligence Estimate (OPIE; Schoenberg, Scott, Duff & Adams, 2002) and Ward – 7 Subtest (Ward, 1990). The OPIE and Ward utilise various selections of Wechsler subtests, along with demographic data in order to estimate Full IQ, in a similar manner to the Wechsler Abbreviated Scales of Intelligence

(Wechsler, 1999). As the RAI assessment is a short, 55-trial assessment with an administration time of approximately 10-15 minutes, comparison to such short-form IQ estimates may be more relevant than full-length psychometric testing. The RAI correlation coefficient reported for the current study for Full IQ ranks above Spinks et al.'s (2009) results for OPIE-3-4 Subtests (.69), OPIE 3-2 Subtests (.58), OPIE 3 - Vocabulary (.45), OPIE 3 – Matrix Reasoning (.35), OPIE 3 – Picture Completion (.44), Shipley Institute of Living Scale (.34), North American Adult Reading Test (.11), Iowa Test of Basic Skills (.23), the Barona demographics estimate (.01) and Crawford demographics formula (.52) for their high IQ (115+) group. Given that the OPIE is a short form of the WAIS-III, the RAI's superior performance in correlating to Full IQ is encouraging. For example, the OPIE-3-4 computes a Full IQ estimate by administering four Wechsler subtests; Vocabulary, Information, Matrix Reasoning and Picture Completion, and has been found to provide robust estimates of Full IQ scores (Schoenberg, Duff, Scott & Adams, 2003; Schoenberg, Duff, Scott, Patton & Adams, 2006). The finding therefore, that the Relational Abilities Index displayed a comparable level of correlation with Full IQ would appear to offer preliminary validation for its potential efficacy in describing intellectual function. The Ward - 7 Subtest form (.83) was the only measure that demonstrated a stronger correlation with Full IQ when compared to the RAI. Given the discrepancy in administration length between the RAI (approx. 15 minutes) and the Ward – 7 Subtest (requiring approximately 40-50 minutes for completion) the lower level of correlation found for the RAI is unsurprising. The Ward-7 Subtest administered over half the subtests required for computation of a Full IQ score on the WAIS, and therefore should in theory demonstrate such predictive ability. However, the finding that the RAI showed a closer relationship to IQ than many short-form and proxy measures further supports the relevance of relational ability measures to IQ.

Furthermore, Spinks et al. (2009) reported a diminishment in correlations for every single measure of IQ when comparing average IQ (defined as 85-115 for this study) and the above average IQ group (115+). As such, the correlation reported for RAI scores and Full IQ scores is particularly noteworthy, due to the high average IQ of the current sample (118.1). It may follow that even stronger correlations may be found between RAI and IQ with an average IQ group. The correlation between RAI and IQ scores in the current study may be further hindered by the relative concentration of RAI scores towards the upper limit of the RAI (M = 48.94). Correlational analyses depend heavily on the variability of data (Abrami, Cholmsky, & Gordon, 2001; Aron & Aron, 2003; Bates, Zhang, Dufek & Chen, 1996, Crocker & Algina, 1986; Glenberg, 1996; Hopkins, 1998). The Pearson product-moment coefficient, as used in this study, is found by dividing the covariance of two variables by the product of their standard deviations. Therefore, as the variability of RAI scores is relatively low, its covariance with the wider distribution found for IQ scores is reduced. This discrepancy in variability results in a lower correlation coefficient between these two measures than what would have been seen if RAI variability was higher (see Goodwin & Leech, 2006).

The high level of correlation between RAI and WAIS Full IQ scores is complemented by previous findings of a significant correlation with Scholastic Ability as measured by the Differential Aptitude Test (.66; Cassidy et al., 2016). The tentative efficacy of the RAI in predicting IQ as well as educational attainment suggests that relational responding may be viewed as an influential factor in school performance. There are a range of factors which influence educational achievement, such as teaching style (Aitkin & Zuzovsky, 1994; Ebmeir & Good, 1979), teacher clarity (Hattie, 2009), school environment (Kwesiga, 2002), self-efficacy (Collins, 1982; Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002), conscientiousness (Spinath, Freudenthalter & Neubauer, 2010) and student's academic discipline (ACT, 2008). However it is generally accepted that individual differences in intelligence are the largest contributor to the variance in educational attainment (Deary, Strand, Smith & Fernandes, 2007; Jensen, 1998; Kaufman, Reynolds, Liu, Kaufman & McGrew; 2012; Lynn & Meisenberg, 2010; Mackintosh, 1998; Rinderman, 2007). Therefore, it would follow that, due to its close relationship with intellectual performance, relational responding may be of importance to school performance. The findings of Study 2 might therefore further strengthen the case that relational responding training may be of benefit to educational performance thus underlining the potential benefit this relational skills program may be able to provide in an educational context.

The current correlational analysis of RAI and IQ scores support previous studies implicating the importance of relational responding to intellectual performance. In a comprehensive study of relational responding and the WAIS-III, O'Hora et al. (2008) found significant correlations between performance on a temporal relations task and Full, Verbal and Performance IQ. The results of the current study support this covariance, with even stronger correlations found for each scale. In line with the findings of O'Hora et al., the current study found significant correlations between RAI scores and Verbal Comprehension, as well as Perceptual Organisation. While O'Hora et al. did not report significant correlations between the Working Memory subindex and relational ability, correlations between RAI and this subindex reached significance in the current analysis.

The finding that the current relational assessment displayed a stronger correlation with the three main IQ indices in comparison to the temporal relational assessment administered by O'Hora et al. (2008) carries a number of implications. Firstly, it may suggest that the relational frames included in the current assessment (coordination/distinction and comparison) exert a more pronounced impact on intellectual performance than temporal relations. While the relational battery employed by O'Hora et al. focused on temporal relations, the RAI assessed responding to 55 MORE THAN/LESS THAN and SAME/OPPOSITE relational tasks. As previously discussed, the relational frame of coordination (as assessed by same/opposite relational tasks) is perhaps the most salient example of relational ability's contribution to intelligence, and in particular, verbal ability (Stewart et al., 2013). The establishment of word-word and word-object coordination relations underlie verbal communication, and the rapid expansion in the breadth and depth of these frames (alongside other relations) has been posited as a possible explanation of the "language explosion" (Stewart & Roche, 2013). In addition, comparison relations are essential to mathematical abstraction and numeracy. Temporal relations are considered a subset of comparison relations and therefore represent a more limited aspect of relational skill. The frames of coordination and comparison would therefore appear to be relevant to many additional domains of intelligence as assessed by the traditional IQ measures. For example, there are a number of WAIS subtests that are clearly dependent on coordination/sameness, such as Vocabulary and Information, and comparison relations, such as Arithmetic. In addition, the Similarities subtest probes for both sameness and comparison relations by asking participants to identify how one item is the same as another. In contrast, WAIS test items that explicitly assess before/after relations are far less common. It may be possible that findings of correlations between IQ and temporal relational task fluency may therefore reflect the contribution of generalised comparison relations to intellectual performance, rather than temporal relations specifically. This is of course a speculative point, but it is one that can be addressed empirically.

Secondly, the possibly stronger relationship between coordination and comparison relational skills and IQ may indicate that common IQ measures target sophistication in certain relational frames more so than others. While relational skills have been consistently shown to be of importance to intelligence, there is a possibility that IQ tests do not adequately assess the breadth and depth of relational ability. Conversely, it may be suggested that intellectual performance relies on some relational frames more than others. For example, as mentioned previously, the full WAIS battery is limited in regards to explicit assessment of temporal relations. It may be the case that certain relational frames load more heavily onto IQ measures than others. Indeed, the current study indicates that, in comparison to temporal relations, the relational frames of coordination and comparison may provide a greater contribution to WAIS IQ scores. The wider range of relational frames, such as hierarchical, analogical and deictic relations may also represent varying levels of influence on IQ scores. As both IQ assessments and relational responding assessments boast a well-established theoretical and empirical basis, the relative contribution of each relational frame to intelligence warrants further investigation.

At this point, it is important to delineate the distinction currently made between intelligence and IQ. As discussed in the introduction chapter, it is imperative to note that the concept of IQ was intended to reflect an approximation of intellectual performance, rather than equate to intelligence itself. Intelligence and IQ, while sometimes considered mutually inextricable, are separate concepts and should be treated accordingly. The current argument proposes that while intelligence, albeit as a partially culturally defined concept, may be considered heavily dependent on specific relational frames, this dependence may not be accurately reflected in traditional IQ measures. However, it must also be noted that due to the relative dearth of understanding regarding the contribution of relational responding to intelligence *in isolation* of IQ scores, this argument is admittedly speculative. In fact, it may not be possible, by definition, to assess intelligence without assessing IQ. This of course is the perennial circular reasoning problem that arises when one tries to study a hypothetical construct independent of the measures used to quantify it. Therefore, the overlap between the relational ability, IQ and the wider concept of intelligence must also be further examined.

In comparison to O'Hora et al. (2005; 2008), correlations between relational ability and IQ subtests, and particularly Verbal subtests, were far more common. Previous significant correlations for Vocabulary, Similarities, Arithmetic and Information were observed here (at higher *r* values), along with correlations between RAI scores and scores for the Digit Span and Comprehension. In fact, only one of the seven verbal subtests (Letter-Number Sequencing) failed to show statistically significant correlations with the RAI, albeit marginally (p = .08). This would imply that there is a shared variance between relational responding and almost all aspects of verbal intelligence as defined by the WAIS.

Correlations found for RAI scores and Performance subtests were much weaker than those witnessed between RAI scores and Verbal subtests, with only one (Block Design) of six subtests displaying a significant correlation. This correlation is noteworthy as it is one of the most validated nonverbal measures of fluid intelligence, and is a feature of a wide range of psychometric tests (Differential Ability Scales, Elliot, 1990; Kohs Block Design Test, Kohs, 1920; WAIS, Wechsler, 1955, 1981, 1997, 2008;

WASI, Wecschler, 1999, 2011; WISC, Wechsler, 1949,1974, 1991, 2004; WPPSI, Wechsler, 1967, 1989, 2002, 2012). Block Design is considered to be a measure of nonverbal problem solving and visuospatial reasoning (Groth-Marnat, 1984), and is regarded as a reliable measure of fluid intelligence. This result further supports O'Hora et al.'s (2008) assertion that Relational Frame Theory predicts this relationship, because the Block Design subtest requires *pragmatic verbal analysis.* Pragmatic verbal analysis involves arbitrarily applicable relational responding under the control of physical-world relations. The block design subtest represents a test for pragmatic verbal analysis as it requires individuals to arrange a formulation of blocks identical to a model by recognising differences between their current formation and the desired formation. Therefore, in this way, the clear overlap between relational sophistication and the skills essential to complete Block Design items may account for the significant correlations observed.

The consistent trend of RAI scores to correlate more strongly with verbal subscales is of interest, replicating previous findings (O'Hora et al. 2005; 2008) and further implicating the centrality of relational responding to even high-level verbal ability (Cassidy et al., 2011; Stewart et al., 2013). In a previous correlational study involving three subtests of the WAIS, O'Hora et al. (2005) found that participants who successfully completed a complex relational task performed significantly better on the two Verbal subtests (Vocabulary & Arithmetic) but not on the Performance subtest (Digit-Symbol Coding). The current study would replicate such findings, with significant correlations found for the Vocabulary and Arithmetic subtests, but not Digit Symbol Coding. More generally, the discrepancy between correlations of RAI and Verbal and Performance subtests was present in the current study, as relational ability displayed a consistently stronger relationship to the former. This discrepancy further implicates the presence of a more defined relationship between relational ability and the verbal aspect of IQ performance.

The relationship between Verbal IQ test items and relational ability is particularly noteworthy given that Verbal IQ may be viewed in some respects as a more stable measure of intellectual performance. Performance IQ has been suggested to be more susceptible to extraneous variables such as task persistence, mood and attention (Kaufman, 1990, Njiokiktjien & Verschoor, 1998, Sackheim et al., 1992). Test-retest analyses of the WAIS-III also indicate that Verbal IQ is more stable than Performance IQ (The Psychological Corporation, 1997). Therefore, the finding that RAI score predicts Verbal IQ moreso than Performance may propose that the RAI measures skills that are of core importance to intelligence, rather than extra-intellectual, generalised test-taking factors, which may affect performance in IQ tests.

In summary, the current findings indicate a high level of correlation between relational ability and Full IQ. The degree of correlation found between these two metrics was at a level comparable (and at times beyond) that of many traditionally accepted measures and proxy measures of intellectual performance. These findings indicate that the RAI may display predictive validity with IQ, as well as targeting a related if not synonymous behavioural domain. As such, the Relational Abilities Index showed robust predictive validity regarding intellectual performance. This assertion is augmented by the consistent and widespread correlation between relational ability and six of the seven IQ scales and subscales assessed by the WAIS-III (with the exception of Processing Speed). Furthermore, this study offers further support for the proposal that relational responding is of fundamental importance to intelligence, and verbal ability in particular, as evinced by the high levels of co-linearity between RAI and Verbal subtest scores.

4.2.3 Post-intervention improvements in relational skill

In analysing RAI assessments both pre- and post-intervention, there was statistically significant rise in RAI scores following training from Time 1 to Time 2. The eta squared statistic indicated a large effect size. Not only is the significant increase in mean RAI scores from 50.4 to 53.2 (out of a possible 55) noteworthy, there is a shift in standard deviation values from pre-training to post-training towards a more tightly concentrated distribution. Excluding participants who did not complete training, it is illuminating to report that while baseline scores ranged from 40-54, this spread was reduced to 50-55 for follow-up assessment scores. Every participant who finished the training program produced a post-training score of at least 50 out of 55, while 4 participants gained full marks in the post assessment as a result of training. The concentration of post-training approaching at the upper limit of RAI scores further supports the efficacy of the training program in fostering a very sophisticated and deep understanding of relational concepts.

Due to the formulation of the RAI assessment, it is highly unlikely that the rises in scores were due to a practice. While baseline and follow-up assessments took the exact same format and collection of trials, there were variations in directionality of trial statement, the relation probed in each trial and the position of the response buttons. In addition, the stimulus sets utilised for each assessment were completely different from each other, ensuring no relational task was seen twice. In fact, the multiple-exemplar format of the assessment is specifically designed to avoid content-based learning, as the goal of this protocol is to produce generalised, content-free skills. The second assessment therefore functioned as an alternate (but equally difficult) version of the first. To complement this, the short duration of the test (10-15 minutes approx.) and the large test-retest interval would also appear to lessen the likelihood of practice effects.

Therefore, due to the consistent and significant increases in RAI scores witnessed, it can be concluded that the training program isolated, targeted and successfully trained relational skills.

4.2.4 Relationship between RAI score and IQ at follow-up

In comparison to the correlation analysis of baseline scores in the previous study, RAI scores at follow-up displayed weaker correlations with IQ scores and subscores. The strong correlations found for RAI and six of the seven IQ indices and subindices at baseline was not replicated at follow up. In general, the relationship between RAI scores and IQ at follow-up was less clear-cut. In fact, follow-up RAI scores displayed significant correlations for only one main IQ scale, Verbal IQ and one subindex, Verbal Comprehension. Correlations between RAI at Time 2 did not reach statistical significance for Full IQ, or Performance IQ at Time 2. Furthermore, there were no significant correlations detected for any of the other subscales. This would appear to suggest that IQ and relational ability are more closely related before the training intervention that afterward.

To complement the finding of a weaker relationship between post-intervention IQ and RAI scores in relation to baseline comparisons, baseline RAI was also more closely associated to follow-up IQ scores. Specifically, there was a higher degree of correlation between baseline RAI scores and follow-up IQ scores, than between follow-up RAI and follow-up IQ scores. This would imply a less close relationship between relational ability and intellectual performance following intervention when compared to performance before training. Relational Abilities Index scores at Time 1 correlated very strongly with Full, Verbal and Performance IQ scores at Time 2. On the other hand, follow-up RAI scores only demonstrated a significant correlation with one of these three main indices at follow-up - Verbal IQ. In addition, RAI scores at Time 1 displayed significant correlations with follow-up Verbal Comprehension scores, but also follow-up Working Memory. In comparison, follow-up RAI scores correlated significantly with post-intervention scores for Verbal Comprehension only. The strength of this relationship was in fact weaker than that seen between baseline RAI and follow-up Verbal Comprehension. The trend for post-intervention IQ scores to correlate more strongly with baseline relational continued into the domain of Verbal subtests. While only one post-intervention subtest demonstrated a significant correlation with follow-up RAI, six such subtests were significantly correlated with baseline RAI. In summary, relational ability and intellectual performance appear to be more closely connected before training, indicating that follow-up IQ scores are in some way less dependent on relational ability.

It is difficult to ascertain whether the aforementioned finding is a statistical artefact of the decreased variability of post-intervention RAI scores, or possibly an indication that after training, intellectual performance is less reliably predicted by the level and nature of relational ability. It may be that the relational training results in improvements beyond the scope of the RAI, which in turn lead exponential increases in intellectual function, although this is an extremely tentative suggestion. On the other hand, a more grounded assertion would be that despite the training intervention, follow-up IQ is still closely tied to pre-intervention levels of relational ability. While RAI scores displayed a significant rise following training, it is interesting to find that despite this improvement in relational ability, IQ is still dependent to an extent on baseline relational ability. Further investigation is required to clarify this finding, as it may suggest that the current relational training program results in relational skill improvements that do not entirely reflect themselves in IQ gains.

Of relevance to the suggestion that relational skills gains did not translate in intellectual improvement, correlational analyses indicate that change in RAI scores did not correlate significantly with Full IQ change. However, it must be noted that this may be indicative of the relative concentration of RAI towards the upper limit. Furthermore, it was also found that the relationship between training and RAI improvements may not be simply linear, as RAI changes did not correlate significantly with days required to complete training or total trials completed. Therefore, it appears that IQ rises were not a function of time spent training.

The extent to which an individual benefited (in terms of IQ increases) from the training program appears to be more closely linked to their pre-intervention level of relational ability. Specifically, changes in RAI scores negatively correlated with baseline RAI scores and also with percentage of correct responses across the entire training program. This finding provides evidence to suggest that participants who presented with higher levels of relational performance at the outset gained less than those with weaker relational skills. It appears that, for the current study, RAI gains depended more heavily upon the difficulty level relative to the individual. Therefore, it appears that the current relational training protocol may require the addition of more complex relational tasks to challenge advanced participants. By including more difficult tasks, it may provide advanced individuals with more opportunity to learn and improve their relational ability. This issue will be discussed in detail in the subsequent section.

4.3 Ceiling effect

Due to the relatively advanced level of intellectual ability displayed by the current sample, a possible limit to the potential IQ rises possible may have compromised the current findings. While previous studies (Cassidy et al., 2011; 2016) reported large IQ gains in populations within the average IQ range, the current study endeavoured to ascertain whether such gains were consistent for those possible for above average IQ individuals. The current study reports Full IQ gains that are significantly lower to those seen in the previous samples (Cassidy et al; 2011, 2016). The current sample displayed a more advanced level of relational responding pre-intervention when compared to the samples studied previously in both Cassidy et al. studies. This increased level of sophistication, in conjunction to the lower IQ gains found in the current study would therefore seem to implicate the possibility of a ceiling effect.

In addition, the impact of advanced relational proficiency at baseline on the training effects was also assessed. With a RAI mean of 48.9 and a majority (67%) of participants achieving baseline scores at 50 and above, the potential for improvement in performance is somewhat limited by the scarcity of "extra" points to gain. The previous Cassidy et al. (2016) study involved samples in the average range of IQ and found mean RAI increases of 14.7 and 6.6 for primary and secondary school students, respectively. In effect, primary school children displayed considerably larger increases in RAI scores than were observed in the current study. Of particular relevance is that, when comparing the respective mean RAI scores of primary, secondary and college students, primary school students displayed the lowest mean RAI score. This may indicate that higher levels of baseline relational ability may diminish the post-intervention improvements in RAI scores. Correlational analyses from the current study support this suggestion, as RAI gain was correlated negatively with starting RAI.

Due to the inverse correlation between baseline RAI and RAI increase, it would follow that the relational skills training may have been too simplistic for the current sample of high functioning adults. Specifically, due to their already high level of sophistication with relational concepts, the current sample of college students did not stand to gain as much as the younger, less advanced group analysed in Cassidy et al.'s (2011) study. As such, a ceiling effect to gains (literally the upper limit for scores in many cases) may have restricted the sample's potential IQ gains. It comes as no surprise, therefore, that individuals with lower RAI scores generally displayed greater gains at follow-up. For future investigations, it would be diligent to increase the difficulty level of relational tasks for more intellectually well developed individuals. Once again, the strong negative correlation found for RAI changes and the total percentage of correct responses across the training period would support this recommendation. Put simply, participants who displayed less proficient relational responding (i.e., produced more incorrect responses) gained more than those who produced a higher percentage of correct responses.

Participants who displayed a more advanced mastery of relational concepts at baseline may therefore benefit from the introduction of more varied and complex relational tasks. The intervention protocol used here only trained and tested for two of the simpler relational frames: SAME/OPPOSITE and MORE THAN/LESS THAN (albeit to a very high level of complexity). The current sample displayed an already advanced mastery of these two frames at baseline, and as such, their relational skill was not developed as much as the less advanced primary school sample studied in Cassidy et al. (2011). Therefore, more advanced relational frames, such as deictic, temporal or hierarchical relations, and more complex trials with an increased number of nodes, may provide the more advanced individual with the opportunity to develop further intellectually. The inclusion and subsequent training of additional relational frames may result in a wider relational sophistication as well as a more comprehensive repertoire of relational skills. The inclusion of training protocols for deictic, temporal or hierarchical relations in particular may facilitate further benefits for numerous cognitive abilities that are of defined importance to intellectual performance, such as arithmetic, logic, metaphor, analogy and perspective taking (O'Hora et al., 2005; 2008; Gore et al., 2010; Barnes et al., 1997).

McHugh et al. (2004) previously outlined a protocol for assessing complex deictic relations using ME-YOU, HERE-THERE and NOW-THEN relations. Such a protocol could be utilised in the current training framework to assess and improve deictic relational responding. For example, temporal relations could be easily integrated in the form of "A is after B, B is after C"; Is A before C?". Due to the relative imprecision of the contextual cues "before and "after" an additional response option could be included to allow participants to respond that the correct response is not well defined. For example, in the event that A is after B and C is after B, the exact relation between A and C is not combinatorilly entailed. Assessing such unspecified relations (as per Roche & Barnes, 1996) would allow for a far more nuanced assessment of relational skills repertoires. In addition, the inclusion of deictic relations affords an increased level of complexity. As demonstrated by McHugh et al., relational tasks of the type seen above can be altered to include extra relational frames and reversals. For example, for the previously outlined deictic relation, participants may also be given the cue; "if there was here and here was there" in order to add another level of complexity.

In summary, it was found that the current relational training fostered greater improvements in participants with lower levels of baseline relational ability. As the current analysis aimed to assess the boundary conditions of the current training intervention, the finding that the protocol presently in use may be limited in its ability to improve high baseline levels of relational ability implicates a definite ceiling effect to its efficacy. The current findings therefore suggest that a substantial modification of this protocol in order is required to provide benefits to relational ability of the magnitude reported by previous analyses (Cassidy et al. 2011, 2016).

4.4 Efficacy of relational training in increasing IQ scores

While previous studies (Cassidy et al. 2011, 2016) identified a significant effect of relational training on Full IQ scores, Study 1 found that relational skills training did not exert a significant effect on most IQ scores when compared to a control group matched for baseline IQ. Overall, results indicate a large increase in scores for both groups across two of the three main measures of IQ (i.e., Full & Performance IQ). While the experimental group displayed mean gains of around 9 points in Full IQ, control participants also demonstrated a significant rise of 6 points. The difference in rises on this scale were not significant. Of great interest were the increases in scores on Performance IQ, with both groups displaying increases of over 10 points on this scale. While the experimental group displayed considerably larger increases in Verbal IQ (5.9 points) compared to the control group (3.1 points), this discrepancy also did not result in a significant effect for the intervention. The failure to finding an effect of the training intervention of the three main IQ indices comes with the considerable caveat that large and statistically significant rises were witnessed in the non-intervention group. As improvements in performance for control participants can be expected as a result of reassessment alone (Wechsler, 1997), a 3-month testing interval was maintained in an attempt to reduce the possible effects of practice. However, due to the significant increases displayed by the control group for Full and Performance IQ, this testing interval did not eliminate practice effects. In summary, the current findings do not replicate the findings of Cassidy et al (2011, 2016), as training did not exert a significant effect on training when compared to a control group. The issue of practice effects will be discussed later in further detail (see Section 4.5.4).

In terms of Verbal subindices, there was a significant effect of relational training on Verbal Comprehension scores. While there was a significant main effect for time for both groups, the mean increase in scores was larger for the experimental group (6 points) when compared to the control group (less than 2 points) on this subindex. These results indicate that membership of the training group resulted in significantly greater increases in Verbal Comprehension scores when compared to the control group. This finding would therefore support the efficacy of the relational skills intervention in improving Verbal Comprehension scores.

The finding that relational skills training resulted in increases in Verbal Comprehension scores is perhaps unsurprising given the theoretical overlap between relational ability and Verbal Comprehension subtests. Verbal Comprehension refers to an individual's vocabulary, verbal expression, verbal knowledge and to what extent he/she can conceptualise verbal information (Groth Marnat, 1984). Groth Marnat also proposes that higher scores on this subindex translate into practical benefits in educational contexts, due to the defined importance of verbal communication in academic pursuits. Therefore, any training intervention that can improve the underlying skills relevant to this scale may provide a wide range of benefits to general intellectual function.

The effect of relational training was not replicated for Working Memory, with no significant interaction effect between intervention type and time. Uniquely among the other IQ scales and subscales analysed in this study, there was no main effect for time regarding Working Memory scores. In terms of Perceptual Organisation and

Processing Speed, there was no significant interaction effect between intervention type and time for either subindex. In line with the general trend of results, there was a large main effect for time on both measures. In summary, these results indicate the relational training did not exert a significantly different effect on Full, Verbal and Performance IQ scores. In addition, scores for Working Memory, Perceptual Organisation and Processing Speed did not appear to benefit to a great extent from relational training when compared to no intervention. Relational training did, however, display a clear effect on improving performance on the Verbal Comprehension sub-index of IQ, far beyond what could be explained by a practice effect.

4.5 Potential Modifications to the Relational Abilities Index

The investigations of the current thesis uncover a number of possible modifications that may improve the capacity of the Relational Abilities Index to provide a more nuanced and sensitive measure of relational responding. As is the case with the relational training program, the current composition of the RAI may benefit from a number of alterations including: (1) the addition of more complex test items, (2) the inclusion of additional relational frames such categorical and deictic relations and (3) a reorientation of the scoring system sensitive to the relative difficulty of trials. Each of these modifications will now be discussed in isolation in regards to the improvement such alterations may provide for the current RAI scale.

As previously discussed in relation to the relational training program (see Section 4.3), the concentration of participant scores towards the upper limit of the RAI calls for the extension of the current 55 trial structure to include more challenging and complex trials. To achieve this, the RAI should be adjusted to include more complex relational frames. The lack of sensitivity evinced by the current RAI structure is exemplified by the relative narrowness of RAI score distribution (SD = 5.8) when compared to Full IQ (SD = 13). It therefore follows that, for high IQ and RAI individuals, the RAI does not currently allow for the subtle differentiation of relational ability at the upper end of performance. In essence, the performances of individuals at the upper end of relational responding skill are crudely grouped together due to the RAI's failure to accurately reflect individual differences for high performance participants. In order to rectify this issue, the RAI should undergo restructuring similar to that proposed for the relational training program. While there are 55 levels required to be completion in the relational training program, there are an additional 15 levels of higher complexity that are supplementary to the 55 core levels. These 15 levels include four relational premises followed by a relational question and represent a significant increase in task difficulty. The first step in increasing the difficulty level of the RAI should be the inclusion of 15 extra test items that represent the 15 additional levels. Indeed, the most readily available option in increasing task difficulty is the inclusion of tasks composed of greater numbers of relational premises. By requiring participants to respond to tasks involving 4 or 5 relational frames, the RAI will be made more sensitive to individual differences in the advanced range of performance. Therefore, as the relational training program is expanded with the inclusion of more complex relational tasks, the RAI should continue to mirror this new structure.

A second proposed modification to the RAI is the inclusion of additional relational frames, such as categorical, deictic and temporal relations. As discussed previously, the contribution of such relational frames to intellectual performance is well-established (Gore et al 2010; O'Hora et al 2008; McHugh et al. 2004). The current structure of the RAI assesses only more than/less than and same/opposite relations, and as such, facilitates a measurement of a relatively limited collection of relational skills.

Therefore, the inclusion of these additional relational frames may provide a more accurate measurement of the wider range of relational frames that constitute relational responding skill. Furthermore, the addition of these frames will also allow for increases task complexity, as the range of relational frames can be easily combined to create a task assessing relational responding across numerous relational frames (see McHugh et al 2008).

A third proposed modification of the RAI is a restructuring of its scoring system to reflect the increasingly complex nature of trials. The current RAI scoring system offers limited utility in describing the nature of each participant's performance. As the scoring system employed by the RAI is a summated scale, it does not reflect the difference between an incorrect response on the first item or the last, valuing every trial as being equal in score regardless of their respective difficulty. As such, the introduction of a modified Guttman scale format should be investigated for the RAI. Guttman scaling (sometimes referred to as cumulative scaling) is designed to establish a one-dimensional continuum of a given variable (Trochim, 2002). Guttman scales arrange test items in a cumulative order (e.g. order of difficulty) allowing the assumption that if an individual correctly responds to a given test item, he/she will also correctly respond to previous test items (Guttman, 1954; Mokken 1970; Manhein, 1977). While often employed to assess an individual's agreement to a set of statements, this scaling method can be applied to the RAI in a manner similar to that seen in traditional psychometric assessments (e.g. WAIS-III). For example, the Wechsler scales of Intelligence outline subtest-specific discontinue criteria that ends each subtest following a certain number of incorrect responses. As such, if an individual receives a score of 25 on the Vocabulary subtest, the examiner knows that this score was accrued in the earlier, less challenging subtest items. In contrast, the RAI does not offer any such indication. As such, the restructuring of the scaling system (from summative to cumulative) employed by the RAI would represent an improvement in the utility of this scale.

4.6 Potential Covariates of relational training efficacy

The current study identified a number of covariates, which may have exerted a significant influence on the efficacy of relational skills training in improving intellectual performance. As this investigation was designed to assess the boundary conditions that may impact the effectiveness of relational training in increasing IQ scores, potential confounding variables (age & starting IQ) were outlined from the outset. In addition, a number of potentially influential variables emerged during analysis, such as practice effects, stability of working memory scores and the influence of timing on performance. As such, the effect of each of these variables will be discussed subsequently.

4.6.1 Age

As this is the first study investigating the efficacy of this relational skills training program with adults, age was also investigated as a possible mediating factor in IQ changes. Age was not found to correlate significantly with Full IQ change or RAI change. This comes with the major caveat that while correlational findings suggested no relation between age and IQ change, the rises seen for this adult group (9 points) were much lower than the gains witnessed in a sample of primary school students (23 points). While the current study may provide tentative evidence of a diminished effect for adults, it is also important to note that similarly to RAI scores, experimental participants ages were tightly concentrated with most participants aged in their late teens and early twenties. Therefore the most illuminating finding in this regard is the difference between the levels of IQ increase found in the current study versus previous studies that implemented relational skills training to younger populations, rather than any correlations the current study can provide.

A possible confounding factor in investigating the effect of age is the high level of relational ability displayed by the current sample. Previous research (McHugh et al. 2004) indicates that relational ability may increase into adulthood. As expected due to both their age and IQ level, the current sample of college students displayed high levels of relational skill at the outset. Therefore, it is difficult to extrapolate the effects of this advanced level of skill from the possible influence of age on IQ gains. As the training program was originally developed for use in a less-advanced child population, further adaptations may be required to foster similar ground-breaking IQ rises in older, more advanced populations. While, the issue of ceiling effects, as well as possible modifications to the training program have been discussed previously, the confounding impact of age on IQ gains deserves further analysis.

4.6.2 Starting IQ

While the potential delimiting effect of high baseline RAI scores has been discussed, the effect of baseline IQ on post-intervention outcome measures was also studied. Results indicate that there was not a strong correlation between baseline intellectual ability and increases in IQ following relational training. Baseline scores for Full IQ did not predict changes in Full, Verbal or Performance IQ. Scores for both Verbal IQ and Performance IQ followed a similar trend, as neither score predicted the changes seen on each scale respectively. Participants' original Full IQ scores did not show a significant correlation with increases in relational ability index scores. The lack of a significant relationship in this regard may suggest that relational ability can be improved irrespective of starting intellectual ability. These findings therefore, suggest that starting IQ does not have a delimiting impact on improvements post-intervention, as there appears to be very little relation between baseline IQ measures and the changes in both IQ and RAI scores. However, it must be noted that the IQ rises seen in this high IQ group were considerably lower than those seen previously in average IQ populations (Cassidy et al., 2016). Furthermore, while baseline IQ did not significantly predict change in Relational Ability Index scores, correlation coefficients suggest a considerable inverse relation between these two metrics. This finding would seem to indicate the negative effect starting IQ may possibly exert on the effectiveness of the current relational training. While the current study offer very preliminary findings, further investigations are warranted to isolate and analyse the potential impact of both age and starting IQ in more detail.

4.6.3 Practice effects

The issue of pronounced practice effects displayed by the control group exerted a considerable impact on the current analysis. While rises in RAI scores, serving as a manipulation check, indicate that relational skills were improved, the presence of significant rises in control group IQ scores rendered it difficult to extrapolate any effect of training on IQ scores (with the exception of Verbal Comprehension). While the presence of practice effects did not obscure the entirety of intervention effect, these effects did significantly inhibit the ability of the experimenter to extrapolate intervention effects from practice effects on a number of metrics.

Test-retest gains for the control group were particularly prominent for Performance IQ and the two Performance subscales, Perceptual Organisation and Processing Speed. In fact, gains in Performance subscales for control participants were slightly greater than those seen in the experimental group. As the control group underwent no intervention, these increases can only be explained as being the result of a practice effect due to retesting. The heightened impact of retesting on Performance items is well supported (Basso, Carona, Lowery & Axelrod, 1997; Catron & Thompson, 1979; Kaufman, 1990; Mitrushina & Satz, 1991; Wechsler, 1974). However, the current increases in control participants' Performance scores are beyond those cited in many investigations of IQ practice effects. In one of the first studies of retest effects on the WAIS-III across four age groups (Tulsky, Zhu & Ledbetter, 1997), found increases of 2.5-3.2 and 2.5-8.3 for Full IQ and Performance IQ respectively after a testing interval of 34.6 days. In general, the retest gains reported by Tulsky et al. (1997) were not at the level found in the current study. Basso et al. (2002) found Full IQ and Performance IQ gains of 4.8 and 10.6 respectively after when comparing 3- and 6 – month testing intervals. These increases were similar to those witnessed in the current analysis. Rises reported for Perceptual Organisation (8.3 points) and Processing Speed (7.1 points) however, were not at the level found for our sample. It is important to note that while these were mean increases for participants who were retested after 3 months and participants retested after 6 months, no significant effect for increased time was detected. Estevis, Basso & Combs (2012) implemented a similar design using the WAIS-IV, and supported the previous findings of a lack of difference in practice effects following a 3- or 6-month interval. Such a finding would seem to imply that practice effects of repeated assessment are somewhat resistant to preventative effects of increased testing intervals, and therefore difficult to eliminate.

Perhaps more pertinent to the current study is Rapport et al.'s (1997) finding of a heightened practice effect for individuals with higher starting IQs. Rapport et al.'s sample was divided into three groups: Low Average (80-90), Average (95-105) and High Average (110-120) who were tested at 2-week intervals using the WAIS-R. For the Low Average group, IQ gains were roughly half that seen in the High Average group. Rapport et al.'s (1997) also reported a larger practice effect for Performance IQ versus Verbal IQ, in line with the current findings. For the High Average group, the study found increases of 11 and 13 points for Full and Performance IQ respectively. Given the above average IQ of the control group, the findings of Rapport and colleagues would appear to have been replicated for the most part. The large practice effects found in the current study may thus be accounted for in part by the above average baseline IQ of our sample.

Rapport et al. (1997) propose two main reasons for practice effects: (1) specific elements of an assessment may be memorised and (2) instructions of test procedures are memorised. It is also stated that test items that have easily memorably single solutions (i.e. object assembly or picture arrangement) are more likely to display a practice effect after repeat assessments. To complement this, procedural practice effects are due to repeated exposures to a subtest resulting in increased familiarity with materials, as well as the progressive development of more efficient problem-solving faculties. In its simplest sense, the individual is literally getting better with practice. The development of more effective strategies in particular, is cited as one of the reasons that speeded subtests tend to be most affected by practice effects (Rapport et al. 1997, Sattler, 2001).

Given the educational background and high level of intellectual performance present in the control sample, it is reasonable to suggest that many of the extraneous factors that contribute to performance, such as attention, motivation and task persistence (Duckworth, Quinn, Lynam, Loeber & Stouthamer-Loeber, 2011; McCaffrey et al., 2000; McCaffrey & Westerveld, 1995; Reeve & Lam, 2007; Thorndike, 1949) may be more prevalent than in an average sample. As the sample has displayed a high level of academic achievement, it may follow that such individual may be better "test takers".

As schooling has been reliably shown to positively impact IQ test performance, the practice effects witnessed in the current sample may be explained, at least partially, by participants' continued engagement in third level education. It is well established that education exerts positive effects on intellectual performance (Baker, Salinas, & Eslinger, 2012; Ceci, 1991; Falch & Massih, 2010; Harnqvist, 1968; Husen & Tuijnman, 1991; Herrnstein & Murray, 1994). Indeed, the progressive increases in IQ scores across successive generations, termed the Flynn effect, has been suggested to result, in no small part, due to increased participation in formal education over the past century (Flynn, 2006; Baker, Eslinger, Benavides, Peters, Dieckmann & Leon, 2015). Perhaps the earliest investigation into the effect of schooling on IQ, conducted by Freeman (1934), concluded that without the mental stimulation schooling can provide, "intellectual development will be seriously limited or aborted." (p.115). Falch and Massih (2010) found that four to five additional years of schooling increases IQ by approximately one standard deviation. Other estimates suggest extra education can result in IQ rises of 2-4 points (Winship & Korenman, 1997; Neal & Johnson, 1996; Cliffordson & Gustafsson, 2008). In addition, it has been widely accepted, according to some authors, that each year of schooling increases IQ by one point (Jencks, 1972; Herrnstein & Murray, 1994). A particularly illuminating study, carried out by Baltes and Reinert (1969), investigated the impact of one extra year of schooling on intellectual performance in a sample of eight-year-old German children. As the German school system required children to be three years of age before April 1 in order to start school that September, Baltes and Reinert recruited three samples of children (aged 8, 9 & 10) born just before or just after that April deadline. This allowed the comparison of two groups of children who were virtually the exact same age, with one group receiving an extra year of schooling. Using the German Begabungs assessment system, which was based on Thurstone's classification of Primary Mental Abilities, the study found a considerable effect of the extra year in education on test scores. In fact, scores for 8 year olds with extra schooling more closely resembled scores for 10 year olds without the extra year, rather than their less schooled peers. Such findings suggest that participants attending college between assessments may have contributed to the IQ rises witnessed in the current thesis.

Formal education has been associated with improving a variety of cognitive skills and functions that contribute to intellectual performance (Ceci, 1991), such as working memory and problem solving (Nisbett, 2009), abstract reasoning and cognitive flexibility (Stevenson & Chen, 1989; Diamond & Lee, 2011) and verbal ability (Carlsson, Dahl, Ockert & Rooth, 2015). While each control participant had completed only 3 extra months of schooling between baseline and follow-up assessments, the results of Carlsson et al. suggest that discernible improvements in verbal ability can be found after just 10 extra days of schooling. Pascarella and Terenzini (1991) propose that while attending college results in significant improvements in cognitive performance, the bulk of this growth happens in the first two years of attendance. As most of the current sample was in the first year of college, this point is particularly noteworthy. Pascarella and Ternezini go further to suggest that no other social institution offers such a setting and impetus for substantial cognitive growth due to its capacity to train "logic, critical thinking and the evaluation of alternative ideas and course of action" (p.59). Therefore, it is not unreasonable to suggest that due to the underlying cognitive effects schooling can provide, control participants participation in third level education may have contributed to IQ rises.

Younger populations may also display heightened practice effects when compared to older populations (Catron & Thompson, 1979; Horton, 1992; Salthouse, 2010). Specifically, the WAIS-III Technical Manual (The Psychological Corporation, 1997) reports that older age groups show higher levels of score stability after retesting in comparison to the age group studied in the current study. While correlational analyses in this study did not find that age predicted IQ gain, it is important to note that due to the close concentration of participants in their late-teens/early twenties, the possibility of more stringent investigation was precluded.. It may be that the control sample represents the "perfect storm" of practice effects – above average IQ, high educational attainment, young, healthy and free from any learning difficulties.

4.6.4 Post-intervention changes in timed and untimed subtest scores

Previous analyses (Basso, Bornstein & Lang, 1999, Dodrill & Troupin, 1975; Rapport et al. 1997) indicate that increases in IQ subtests that involve a timed element may underlie much of the practice effects found following multiple assessments. The development of more effective strategies in particular, is cited as one of the reasons that speeded subtests tend to be most affected by practice effects (Rapport et al. 1997, Sattler, 2001). Differences between timed and untimed scores across the study were therefore identified as an area of interest. The 13 subtests administered were divided into those which involved a time limit or rewarded speed of response (Picture Completion, Digit-Symbol Coding, Block Design, Arithmetic, Picture Arrangement & Symbol Search) and those that did not involve a timed aspect (Vocabulary, Similarities, Digit Span, Matrix Reasoning, Information, Comprehension & Letter-Number Sequencing). Mean standardised scores for both timed and untimed subtests were then computed for each group. Results indicate that there was no significant difference for baseline standardised timed scores between the control and experimental groups. There were statistically significant rises in scores for both across the two testing periods, with the experimental group producing slightly smaller score increases. Results from a mixed between-within ANOVA found that there was no significant interaction effect between intervention type and time. In summary, the training did not exert a significantly different effect on subtest that involved a timed aspect.

While there was no significant effect of training on timed subtests, analyses of untimed subtests provided contrasting findings. While baseline differences between the control and experimental groups for untimed subtests were insignificant, there appeared to be significance between-group differences in rises for these subtests. There was no significant rise in scores for the control group from Time 1 to Time 2. There was however, a significant rise witnessed in the experimental group from baseline to follow-up. Due to these increases being found for experimental participants only, there was a significant different between the control and experimental groups untimed subtest scores at follow-up. The second ANOVA reported that the interaction effect between intervention type and time was approaching significance. However, due to the robust discrepancy witnessed between groups, it could be hypothesised with reasonable confidence that this interaction would reach significance in a larger sample. While there was no significant interaction effect regarding rises in Full IQ scores, these results indicate a qualitative difference in the nature of rises seen between the control and intervention groups, on the subtest level.

Further informing the previous finding was the result that the vast majority of the control group's IQ rises were due to increases in scores on timed subtests. In fact, ,every subtest that demonstrated significant rises for this group (Picture Completion, Block Design, Digit Symbol Coding, Picture Arrangement, Symbol Search) rewarded quickness of response by either giving bonus points or by enforcing a time limit. In addition, only one subtest which rewards response speed, Arithmetic (a verbal subtest) did not increase significantly between testing periods. All other subtests did not have a timed element, and did not show significant rises for control participants. The discrepancy between mean increases in timed subtest scores (9.8 points) versus untimed subtest scores (0.9 points) for control participants, demonstrates that the former accounts for the control groups IQ gains virtually in their entirety. Given the mean Full IQ of the control group at baseline, the increase of 10 standardised points witnessed equates to roughly 6-7 Full IQ points, which is consistent with Full IQ gains found in the control group. To complement this, scores on untimed subtests displayed a rise of less than one point for control participants. As improvements in timed subtest performance was the largest contributor to the control group's IQ increases, the finding that this was not the case for the experimental group may help extrapolate the effects of relational training previously obscured by practice effects.

Regarding the experimental group, there were significant rises on both timed and untimed subtest scores. While increases in mean timed subtest scores (8 points) were slightly lower than those seen in the control group (9.8 points), increases in untimed subtests scores (4.4 points) for this group far outweigh those of the control group (0.9 points). This stark contrast goes a long way in delineating the impact of the training program in comparison to non-intervention, as it proposes that such a training regimen leads to qualitatively distinct accruements in intellectual performance. The large increases in IQ scores for control participants contributed to our failure to find an interaction effect in support of the training program, but it now appears that the training may lead to more than can be expected given a practice effect. As indicated by the differential effect on Verbal Comprehension scores, relational training was found to lead to significantly greater improvements in performance on untimed subtests. Given these findings, it comes as no surprise that the Verbal Comprehension index is the only one of the seven IQ indices and subindices that is computed solely on the basis on untimed subtests (Vocabulary, Information & Comprehension). Each of the other six metrics is calculated by using at least one timed subtest.

While the discrepancy between groups in rises between timed and untimed subtests is prominent in the current analyses, the relative stability of these scales is an important consideration. On the whole, untimed subtests display greater test-retest stability in comparison to timed tests. For example, four of the five subtests which demonstrate the highest stability coefficients for this age group are untimed subtests (Information, Vocabulary, Digit Span & Block Design) according to the WAIS-III Technical Manual (The Psychological Corporation, 1997). The experimental group demonstrated significant increases in scores for three of these four subtests (Information, Vocabulary and Block Design), while control participants displayed significant rises for Block Design only. Therefore it may be proposed that there are significantly different processes at play underlying the IQ gains witnessed in each group. The experimental group produced rises in subtests that for the most part, appear less susceptible to practice effects and are more stable at retest. The control group on the other hand, demonstrated rises in timed subtests, which have previously been implicated as being vulnerable to practice. It could therefore be proposed that the rises in IQ scores displayed by experimental participants may be more reflective of a genuine increase in intellectual performance than those seen in the control group, even though the follow-up IQ gains were not significantly different in most cases. The rises observed in the control group are more likely to be accounted for by increased performance resulting from memorisation and the development of more effective problem solving strategies due to practice, rather than an improvement in intellectual ability.

4.6.5 Stability of Baseline Working Memory scores

While the significant increases in scores on IQ indices and subindices was pervasive across both groups included in this study, these increases were not uniform across all IQ metrics. As mentioned previously, Working Memory scores remained relatively unchanged for both the experimental and control group. Placing this into the wider context of previous investigations into intellectual interventions, the failure of the current training program to produce any effect on Working Memory carries a number of implications.

Firstly, the lack of increases in Working Memory scores indicates that the current relational program does not appear to exert a beneficial effect on working memory capacity. Some of the most noteworthy research on intellectual enhancement in recent times has focused on improving levels of Working Memory as a means of ameliorating intellectual performance (e.g. Buschkuehl et al., 2008, Jaeggi et al., 2008). While the relational skills intervention implemented in the current study does not target memory skills, it is of interest to find that rises seen in the experimental group were not "inflated" by increases in Working Memory capacity. For the experimental group, the Working Memory subindex displayed a statistically insignificant mean rise of just over one point, and therefore could not account for a substantial portion of the IQ rises seen. In fact, Working Memory was the only IQ index or subindex that did not rise significantly over time. This finding would indicate that if such working memory training can lead to the increases in intelligence the authors propose (Buschkuehl et al. 2008; Jaeggi et al., 2008), such training may not the only means by which intelligence can possibly be improved. While the intervention group did not display significantly greater rises than the control group in Full IQ, the current data suggests a qualitative difference in the rises seen. Therefore, the training program may produce benefits for

IQ that are distinct not only from practice effects, but also the those demonstrated in Jaeggi and colleagues' studies. Although this suggestion comes with the considerable caveat that there was not a significance difference in between group Full IQ rises, the discrepancy in the nature of IQ rises may indicate that there are numerous training protocols which may lead to improvements in intellectual performance.

In comparison to training procedures outlined by Jaeggi and colleagues, the current training program may produce more generalized effects due to its efficacy in raising Verbal Comprehension scores. Concerns have persistently been raised over the generalisability of working memory improvements to intellectual function (Ackerman, Bier & Boyle, 2005; Colom, Abd, Quiroga, Shih & Flores- Mendoza, 2008; Kane, Hambrick & Conway, 2005; Moody, 2009). In comparison, due to the potential impact on Verbal Comprehension, the current training program harbours the possibility of genuine improvements in intellectual function, rather than improved performance on IQ proxy measures. According to Groth-Marnat (1984), Verbal Comprehension assesses an individual's vocabulary, verbal expression, verbal knowledge and to what extent he/she can conceptualise verbal information. Verbal Comprehension and vocabulary in particular, are consistently shown to be among the most reliable predictors of an individual general intellectual ability (Cunningham & Stanovich, 2004, Jensen, 1980; Matarazzo, 1972). The generalisability of Verbal Comprehension scores is therefore well established. It may be suggested, with relative confidence that Verbal Comprehension scores refer to a set of skills that are of great importance to intelligence. As such, an intervention that displays the capacity to improve the skills relevant to this scale may be of pronounced benefit to general intellectual ability, as well as educational attainment specifically. It is in this way that the current intervention may prove invulnerable to the criticisms commonly levelled at working memory training programs.

4.7 Summary

While the current study displayed diminished effect due to a number of possible confounding effects, our results complement previous work outlining the relevance of relational responding to verbal and intellectual skills. Furthermore, while the effects found with the current sample are not as large as those reported in the Cassidy et al (2011, 2016) studies, it is important to note that training did exert a considerable effect on numerous IQ domains which cannot be easily accounted for by chance or practice effects. Such findings might therefore support the tentative efficacy of a relational skills training protocol in increasing intellectual performance despite the current study's failure to find significant intervention effects on the main metrics of IQ.

Due to the weaker effect found for the current sample, the boundary effects and delimiting factors present must be considered more closely. While the primary focus of the current study was to ascertain whether a relational skills training program could result in significant IQ gains in comparison to a control group, the current analysis also shed light on possible boundary effects of such an intervention. While starting IQ did affect post-intervention IQ gains, results indicate that baseline relational ability is inversely correlated with post-intervention IQ and RAI gains. This finding may indicate the current training protocol was ineffective in improving relational responding in more advanced participants. As such, modifications to this training procedure may be required to produce more significant IQ rises in more advanced populations. While the current analysis identified a number of possible boundary conditions of the current intervention, much remains to be understood in terms of environmental and interpersonal variables that may exert an influence on the effectiveness of relational training..

4.8 Limitations of the current study

There are a number of factors that may have potentially limited the reliability and generalisability of the current investigation. While the current analysis intended to assess the possible effects of adult age on the efficacy of a relational training procedure in improving IQ, the low level of variance in recruited participants' age limited the extent to which these factors could be analysed. Correlational analyses display a pronounced reliance on the variability of data (Bates et al., 1996; Hopkins, 1998). Due to the relatively tight concentration of participant age, the current investigation lacked the nuanced analysis that could have been provided by a greater age range and distribution. While the current finding of a diminished efficacy in improving adult IQ (in comparison to previous analyses of its effect on child IQ) is illuminating in itself, further investigation is needed to provide a comprehensive analysis of the effect of age on the efficacy of relational training in improving intellectual performance. As the current data set represents a limited range of adult age, subsequent studies may assess relational training and its efficacy across a number of child and adult age groups. Such a design could provide correlational analysis and between-group comparisons across numerous age groupings in order to study age as a potential confound of training efficacy.

The current analysis may also have benefitted from a larger number of participants. While an appeal to the need for a larger sample in the search for significant variable effects might signal a lack of emphasis on tighter experimental control, in the case of the current study the need for a large sample is due to known processes relating to the increased pronouncement of practice effects for higher IQ samples. Specifically, a large sample size may have allowed for an investigation of relational training efficacy considerably less affected by the large practice effects found in the current analysis. In effect, such processes reduce the inter-group differences because there is greater overlap in-group means. This represents a reduction in statistical power to extract out the effects of the intervention over the known high starting IQ confound. Thus, in this case the appeal for larger samples does not represent a chasing of statistical significance at the expense of better experimental control, but a measure to lessen the deleterious effects of practice effects found in control groups. While the behavior analytic tradition is characterized by an emphasis on low n, or even single subject designs (see Skinner, 1966, Sidman, 1960), it is expected that given a larger sample, further between-group differences may have emerged for the current design. Given the clear trend for experimental participants to display considerably larger rises on both Full IQ and Verbal IQ when compared to the control group, such between-group differences in IQ rises may have reached statistical significance given a larger sample. A larger sample may thus further inform the preliminary positive findings posited in support of relational training in raising IQ, as outlined in previous analyses (Cassidy et al., 2016) and in the current study to a lesser extent. Conversely, it may also be the case that these effects “wash-out” when analysed in larger cohorts. As such, the current research stream may benefit from the implementation of larger scale randomized control trails investigating effectiveness of relational training in ameliorating intellectual performance.

The failure to assess relational ability in control participates at both baseline and follow-up precluded the possibility of ruling out practice effects as the mechanism underlying RAI score increases. While the current design represented the first randomized control trial of the current relational training procedure, it failed to assess the stability of relational ability in a control group. The collection of RAI data for control participants would have provided the opportunity to assess between group differences in RAI scores change for both intervention and non-intervention groups. As discussed previously (see Section 4.4), the possibility that practice may account for the post-intervention RAI score increases found for the experimental group is unlikely. However, this possibility could not be empirically tested due to the lack of control data in this regard. Future analyses should therefore assess the stability of RAI scores following no intervention over an appropriate time period, as this may provide an a clear indication of how much of post-intervention RAI rises can be accounted for by relational training.

Finally, in comparison to previous analyses of relational training in improving IQ, there was significantly less experimental control regarding where and when the relational training took place. Participants trained at their leisure, therefore completing training in an uncontrolled environment. While it is difficult to estimate the potential deleterious effects of extra-experimental environmental effects (such as noise, distraction, participant fatigue), increased control over the regularity of training sessions may have been beneficial in ensuring the quality and quantity of training sessions. A small number of participants displayed relatively irregular patterns of training at times by failing to complete the 2-3 weekly sessions recommended by the experimenter. While the experimenter was in a position to prompt participants if there were breaks in the training regimen, this was not always effective in re-engaging participants in training. Therefore, the diminished experimental control over the regularity and location of training may be considered a clear deficiency in the design of the current investigation and at present has had an unknown effect on the outcomes of the intervention.

4.9 Implications for future research

As the current investigation represents the first analysis of a relational training protocol administered to a sample of adults, it bears a number of important implications for future research. Study into the effectiveness of a relational skills training protocol is still in its early stages. Nevertheless, the current investigation aimed to explore possible boundary conditions regarding the efficacy of such a protocol, and as such, identified a number of findings that require consideration. Most notably, the failure of the current study to replicate the magnitude of IQ rises reported in previous studies (Cassidy et al., 2011, 2016), raises questions regarding the generalisability of previous findings to adult populations. The current findings may suggest that the relational training program is ineffective (or currently insufficient) in improving adult intellectual performance. While the possible delimiting effect of maturational age on the effectiveness of the current training procedure has been discussed previously, much still remains to be learned regarding the relationship between age (i.e., current stage of intellectual development) and interventions designed to improve intellectual performance.

It appears that ceiling effects may have further hindered the recorded efficacy of relational training in improving IQ. Due to the finding of an inverse relation between baseline relational ability and IQ gains, the possibility that a more challenging and comprehensive repertoire of relational tasks may lead to greater IQ increases warrants further investigation. As the current sample displayed high levels of relational ability at baseline, it would follow that participants did not stand much to gain from the subsequent training. As there was an inverse relationship between percentage of correct responses and IQ increases, it appeared that those who displayed lower levels of relational ability displayed greater IQ gains. The training protocol employed may therefore require the addition of more complex and varied relational tasks to provide a challenge and opportunity to improve for more advanced participants. One possible modification to the training procedure would be the inclusion of a wider range of relational frames (e.g., temporal, hierarchical and analogical relations), as a number of such frames have been implicated in intellectual performance (see Section 4.3). The inclusion of additional relational frames as well as the integration of more advanced levels of relational difficulty may extend the efficacy of training in improving relational and intellectual ability in advanced populations. In addition, as previously discussed, the current findings highlight a number of possible limitations in the structuring and scoring system underlying the current RAI assessment. By completing the modifications recommended in the current thesis, the descriptive power and validity of the RAI may be substantially enhanced.

One of the most salient implications of the current findings is the identification of possible boundary effects to the efficacy of relational training. Subsequent studies should aim to extrapolate the effects of these boundary conditions statistically, in order to gain a better understanding of their relative effects on IQ increases. The current study is not sufficient in assessing the effects of such variables (e.g. age, starting IQ) in isolation of each other, and to an extent the effects of these confounds on the current sample may be relatively inextricable. Future investigations may implement targeted analyses of each of these confounds in order to clarify the respective influence of each variable. Future investigations may implement training interventions in samples which control of each of these potential boundary conditions. For example, it may be worthwhile to assess the effects of relational training between groups of varying IQ levels, while controlling for age, and vice versa.

The current findings propose that, while the relational training did not exert a significantly greater effect on the main IQ indices when compared to a control group, there was a clear discrepancy in the nature of IQ rises seen between groups. Specifically, the finding that relational training produced much greater effect on the untimed subtests of the WAIS when compared to control group rises would indicate that relational ability makes a significantly greater contribution to the aspects of intelligence assessed by these subtests. Due to the fact that most untimed subtests are Verbal subtests, the findings of the relational training program's considerable efficacy in improving scores on these scales is unsurprising. However, while the relationship between relational and verbal ability is well established (see Section 1.8), the contribution of relational training to untimed subtests is less clear and therefore requires further delineation. This preliminary finding requires empirical replications, as well as theoretical consideration in order to gain a better understanding of why relational responding is of such importance to the untimed subtests and timed subtests.

Findings from the current study suggest that the current relational training program did not lead to improved processing speed levels or performance on timed test items. Due to the relevance of processing speed to intelligence, subsequent research may focus on devising a RAI fluency score (time x accuracy) as well as measures designed to train participants to respond more quickly, as well as more accurately, to relational tasks. The computation of a RAI fluency score may offer descriptive utility regarding levels of relational ability displayed by a given individual. As the RAI score (out of 55) is the main metric of the current training program, it is not sensitive to differences in response time between participants. For example, the current relational training program does not quantify the difference between a participant who completes the RAI assessment with perfect accuracy in 5 minutes, versus a participant who completes the assessment with perfect accuracy in 15 minutes. In addition, the development of training procedures that require participants to respond in progressively shorter time frames should therefore be considered. As the current time limit imposed on relational tasks is 30 seconds, this window could be halved once mastery of a given level is achieved. It is hoped that by progressively shortening the time limit allocated the training stages, more rapid responding to relational tasks may be trained.

4.10 Conclusion

The current thesis aimed to assess (1) whether a relational skills training intervention could significantly increase scores on a traditional IQ assessment and (2) the relationship between relational ability (as measured by the RAI) and intellectual performance. Results from the second study indicate an extremely high level of correlation between relational ability and all seven scales and subscales of IQ, as assessed by the WAIS. This finding offers further support to previous analyses implicating the fundamental importance of relational responding to intelligence. Furthermore, the current findings also suggest that the RAI offers considerable utility as a proxy measure of intelligence.

The current investigation failed to identify a significant effect for group on the three main indices of IQ (Full, Verbal & Performance). However, analyses of variance indicated that the relational training resulted in a significantly greater efficacy in improving performance on the Verbal Comprehension subindex of IQ when compared to non-intervention. As this is first controlled examination of the effectiveness of the current relational training program using adults, the failure to find significant effects for Full IQ may be very important scientifically. This finding may indicate that relational skills training is not an effective means of improving intellectual performance as assessed by the three main indices of IQ. However, this conclusion may be premature due to the triangulation of well-established theory, considerable correlational work supporting the relevance of relational skill to intelligence, as well as previous research investigations (Cassidy et al. 2011, 2016) that would challenge this conclusion.

Following closer inspection of the increases in subtest scores, a clear discrepancy was found between groups in terms of the type of subtest that showed significant rises from baseline to post-intervention. Only the experimental group displayed significant rises in scores for subtests that did not include a timed element. On the other hand, virtually the entirety of IQ gains of the control group appear to be due to improved performance on timed subtests only. This qualitative difference in the nature of IQ rise is noteworthy due to the fact that untimed subtests are more stable, less vulnerable to practice and more reflective of intellectual performance when compared to timed subtests, It would follow therefore that, in comparison to timed subtests, performance on the untimed subtests may rely more heavily on relational ability. The current analysis is the first to detect such an effect for relational skills training, and therefore offers a significant contribution to the analysis of the relationship between relational responding and intelligence.

The current investigation also identified possible boundary conditions for the efficacy of relational training in raising IQ scores. Specifically, the current analysis provides a more detailed examination of the conditions within which previously reported effects may occur. For example, this investigative study found that high levels of baseline IQ and relational ability may adversely affect the impact of the current relational training program. In addition to the effect of starting IQ, age may also be considered a potential confounding factor due to the failure of the current analysis to replicate the training effects previously reported in child populations. While the current analysis identifies such variables as confounding factors, a more stringent and focused study of these variables is warranted because understanding them will form the very basis of a full understanding of the behavioural processes underlying the effects of relational skills interventions.

In conclusion, the current investigation cast some light on the possible processes underlying a relational skills training intervention. This began with an extensive correlational analysis of the relationship between IQ and relational skill levels. While significant and widespread IQ gains for individuals were not observed in the intervention study, the current study opened up definite avenues of further empirical enquiry that can help to more clearly explicate the processes involved in the training and impact of relational skill training interventions.

References

Abrami, P. C., Cholmsky, P., & Gordon, R. (2001). *Statistical analysis for the social sciences: An in- teractive approach*. Needham Heights, MA: Allyn & Bacon.

Abramson, C. I. (2013). Problems of Teaching the Behaviorist Perspective in the Cognitive Revolution. *Behavioral Sciences*, *3*(1), 55–71. http://doi.org/10.3390/bs3010055

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131,* 30–60.

ACT. (2008). Measuring college readiness: The national graduating class of 2008. Iowa City, IA: Author.

Aitkin, M., & Zuzovsky, R. (1994). Multilevel Interaction Models and their Use in the Analysis of Large‐scale School Effectiveness Studies. *School Effectiveness And School Improvement*, *5*(1), 45-73. http://dx.doi.org/10.1080/0924345940050104

Ali, A., Ambler, G., Strydom, A., Rai, D., Cooper, C., & McManus, S. et al. (2012). The relationship between happiness and intelligent quotient: the contribution of socio-economic and clinical factors. *Psychological Medicine*, 43(06), 1303-1312.

Anastasi, A. (1990). *Psychological testing (6th ed.).* Englewood Cliffs, NJ: Prentice-Hall.

Anderson, M. (2004). Sex differences in general intelligence. In R. L. Gregory (Ed.), *The Oxford companion to the mind.* Oxford: Oxford University Press.

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.

Andrews, G. & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology, 45,* 153-219.

Aron, A., & Aron, E. N. (2003). *Statistics for psychology* (3rd ed.). Englewood Cliffs, NJ: Prentice- Hall.

Aronson, J., Fried, C., & Good, C. (2002). Reducing the Effects of Stereotype Threat on African American College Students by Shaping Theories of Intelligence. *Journal Of Experimental Social Psychology*, *38*(2), 113-125. http://dx.doi.org/10.1006/jesp.2001.1491

Axelrod, B. N. (2002). Validity of the Wechsler Abbreviated Scale of Intelligence and other very short forms of estimating intellectual functioning, *Assessment, 9*, 17–23.

Baars, B. (1986). *The cognitive revolution in psychology*. New York: Guilford Press.

Baer, D., Peterson, R., & Sherman, J. (1967). The development of imitation by reinforcing behavioral similarity to a model1. *J Exp Anal Behav*, *10*(5), 405-416. http://dx.doi.org/10.1901/jeab.1967.10-405

Baer, J. (1993). *Creativity and divergent thinking*. Hillsdale, N.J.: L. Erlbaum Associates.

Baker, D., Eslinger, P., Benavides, M., Peters, E., Dieckmann, N., & Leon, J. (2015). The cognitive impact of the education revolution: A possible cause of the Flynn Effect on population IQ. *Intelligence*, *49*, 144-158. http://dx.doi.org/10.1016/j.intell.2015.01.003

Baker, D., Salinas, D., & Eslinger, P. (2012). An envisioned bridge: Schooling as a neurocognitive developmental institution. *Developmental Cognitive Neuroscience*, *2*, S6-S17. http://dx.doi.org/10.1016/j.dcn.2011.12.001

Baltes P., & Reinert, G. (1969) *Cohort effects in cognitive develop- ment in children as revealed by cross-sectional sequences.* Dev Psychol, 1, 169–77.

Barbera, M., & Rasmussen, T. (2007). *The verbal behavior approach*. London: Jessica Kingsley Publishers.

Barnes, D. (1994). Stimulus equivalence and relational frame theory. *The Psychological Record, 44,* 91-124.

Barnes, D., & Browne, M., & Smeets, P.M., & Roche, B. (1995). A transfer of functions and a conditional transfer of functions through equivalence relations in three to six year old children. *The Psychological Record, 45*, 405-430.

Barnes, D., Hegarty, N., & Smeets, P. (1997). Relating equivalence relations to equivalence relations: A relational framing model of complex human functioning. *The Analysis of Verbal Behavior, 14,* 1-27.

Barnes, D., & Roche, B. (1997). Relational frame theory and the experimental analysis of human sexuality. *Applied And Preventive Psychology*, *6*(3), 117-135. http://dx.doi.org/10.1016/s0962-1849(97)80001-7

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

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. (2001). The development of self and perspective-taking: A relational frame analysis. *Behavioral Development Bulletin*, *10*(1), 42-45. http://dx.doi.org/10.1037/h0100482

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., Regan, D., Barnes-Holmes, Y., Commins, S., Walsh, D., & Stewart, I. et al. (2005). Relating Derived Relations as a Model of Analogical Reasoning: Reaction Times and Event-Related Potentials. *J Exp Anal Behav*, *84*(3), 435-451. http://dx.doi.org/10.1901/jeab.2005.79-04

Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., & Walsh, D. et al. (2005). Derived Stimulus Relations, Semantic Priming, and Event-Related Potentials: Testing a Behavioral Theory of Semantic Networks. *J Exp Anal Behav*, *84*(3), 417-433. http://dx.doi.org/10.1901/jeab.2005.78-04

Basso, M. R., Bornstein, R. A., & Lang, J. M. (1999). Practice effects on commonly used measures of executive function across twelve months. *The Clinical Neuropsychologist*, *13*(3), 283-292.

Basso, M. R., Carona, F. D., Lowery, N., & Axelrod, B. N. (2002). Practice effects on the WAIS-III across 3 and 6-month intervals. *The Clinical Neuropsychologist, 16(1),* 57-63.

Bates, B.T., Zhang, S., Dufek, J.S. & Chen, F.C. (1996). The effects of sample size and variability on the correlation coefficient. *Medicine and Science in Sports and Exercise*, *28*(3), 386-391.

Baxendale, S. (2011). IQ and Ability Across the Adult Life Span. *Applied Neuropsychology, 18,* 164-167.

Bayley, N. (1940). *Studies in the development of young children*. Berkeley: University of California Press.

Beckwith, J. (1999). Simplicity and complexity: Is IQ ready for genetics? *Current Psychology of Cognition, 18*, 161-169.

Bennett, G., Seashore, H,. & Wesman, A. (1990). *Manual for the Differential Aptitude Tests*. San Antonio: Psychological Corporation.

Berens, N., & Hayes, S. (2007). Arbitrarily Applicable Comparative Relations: Experimental Evidence for a Relational Operant. *J Appl Behav Anal*, *40*(1), 45-71. http://dx.doi.org/10.1901/jaba.2007.7-06

Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin, 105*, 467–477.

Bevan, W. (1982). A sermon of sorts in three plus parts. *American Psychologist*, *37*(12), 1303-1322. http://dx.doi.org/10.1037//0003-066x.37.12.1303

Bevan, W. (1991). Contemporary psychology: A tour inside the onion. *American Psychologist, 46,* 475– 483.

Bevan, W, & Kessel, F. (1994). Plain truths and home cooking: Thoughts on the making and remaking of psychology. *American Psychologist, 49*, 505-509.

Black, E. (2003). *War Against the Weak: Eugenics and America's Campaign to Create a Master Race.* New York: Four Walls Eight Windows.

Block, N. (1981). Psychologism and behaviorism. *Philosophical Review* XC(1), 5–43

Bradway, K. (1944). IQ Constancy on the Revised Stanford-Binet from the Preschool to the Junior High School Level. *The Pedagogical Seminary And Journal Of Genetic Psychology*, *65*(2), 197-217. http://dx.doi.org/10.1080/08856559.1944.10533295

Binet, A. (1904a). Nos commissions de travail [Our working committee]. *Bulletin de la Société Libre pour l'Etude Psychologique de l'Enfant*, *4*(14), 337–346.

Binet, A. (1904b). Sommaire des travaux en cours à la Société de psychologie de l'enfant [Summary of the work in progress at the society of child psychology]. *L'Année Psychologique*, *10*, 116–130.

Binet, A., & Henri, V. (1895b). La mémoire des phrases [Memory phrases]. *L'Année Psychologique*, *1*, 24–59.

Binet, A., & Henri, V. (1896). La psychologie individuelle [Individual psychology]. *L'Année Psychologique*, *2*, 411–465.

Binet, A., & Simon, T. (1905). Méthode nouvelle pour le diagnostic du niveau intellectuel des anormaux [A new method for the diagnosis of the abnormal intellectual levels]. *L'Année Psychologique, 11,* 191-244.

Birren, J., & Morrison, D. (1961). Analysis of the Wais Subtests in Relation to Age and Education. *Journal Of Gerontology*, *16*(4), 363-369. http://dx.doi.org/10.1093/geronj/16.4.363

Boring, E. (1923). *Intelligence as the Tests Test It. New Republic,* 36, 35-37.

Buck v. Bell , 274 U.S. 200 (1927).

Bugg, J., Zook, N., DeLosh, E., Davalos, D., & Davis, H. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain And Cognition*, *62*(1), 9-16. http://dx.doi.org/10.1016/j.bandc.2006.02.006

Buschkuehl, M., Jaeggi, S., Hutchison, S., Perrig-Chiello, P., Däpp, C., & Müller, M. et al. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology And Aging*, *23*(4), 743-753. http://dx.doi.org/10.1037/a0014342

California Department of Education (1997). *Best Practices for Designing and Delivering Effective Programs for Individuals with Autistic Spectrum Disorders:*

Recommendations of the Collaborative Work Group on Austistic Spectrum Disorders. Sacramento: California Department of Education.

Carlsson, M., Dahl, G., Öckert, B., & Rooth, D. (2015). The Effect of Schooling on Cognitive Skills. *Review Of Economics And Statistics*, *97*(3), 533-547. http://dx.doi.org/10.1162/rest_a_00501

Carpentier, F., Smeets, P., & Barnes-Holmes, D. (2003). Matching unrelated stimuli with same discriminative functions: training order effects. *Behavioural Processes*, *60*(3), 215-226. http://dx.doi.org/10.1016/s0376-6357(02)00124-9

Carpentier, F., Smeets, P. M., Barnes-Holmes, D., & Stewart, I. (2004). Matching derived functionally-same relations: Equivalence-equivalence and classical analogies. *The Psychological Record, 54*, 255-273.

Cassidy, S., Roche, B., Colbert, D., Stewart, I. & Grey, I. (2016) A Relational Frame Skills Training Intervention to Increase General Intelligence and Scholastic Aptitude. *Learning and Individual Differnces, 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 human intelligence. *European Journal of Behavior Analysis, 11*, 37-51.

Catania, A. C. (1996). On the origins of behavior structure. In T. R. Zentall & P. M. Smeets (Eds.) *Stimulus class formation in humans and animals* (pp. 3- 12). North Holland: Elsevier Science.

Catania, A. C. (1998). *Learning* (4th ed.). New Jersey, Prentice-Hall.

Catron, D. W., & Thompson, C. C. (1979). Test–retest gains in WAIS scores after four retest intervals. *Journal of Clinical Psychology, 35*, 352–357.

Cattell, R. (1971). *Abilities*. Boston: Houghton Mifflin.

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

Cheramie, G. M., Stafford, M. E., Mulkey, B., Streich, K., & Tagle, L. (2014). School Psychologists' perceptions of due process hearings. *Research and Practice in the Schools, 2*(1), 41-52.

Chomsky, N. (1959). Review of Skinner's Verbal Behavior. *Language, 35*, 26-58.

Cliffordson, C., & Gustafsson, J. E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence,* 36, 143-152.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early Intensive Behavioral Treatment. *Journal Of Developmental & Behavioral Pediatrics*, *27*(Supplement 2), S145-S155. http://dx.doi.org/10.1097/00004703-200604002-00013

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed). Hillsdale: Erlbaum.

Collins, J. L. (1982). *Self-efficacy and ability in achievement behavior*. Paper presented at the meeting of the American Educational Research Association, New York.

Colom, R., Abad, F. J., Quiroga, M. A., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence., 36,* 584–606.

Covin, T. (1977). Stability of the WISC-R for 9-year-olds with learning difficulties. *Psychological Reports*, *40*(3c), 1297-1298. http://dx.doi.org/10.2466/pr0.1977.40.3c.1297

Creasy, M. (1959). Some Criticisms of Factor Analysis with Suggestions for Alternative Methods. *The British Journal Of Psychiatry*, *105*(440), 755-761. http://dx.doi.org/10.1192/bjp.105.440.755

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt Brace Jovanovich.

Cronbach, L. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671-684. http://dx.doi.org/10.1037/h0043943

Cullinan, V., Barnes-Holmes, D., & Smeets, P.M. (2001). A precursor to the relational evaluation procedure: Searching for the contextual cues that control equivalence responding. *Journal of the Experimental Analysis of Behavior; 76*(3), 339–349.

Cunningham, W., Clayton, V., & Overton, W. (1975). Fluid and Crystallized Intelligence in Young Adulthood and Old Age. *Journal Of Gerontology*, *30*(1), 53-55. http://dx.doi.org/10.1093/geronj/30.1.53

Dale, R. (2005). Cognitive and behavioral approaches to language acquisition: Conceptual and empirical intersections. *Behavior Analyst Today, 5*, 336–359.

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

Deary, I., Whalley, L., & Crawford, J. (2004). An instantaneous measure of a lifetime's cognitive change. *Intelligence, 32*, 113-119.

Deary, I., Whiteman, M., Starr, J., Whalley, L., & Fox, H. (2004). The Impact of Childhood Intelligence on Later Life: Following Up the Scottish Mental Surveys of 1932 and 1947. *Journal Of Personality And Social Psychology*, *86*(1), 130-147. http://dx.doi.org/10.1037/0022-3514.86.1.130

De Bruin, G.P. (2004). Problems with the factor analysis of items: Solutions based on item response theory and item parceling. *South African Journal of Industrial Psychology, 30,* 16–26.

De Houwer, J. (2011). Evaluative conditioning: Methodological considerations. In K. C. Klauer, C. Stahl, & A. Voss (Eds.) *Cognitive methods in social psychology* (pp. 124-147). New York: Guilford.

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.

De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional-cognitive framework for attitude research. *European Review Of Social Psychology*, *24*(1), 252-287. http://dx.doi.org/10.1080/10463283.2014.892320

Diamond, A., & Lee, K. (2011). Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old. *Science*, *333*(6045), 959-964. http://dx.doi.org/10.1126/science.1204529

Dickens, W., & Flynn, J. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, *108*(2), 346-369. http://dx.doi.org/10.1037/0033-295x.108.2.346

Dickinson, M., & Hiscock, M. (2010). Age-related IQ decline is reduced markedly after adjustment for the Flynn effect. *Journal Of Clinical And Experimental Neuropsychology*, *32*(8), 865-870. http://dx.doi.org/10.1080/13803391003596413

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. http://dx.doi.org/10.1016/j.rasd.2014.05.005

Dixon, M. R. (2014a). *The PEAK relational training system: Direct training module*. Carbondale: Shawnee Scientific Press.

Dixon, M. R. (2014b). *The PEAK relational training system: Generalization module*. Carbondale: Shawnee Scientific Press.

Dodrill, C. B., & Troupin, A. S. (1975). Effects of repeated administrations of a comprehensive neuropsychological battery among chronic epileptics. *Journal of Nervous and Mental Disease*, *161*(3), 185-190

Dollard, J., & Miller, N. (1950). *Personality and psychotherapy*. New York: McGraw-Hill.

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

Dymond, S., & Barnes, D. (1994). A transfer of self-discrimination response functions through equivalence relations. *Journal of the Experimental Analysis of Behavior, 62*(2), 251-267. http://dx.doi.org/10.1901/jeab.1994.62-251

Dymond, S., & Roche, B. (2013). *Advances in relational frame theory*. Oakland, CA: Context Press.

Ebmeier, H., & Good, T. (1979). The effects of instructing teachers about good teaching on the mathematics achievement of fourth grade students. *American Educational Research Journal*, 16(1), 1-16.

Educational Testing Services (2014). *IQ Estimates by College Major*. Retrieved from www.statisticbrain.com/iq-estimates-by-intended-college-major/

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive Behavioral Treatment at School for 4- to 7-Year-Old Children with Autism: A 1-Year Comparison Controlled Study. *Behavior Modification*, *26*(1), 49-68. http://dx.doi.org/10.1177/0145445502026001004

Elliott, C. D. (1990). *Differential Ability Scales*. San Antonio, TX: The Psychological Corporatation.

Engle, R., Tuholski, S., Laughlin, J., & Conway, A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal Of Experimental Psychology: General*, *128*(3), 309-331. http://dx.doi.org/10.1037/0096-3445.128.3.309

Ertel, S. (2013). *Factor analysis*. Göttingen: Universitätsverlag Göttingen.

Esher, F., Raven, I., & Earl, C. (1942). Discussion on testing intellectual capacity in adults. *Proceedings of the Royal Society of Medicine,* 35, 779-785.

Estevis, E., Basso, M.R. & Combs, D. (2012). Effects of Practice on the Wechsler Adult Intelligence Scale-IV Across 3- and 6-Month Intervals, *The Clinical Neuropsychologist*, 26:2, 239-254.

Falch, T., & Sandgren Massih, S. (2010).The effect of education on cognitive ability. *Economic Inquiry*, *49*(3), 838-856. http://dx.doi.org/10.1111/j.1465-7295.2010.00312.x

Fancher, R. (1985). Spearman's original computation of g: A model for Burt?. *British Journal Of Psychology*, *76*(3), 341-352. http://dx.doi.org/10.1111/j.2044-8295.1985.tb01957.x

Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York: W. W. Norton & Company.

Fattu, N., Mech, E., & Kapos, E. (1954). Some statistical relationships between selected response dimensions and problem-solving proficiency. *Psychological Monographs: General And Applied*, *68*(6), 1-23. http://dx.doi.org/10.1037/h0093671

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53*, 345-358.

Fletcher, R.B. & Hattie, J. (2011). Intelligence and intelligence testing. New York: Routledge: Taylor & Francis Group.

Flynn, J. (1998). Rising IQ Scores: Implications for the Elderly. *Australasian Journal On Ageing*, *17*(3), 106-107. http://dx.doi.org/10.1111/j.1741-6612.1998.tb00047.x

Flynn, J. (2007). *What is intelligence?*. Cambridge, UK: Cambridge University Press.

Flynn, J. (2009). *What is intelligence?*. Cambridge, UK: Cambridge University Press.

Fowler, R. (1990). Psychology: The core discipline. *American Psychologist*, *45*(1), 1-6. http://dx.doi.org/10.1037//0003-066x.45.1.1

Freeman, F. (1934). II. Contributions from Psychology. *Research Quarterly. American Physical Education Association*, *5*(sup4), 29-30. http://dx.doi.org/10.1080/23267402.1934.10761669

Galizio, M., Stewart, K., & Pilgrim, C. (2001). Clustering in artificial categories: An equivalence analysis. *Psychonomic Bulletin & Review, 8*, 609–614.

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Gardner, H. (1993). *Multiple intelligences*. New York, NY: Basic Books.

Gardner, H. (1983). *Frames of mind*. New York: Basic Books.

Gelb, S., (1986, April) *From moral imbecility to maladaptive behaviour: The social construction of educable mental retardation*. Paper presented at the annual meeting of the American Education Research Association, San Francisco, CA.

Gentner, D., & Loewenstein, J. (2002). Relational language and relational thought. In J. Byrnes, & E. Amsel (Eds.), *Language, Literacy, and Cognitive Development* (pp. 87-120). Mahwah, NJ: LEA.

Gentner, D., & Rattermann, M. J. (1991) . Language and the career of similarity . In S. A. Gelman & J. P. Byrnes (Eds .), *Perspectives on thought and language : Interrelations in development* (pp.225-277). London: Cambridge University Press.

Ghisletta, P., & Lindenberger, U. (2004). Static and Dynamic Longitudinal Structural Analyses of Cognitive Changes in Old Age. *Gerontology*, *50*(1), 12-16. http://dx.doi.org/10.1159/000074383

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature*, *410*(6831), 930-933. http://dx.doi.org/10.1038/35073582

Glenberg, A. M. (1996). *Learning from data: An introduction to statistical reasoning*. Mahwah, NJ: Erlbaum.

Goddard, H. H. (1912). *The Kallikak Family: A study in the heredity of feeble-mindedness*. New York: Macmillan.

Goddard, H.H. (1914). F*eeble-mindedness: Its causes and consequences*: New York: Macmillan.

Goddard, H.H. (1917) Mental tests and immigrants. *Journal of Delinquency, 2,* 243–277.

Gomez, S., Lopez, F., Martin, C.B., Barnes-Holmes, Y., Barnes-Holmes, D. (2007) Exemplar training and a derived transformation of functions in accordance with symmetry and equivalence. *The Psychological Record, 57*, 273-293.

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *Journal of Experimental Education, 74*, 251–266.

Gore, N. J., Barnes-Holmes, Y., & Murphy, G. (2010). The Relationship between Intellectual Functioning and Relational Perspective-Taking. *International Journal of Psychology & Psychological Therapy, 10,* 1-17.

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment, 68*(3), 532-560.

Goswami, U. (1991). Analogical Reasoning: What Develops? A Review of Research and Theory. *Child Development*, *62*(1), 1. http://dx.doi.org/10.2307/1130701

Gottfredson, L. (1998). *The general intelligence factor*. New York: Scientific American, Inc.

Gottfredson, L. (2003a). Dissecting practical intelligence: Its claims and evidence. *Intelligence, 31,* 343–397. doi:10.1016/S0160-2896 (02)00085-5

Gould, S. (1981). *The mismeasure of man*. New York: Norton.

Gregory, R. (1994). Seeing intelligence. In J. Khalfa (Ed). *What is intelligence?* (pp. 13-26). Cambridge, England: Cambridge University Press.

Groth-Marnat, G. (1984). *Handbook of psychological assessment.* New York: Van Nostrand Reinhold.

Guilford, J. (1936). *Psychometric methods*. New York: McGraw-Hill Book Company, Inc.

Guttman, L. (1954). An Outline of Some New Methodology for Social Research. *Public Opinion Quarterly*, *18*(4), 395. http://dx.doi.org/10.1086/266532

Hale, S., & Jansen, J. (1994). Global processing-time coefficient characterize individual and group Ddifferences in cognitive speed. *Psychological Science*, *5*(6), 384-389. http://dx.doi.org/10.1111/j.1467-9280.1994.tb00290.x

Halford, G. S. (1982). *The development of thought.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Harmon, K., Strong, R., & Pasnak, R. (1982). Relational responses in tests of transposition with Rhesus monkeys. *Learning And Motivation*, *13*(4), 495-504. http://dx.doi.org/10.1016/0023-9690(82)90006-6

Härnqvist, K. (1968). RELATIVE CHANGES IN INTELLIGENCE FROM 13 TO 18: Results. *Scandinavian Journal Of Psychology*, *9*(1), 65-82. http://dx.doi.org/10.1111/j.1467-9450.1968.tb00519.x

Harris, S., & Handleman, J. (1994). *Preschool education programs for children with autism*. Austin, Tex.: PRO-ED.

Hartlage, L., C. & Steele, C., T. (1977) *WISC and WISC-R correlates of academic achievement.* Psychol. Schs., 14,15–18.

Hattie, J. (2009). *Visible learning*. London: Routledge.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19-40). Reno, NV: Context Press.

Hayes, S. C. (1994). Relational frame theory as a behavioral approach to verbal events. Chapter in S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 9-30). Reno, NV: Context Press.

Hayes, S. (1994). *Behavior analysis of language and cognition*. Reno, NV: Context Press.

Hayes, S., & Barnes, D. (1997). Analyzing Derived Stimulus Relations Requires More Than The Concept Of Stimulus Class. *J Exp Anal Behav*, *68*(2), 235-244. http://dx.doi.org/10.1901/jeab.1997.68-235

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

Hayes, S.C., Fox, E., Gifford, E.V., Wilson, K.G., Barnes-Holmes, D. & Healy, O. (2001). Derived relational responding as learned behavior. In. S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of language and cognition*. New York: Plenum Press.

Healy, O., Barnes-Holmes, D., & Smeets, P. (2000). Derived relational responding as generalized operant behavior. *J Exp Anal Behav*, *74*(2), 207-227. http://dx.doi.org/10.1901/jeab.2000.74-207

Herrnstein, R., & Murray, C. (1994). *The bell curve*. New York: Free Press.

Hopkins, K. D. (1998). *Educational and psychological measurement and evaluation* (8th ed.). Needham Heights, MA: Allyn & Bacon.

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

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65,* 185-241.

Horton, A.M. (1992). Neuropsychological practice effects and age: A brief note. *Perceptual and Motor Skills*, *75*, 257–258.

Hossiep, R., Hasella, M., & Turck, D. (1999). *BOMAT - advanced*. Göttingen [u.a.]: Hogrefe, Verl. für Psychologie.

Houwer, J. (2011). Why the Cognitive Approach in Psychology Would Profit From a Functional Approach and Vice Versa. *Perspectives On Psychological Science*, *6*(2), 202-209. http://dx.doi.org/10.1177/1745691611400238

Howe, M. (1989). Separate skills or general intelligence: The autonomy of human abilities. *British Journal Of Educational Psychology*, *59*(3), 351-360. http://dx.doi.org/10.1111/j.2044-8279.1989.tb03109.x

Howe, M. (1990a) *The origins of exceptional abilities.* Oxford:Blackwell.

Howlin, P., Magiati, I., & Charman, T. (2009). Systematic Review of Early Intensive Behavioral Interventions for Children With Autism. *American Journal On Intellectual And Developmental Disabilities*, *114*(1), 23. http://dx.doi.org/10.1352/2009.114:23;nd41

Hull, C. (1920). *Quantitative aspects of the evolution of concepts*. Princeton, N.J.: Psychological review Company.

Husen, T., & Tuijnman, A. (1991). The Contribution of Formal Schooling to the Increase in Intellectual Capital. *Educational Researcher*, *20*(7), 17-25. http://dx.doi.org/10.3102/0013189x020007017

Irwin, D. O. (1966). Reliability of the WISC. *Journal of Educational Measurement, 3,* 287-292.

Jacobs, J., Lanza, S., Osgood, D., Eccles, J., & Wigfield, A. (2002). Changes in Children's Self-Competence and Values: Gender and Domain Differences across Grades One through Twelve. *Child Development*, *73*(2), 509-527. http://dx.doi.org/10.1111/1467-8624.00421

Jacobson, J., Mulick, J., & Green, G. (1998). Cost–benefit estimates for early intensive behavioral intervention for young children with autism—general model and single state case. *Behavioral Interventions*, *13*(4), 201-226. http://dx.doi.org/10.1002/(sici)1099-078x(199811)13:4<201::aid-bin17>3.3.co;2-i

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. http://dx.doi.org/10.1073/pnas.0801268105

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

Jensen, A. (1969). How Much Can We Boost IQ and Scholastic Achievement. *Harvard Educational Review*, *39*(1), 1-123. http://dx.doi.org/10.17763/haer.39.1.l3u15956627424k7

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

Jensen, A. (1998). *The g factor*. Westport, Conn.: Praeger.

Jones, H., & Bayley, N. (1941). The Berkeley Growth Study. *Child Development*, *12*(2), 167. http://dx.doi.org/10.2307/1125347

Joseph, J. (1998). The equal environment assumption of the classical twin method: A critical analysis. *Journal of Mind and Behavior,19,* 325-358.

Journal of Educational Psychology (1921). Intelligence and Its Measurement: A Symposium. *Journal Of Educational Psychology*, *12*(3), 123-147. http://dx.doi.org/10.1037/h0076078

Judson, A., Cofer, C., & Gelfand, S. (1956). Reasoning as an associative process: II. "Direction" in problem solving as a function of prior reinforcement of relevant responses. *Psychological Reports*, *2*(3), 501-507. http://dx.doi.org/10.2466/pr0.1956.2.3.501

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. http://dx.doi.org/10.1016/0022-4405(88)90040-4

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 66–71.

Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence.* Boston: Allyn and Bacon.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman Brief Intelligence Test* (2nd ed.). Bloomington, MN: Pearson.

Kaufman, A., Johnson, C., & Xin Liu,. (2008). A CHC Theory-Based Analysis of Age Differences on Cognitive Abilities and Academic Skills at Ages 22 to 90 Years. *Journal Of Psychoeducational Assessment*, *26*(4), 350-381. http://dx.doi.org/10.1177/0734282908314108

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*, 123–138.

Kimble, G. (1994). A frame of reference for psychology. *American Psychologist*, *49*(6), 510-519. http://dx.doi.org/10.1037//0003-066x.49.6.510

Kishita, N., Ohtsuki, T., & Stewart, I. (2013). The Training and Assessment of Relational Precursors and Abilities (TARPA): A follow-up study with typically developing children. *Journal Of Contextual Behavioral Science*, *2*(1-2), 15-21. http://dx.doi.org/10.1016/j.jcbs.2013.01.001

Kohs, S. C. (1920). The Block-Design Tests. *Journal of Experimental Psychology, 3*(5), 357-376.

Kraut, A. M. (1990). Immigrant attitudes toward physicians in America. *Journal of the American Medical Association, 263*(April 4), 1807–1811.

Krawczyk, D., McClelland, M., Donovan, C., Tillman, G., & Maguire, M. (2010). An fMRI investigation of cognitive stages in reasoning by analogy. *Brain Research*, *1342*, 63-73. http://dx.doi.org/10.1016/j.brainres.2010.04.039

Krawczyk, D., Michelle McClelland, M., & Donovan, C. (2011). A hierarchy for relational reasoning in the prefrontal cortex. *Cortex*, *47*(5), 588-597. http://dx.doi.org/10.1016/j.cortex.2010.04.008

Kwesiga, C.J. (2002). Women’s access to higher education in Africa: Uganda’s experience. Kampala: Fountain publishers Ltd.

Kyllonen, P., Lohman, D., & Woltz, D. (1984). Componential modeling of alternative strategies for performing spatial tasks. *Journal Of Educational Psychology*, *76*(6), 1325-1345. http://dx.doi.org/10.1037/0022-0663.76.6.1325

Layzer, D. (1972). A rejoinder to Professor Herrnstein's comments. *Cognition*, *1*(4), 423-426. http://dx.doi.org/10.1016/0010-0277(72)90007-8

Legg, S., & Hutter, M. A (2007) A collection of definitions of intelligence. In B. Goertzel & P. Wang (Eds.) *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*. Amsterdam: IOS Press.

Lomax, A., & Berkowitz, N. (1972). The Evolutionary Taxonomy of Culture. *Science*, *177*(4045), 228-239. http://dx.doi.org/10.1126/science.177.4045.228

Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior, 69*, 327-354.

Lovaas, O. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55,* 3-9.

Lovaas, O., & Wright, S. (2006). "A reply to recent public critiques...". *Journal Of Early And Intensive Behavior Intervention*, *3*(2), 234-236. http://dx.doi.org/10.1037/h0100335

Lovaas, O., Smith, T., & McEachin, J. (1989). Clarifying comments on the young autism study: Reply to Schopler, Short, and Mesibov. *Journal Of Consulting And Clinical Psychology*, *57*(1), 165-167. http://dx.doi.org/10.1037//0022-006x.57.1.165

Lovecky, D. V. (2004). *Different Minds: Gifted Children With AD/HD, Asperger Syndrome, and Other Learning Deficits*. London: Jessica Kingsley.

Luciano, C., Becerra, I., & Valverde, M. (2007). The Role of Multiple-Exemplar Training and Naming in Establishing Derived Equivalence in an Infant. *J Exp Anal Behav*, *87*(3), 349-365. http://dx.doi.org/10.1901/jeab.2007.08-06

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

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

MacIntyre, A. (1985b). How psychology makes itself true—or false. In S. Koch & D. E. Leary (Eds.), *A century of psychology as science*. New York: McGraw-Hill.

Mackay, A., Connor, L., Albert, M., & Obler, L. (2002). Noun and verb retrieval in healthy aging. *Journal Of The International Neuropsychological Society*, *8*(6), 764-770. http://dx.doi.org/10.1017/s1355617702860040

Mackey, A., Miller Singley, A., & Bunge, S. (2013). Intensive Reasoning Training Alters Patterns of Brain Connectivity at Rest. *Journal Of Neuroscience*, *33*(11), 4796-4803. http://dx.doi.org/10.1523/jneurosci.4141-12.2013

Mackey, A.P., Whitaker, K.J., & Bunge, S.A. (2012). Experience-dependent plasticity in white matter microstructure: reasoning training alters structural connectivity. *Frontiers in Neuroanatomy*, 6(32), 1-9

Maher, B. A. (1985). Underpinnings of today's chaotic diversity. *International Newsletter of Paradigmatic Psychology, 1,* 17–19.

Maltzman, I. (1955). Thinking: from a behavioristic point of view. *Psychological Review*, *62*(4), 275-286. http://dx.doi.org/10.1037/h0041818

Manheim, H. L. (1977) *Sociological Research.* Georgetown: Irwin Dorsey.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: Williams & Williams Company.

Manard, M., Carabin, D., Jaspar, M., & Collette, F. (2014). Age-related decline in cognitive control: the role of fluid intelligence and processing speed. *BMC Neuroscience*, *15*(1), 7. http://dx.doi.org/10.1186/1471-2202-15-7

Marchman, V., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *0*(0), 080130202934054-???. http://dx.doi.org/10.1111/j.1467-7687.2007.00671.x

Martin, T. A., Donders, J., & Thompson, E. (2000). Potential of and problems with new measures of psychometric intelligence after traumatic brain injury. *Rehabilitation Psychology, 45,* 402–408.

Martinez, M. E. (2000). *Education as the cultivation of intelligence*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Matthews, G., Zeidner, M., & Roberts, R. (2002). *Emotional intelligence*. Cambridge, Mass.: MIT Press.

Maurice, C., Green, G., & Luce, S. (1996). *Behavioral intervention for young children with autism*. Austin, Tex.: Pro-Ed.

Mayr, E. (2001). *What evolution is*. New York: Basic Books.

McCaffrey, R. J., Duff, K. & Westervelt, H. J. (2000) Practitioner's guide to evaluating change with neuropsychological assessment instruments. New York: Kluwer/Plenum.

McCaffrey, R. J., & Westervelt, H. J. (1995). Issues associated with repeated neuropsychological assessments. *Neuropsycholical Review, 5*(3), 203–221.

MacCall, J. J. (1970). Economics of information and job search. *Quarterly Journal of Economics*, *84*, 113–126.

McCall, R. B. (1977). Childhood IQs as predictors of adult educational and occupational status. *Science*, *197*, 482-483.

McCall, R. B., Appelbaum, M. I., & Hogarty, P. S. (1973). *Developmental changes in mental performance.* Chicago: Published by the University of Chicago Press for the Society for Research in Child Development.

McEachin, J., Smith, T., & Lovaas, O. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation. 97*, 359–372.

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

McLean, J., Kaufman, A., & Reynolds, C. (1989). Base rates of WAIS-R subtest scatter as a guide for clinical and neuropsychological assessment. *Journal Of Clinical*

Psychology, *45*(6), 919-926. http://dx.doi.org/10.1002/1097-4679(198911)45:6<919::aid-jclp2270450614>3.0.co;2-j

Melby-Lervåg, M. & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology,49*(2)*,* 270-291. doi: 10.1037/a0028228

Metz, Brittany Leigh, "A Comparison of the WJ-III Test of Cognitive Abilities and the WAIS-III" (2005). Theses, Dissertations and Capstones. Paper 762.

Mitrushina, M., & Satz, P. (1991). Effect of repeated administration of a neuropsy-chological battery in the elderly. *Journal of Clinical Psychology 47(6)*, 790-801.

Mokken, (1970) *A theory and procedures of scale analysis: with applications in political research*. The Hague: Mouton.

Moody, D. (2009). Can intelligence be increased by training on a task of working memory?. *Intelligence*, *37*(4), 327-328. http://dx.doi.org/10.1016/j.intell.2009.04.005

Moran, L., Stewart, I., McElwee, J., & Ming, S. (2010). Brief Report: The Training and Assessment of Relational Precursors and Abilities (TARPA): A Preliminary Analysis. *J Autism Dev Disord*, *40*(9), 1149-1153. http://dx.doi.org/10.1007/s10803-010-0968-0

Mortensen E. L., & Kleven M. (1993). A WAIS longitudinal study of cognitive development during the lifespan from ages 50 to 70. *Developmental Neuropsychology,* 9, 115–130.

Murphy, C., & Barnes-Holmes, D. (2009). Derived more–less relational mands in children diagnosed with autism. *Journal of Applied Behavior Analysis*, *42*(2), 253-268. http://dx.doi.org/10.1901/jaba.2009.42-253

Murphy, C., & Barnes-Holmes, D. (2010). Establishing five derived mands in three adolescent boys with autism. *Journal of Applied Behavior Analysis*, *43*(3), 537-541. http://dx.doi.org/10.1901/jaba.2010.43-537

Myers, S., & Johnson, C. (2007). Management of Children With Autism Spectrum Disorders. *Pediatrics*, *120*(5), 1162-1182. http://dx.doi.org/10.1542/peds.2007-2362

Myerson, J., Hale, S., Zheng, Y., Jenkins, L., & Widaman, K. (2003). The difference engine: A model of diversity in speeded cognition. *Psychonomic Bulletin & Review*, *10*(2), 262-288. http://dx.doi.org/10.3758/bf03196491

Naglieri, J. A., & Bardos, A. N. (1997). *General Ability Scale for Adults (GAMA).* Minnetonka, NM: National Computer Systems.

Neal, D., & Johnson, W. (1996). The Role of Premarket Factors in Black-White Wage Differences. *Journal Of Political Economy*, *104*(5), 869-895. http://dx.doi.org/10.1086/262045

Neisser, U., Boodoo, G., Bouchard, T., Boykin, A., Brody, N., & Ceci, S. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*(2), 77-101. http://dx.doi.org/10.1037/0003-066x.51.2.77

Nicolas, S., Andrieu, B., Croizet, J. C., Sanitioso, R. B., & Burman, J. T. (2013). Sick? or Slow? On the origins of intelligence as a psychological object. *Intelligence, 41*(5), 699-711.

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. http://dx.doi.org/10.1037/a0029772

O'Connor, J., Rafferty, A., Barnes-Holmes, D., & Barnes-Holmes, Y. (2009). The role of verbal behavior, stimulus nameability, and familiarity on the equivalence performances of autistic and normally developing children. *The Psychological Record, 59,* 53–74.

O'Hora, D., Pelaez, M., & Banres-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., Banres-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'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 Behavior. *International Journal of Psychology and Psychological Therapy, 9*, 1 - 17.

Perkins, F. T. (1931). A further study of configurational learning in goldfish. *Journal of Experimental Psychology, 14*, 508-538.

Petrill, S., Saudino, K., Cherny, S., Emde, R., Fulker, D., Hewitt, J., & Plomin, R. (1998). Exploring the Genetic and Environmental Etiology of High General Cognitive Ability in Fourteen- to Thirty-Six-Month-Old Twins. *Child Development*, *69*(1), 68-74. http://dx.doi.org/10.1111/j.1467-8624.1998.tb06133.x

Pike, A., Reiss, D., Hetherington, E., & Plomin, R. (1996). Using MZ Differences in the Search for Nonshared Environmental Effects. *J Child Psychol & Psychiat*, *37*(6), 695-704. http://dx.doi.org/10.1111/j.1469-7610.1996.tb01461.x

Plant, W., & Richardson, H. (1958). The IQ of the average college student. *Journal Of Counseling Psychology*, *5*(3), 229-231. http://dx.doi.org/10.1037/h0040731

Plomin, R. (1990). *Nature and nurture: An introduction to human behavioral genetics*. Pacific Grove, CA: Brooks/Cole.

Plomin, R. (1997). *Behavioral genetics*. New York: W.H. Freeman.

Plomin, R., & DeFries, J. (1998). The Genetics of Cognitive Abilities and Disabilities. *Sci Am*, *278*(5), 62-69. http://dx.doi.org/10.1038/scientificamerican0598-62

Pryor, K., Haag, R., & O'Reilly, J. (1969). The creative porpoise: training for novel behavior1. *J Exp Anal Behav*, *12*(4), 653-661. http://dx.doi.org/10.1901/jeab.1969.12-653

Psychological Corporation. (1999). *Wechsler Abbre- viated Scale of Intelligence.* San Antonio, TX: Author.

Putnam, H. (1975b) *Mind, Language and Reality, Philosophical Papers, volume 2*. Cambridge: Cambridge University Press.

Quereshi, M. (1968). Practice effects on the WISC subtest scores and IQ estimates. *Journal Of Clinical Psychology*, *24*(1), 79-85. http://dx.doi.org/10.1002/1097-4679(196801)24:1<79::aid-jclp2270240125>3.0.co;2-c

Ramsden, S., Richardson, F. M., Josse, G., Shakeshaft, C., Seghier, M. L., & Price, C. J. (2013). The influence of reading ability on subsequent changes in verbal IQ in the teenage years. *Developmental Cognitive Neuroscience*, *6*, 30–39. http://doi.org/10.1016/j.dcn.2013.06.001

Rapport, L. J., Brines, D. B., Theisen, M. E., & Axelrod, B. N. (1997). Full scale IQ as mediator of practice effects: The rich get richer. *The Clinical Neuropsychologist*, *11*(4), 375-380.

Raven, J. C. (1990). *Advanced Progressive Matrices: Sets I, II*. Oxford: Oxford University Press.

Rechsly, D., & Rechsly, J. (1979). *Validity of the WISC-R factor scores in predicting achievement and attention for four sociocultural groups.* Journal of School Psychology. 17, 355-359.

Reddy, A. (2008) The Eugenic Origins of IQ Testing Implications for Post-Atkins Litigation. *DePaul Law Review, 57*(3) 667-678.

Reichow, B., & Wolery, M. (2008). Comprehensive Synthesis of Early Intensive Behavioral Interventions for Young Children with Autism Based on the UCLA Young Autism Project Model. *J Autism Dev Disord*, *39*(1), 23-41. http://dx.doi.org/10.1007/s10803-008-0596-0

Reise, S.P. (1999). Personality measurement issues viewed through the eyes of IRT. In S.E. Embretson & S.L. Hersberger (Eds), *The new rules of measurement* (pp. 219-242). Mahwah, NJ: Erlbaum.

Remington, B., Hastings, R., Kovshoff, H., degli Espinosa, F., Jahr, E., & Brown, T. et al. (2007). Early Intensive Behavioral Intervention: Outcomes for Children With Autism and Their Parents After Two Years. *American Journal On Mental Retardation*, *112*(6), 418. http://dx.doi.org/10.1352/0895-8017(2007)112[418:eibiof]2.0.co;2

Reese, H.W. (1968). *The perception of stimulus relations: Discrimination learning and transposition*. New York: Academic Press.

Reeve, C. L., & Lam, H. (2007). The relation between practice effects, test-taker characteristics and degree of g-saturation. *International Journal of Testing, 7,* 225–242.

Rindermann, H. (2007). The relevance of class ability for teaching and development of individual competences. *Unterrichtswissenschaft, 35,* 68–89.

Roche, B., Barnes-Holmes, D., Smeets, P. M., Barnes-Holmes, Y., & McGeady, S. (2000). Contextual control over the derived transformation of discriminative and sexual arousal functions. *The Psychological Record, 50,* 267-291.

Royce, J. (1970). *Toward unification in psychology*. [Toronto]: University of Toronto Press in co-operation with the Center for Advanced Study in Theoretical Psychology of the University of Alberta.

Ryan, P. J. (1997). Unnatural selection: Intelligence testing, eugenics, and American political cultures . *Journal of Social History, 30,* 669–685.

Ryan, J., Sattler, J., & Lopez, S. (2000). Age Effects on Wechsler Adult Intelligence Scale-III Subtests. *Archives Of Clinical Neuropsychology*, *15*(4), 311-317. http://dx.doi.org/10.1093/arclin/15.4.311

Rychlak, J. (1988). Decomposing Psychology's Compositional Problem. *Psyccritiques*, *33*(11). http://dx.doi.org/10.1037/026197

Sallows, G., & Graupner, T. (2005). Intensive Behavioral Treatment for Children With Autism: Four-Year Outcome and Predictors. *American Journal On Mental Retardation*, *110*(6), 417. http://dx.doi.org/10.1352/0895-8017(2005)110[417:ibtfcw]2.0.co;2

Salthouse, T. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403-428. http://dx.doi.org/10.1037//0033-295x.103.3.403

Salthouse, T.A. (2010) Major issues in cognitive aging. New York: Oxford University Press.

Sattler, J. (1998). *Clinical and forensic interviewing of children and families*. San Diego, CA: Jerome M. Sattler, Publisher.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications.* San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Schaie, K. (1958). Differences in some personal characteristics of “rigid” and “flexible” individuals. *Journal Of Clinical Psychology*, *14*(1), 11-14. http://dx.doi.org/10.1002/1097-4679(195801)14:1<11::aid-jclp2270140104>3.0.co;2-f

Schaie, K., Labouvie, G., & Buech, B. (1973). Generational and cohort-specific differences in adult cognitive functioning: A fourteen-year study of independent samples. *Developmental Psychology*, *9*(2), 151-166. http://dx.doi.org/10.1037/h0035093

Schaie, K. W. (1983). Age changes in adult intelligence. In D. S. Woodruff & J. E. Birren (Eds.), *Aging: Scientific perspectives and social issues* (2nd ed., pp. 137-148). Monterey, CA: Brooks/Cole.

Schaie, K. W. (1994). Adult intellectual development. In R. Corsini (Ed.), *Encyclopedia of psychology* (pp. 33-34). New York: Wiley.

Schlinger, H. (1993). Learned expectancies are not adequate scientific explanations. *American Psychologist*, *48*(11), 1155-1156. http://dx.doi.org/10.1037/0003-066x.48.11.1155

Schlinger, H.D. (1998). Of planets and cognitions: The use of deductive inference in the natural sciences and psychology. *The Skeptical Inquirer, 22*, 49–51.

Schlinger, H.D. (2003). The Myth of Intelligence. *The Psychological Record, 53*(1), 15–33.

Schoenberg, M.R., Duff, K., Scott, J.G., Adams, R.L. (2002). An evaluation of the clinical utility of the OPIE-3 as an estimate of premorbid WAIS-III FSIQ. *Clinical Neuropsychologist, 17,* 308-321.

Schoenberg, M. R., Duff, K., Scott, J. G., & Adams, R. L. (2003). An evaluation of the clinical utility of the OPIE-3 as an estimate of premorbid WAIS-III FSIQ. *The Clinical Neuropsychologist*, *17*, 308– 321.

Schoenberg, M. R., Duff, K., Scott, J. G., Patton, D., & Adams, R. (2006). Prediction errors of the Oklahoma Premorbid Intelligence Estimate-3 (OPIE-3) stratified by 13 age groups. *Archives of Clinical Neuropsychology, 21*, 469-475.

Schreibman, L., & Koegel, R. L. (2005). Training for parents of children with autism: Pivotal responses, generalization, and individualization of intervention. In E.D. Hibbs & P.S. Jensen (Eds.), *Psychological treatments for child and adolescent disorders: Empirically based strategies for clinical practice (2nd edition)*, (pp. 603-631). Washington, DC: American Psychological Association.

Schwartzman, A., Gold, D., Andres, D., Arbuckle, T., & Chaikelson, J. (1987). Stability of intelligence: A 40-year follow-up. *Canadian Journal Of Psychology/Revue Canadienne De Psychologie*, *41*(2), 244-256. http://dx.doi.org/10.1037/h0084155

Seguin, E. (1866). *Idiocy and its treatment by the physiological method.* New York: William Wood.

Shea, V. (2004). A perspective on the research literature related to early intensive behavioral intervention (Lovaas) for young children with autism. *Autism*, *8*(4), 349-367. http://dx.doi.org/10.1177/1362361304047223

Sheinkopf, S. J., & Siegel, B. (1998). Home based behavioral treatment of young children with autism. *Journal of Autism and Developmental Disorders*, *28*(1), 15– 23.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology.* New York: Basic Books. Boston: Authors Cooperative.

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

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

Silva, M. A. (2008). Development of the WAIS-III: A brief overview, history and description. *Graduate Journal of Counseling Psychology, 1*(1), 1-19.

Sim, H. O. (1994). Coping Resources Mediating the Relationship between Stressful Life Events and Delinquent Behavior among Adolescents (Unpublished doctoral dissertation), Oregon State University.

Skinner, B.F. (1953). *Science and human behavior*. New York: Macmillan.

Skinner, B.F. (1957). *Verbal behavior*. New York: Appleton–Century–Crofts.

Skinner, B.F. (1966). An operant analysis of problem solving. In B Kleinmuntz (Ed.) *Problem solving: Research, method and theory* (pp. 133-171). New York: John Wiley & Sons.

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

Siegler, R. (1992). The other Alfred Binet. *Developmental Psychology, 28*, 179-190.

Smeets, P., Barnes, D., & Roche, B. (1997). Functional Equivalence in Children: Derived Stimulus–Response and Stimulus–Stimulus Relations. *Journal Of Experimental Child Psychology*, *66*(1), 1-17. http://dx.doi.org/10.1006/jecp.1997.2378

Smith, B., Smith, T., Taylor, L., & Hobby, M. (2005). Relationship between intelligence and vocabulary 1,2. *Perceptual And Motor Skills*, *100*(1), 101-108. http://dx.doi.org/10.2466/pms.100.1.101-108

Smith, T., Groen, A., & Wynn, J. (2000). Randomized Trial of Intensive Early Intervention for Children With Pervasive Developmental Disorder. *American Journal On Mental Retardation*, *105*(4), 269. http://dx.doi.org/10.1352/0895-8017(2000)105<0269:rtoiei>2.0.co;2

Sontag, L., Baker, C., & Nelson, V. (1955). Personality as a determinant of performance. *American Journal of Orthopsychiatry, 25*, 555–562. doi: 10.1111/j.1939-0025.1955.tb00151.x

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal Of Psychology*, *15*(2), 201. http://dx.doi.org/10.2307/1412107

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

Spearman, C. (1927). *The abilities of man*. New York: Macmillan Co.

Spearman, C., & Jones, L. (1950). Human Ability: A Continuation of the Abilities of Man. *Journal Of The Royal Statistical Society. Series A (General)*, *113*(4), 584. http://dx.doi.org/10.2307/2980883

Spencer, H. (1855). *The principles of psychology.* London: Longman, Brown, Green, and Longmans.

Spinath, B., Freudenthaler, H.H., & Neubauer, A.C. (2010). Domain-specific school achievement in boys and girls as predicted by intelligence, personality and motivation. *Personality and Individual Differences,* 48, 481–486.

Spinks, R., McKirgan, L. W., Arndt, S., Caspers, K., Yucuis, R., & Pfalzgraf, C. J. (2009). IQ estimate smackdown: Comparing IQ proxy measures to the WAIS-III. *Journal of the International Neuropsychological Society*, *15*(4), 590–596. http://doi.org/10.1017/S1355617709090766

Staats, S. (1989). Hope - A comparison of two self-report measures for adults. *Journal of Personality Assessment, 53*(2), 366-375.

Staats, A., & Burns, G. (1981). *Intelligence and child development*. Provincetown, Mass.: The Journal Press.

Staats, A., & Staats, C. (1963). *Complex human behavior*. New York: Holt, Rinehart and Winston.

Stankov, L. (1986). Kvashchev's experiment: Can we boost intelligence?. *Intelligence*, *10*(3), 209-230. http://dx.doi.org/10.1016/0160-2896(86)90016-4

Stanovich, K. E., & Cunningham, A. E. (2004). Inferences from correlational data: Exploring associations with reading experience. In N. Duke & M. Mallette (Eds.), *Literacy research methodologies* (pp. 28-45). New York: Guilford Press.

Stavrou, E. & Flanagan, R. (1996, March). *The stability of WISC—III scores in learning disabled children.* Paper presented at the Annual Convention of the National Association of School Psychologists, Atlanta, GA.

State Commission in Lunacy, State Hospitals, and Care, Custody, Apprehension, Commitment of Insane and Other Incompetent Persons (1913) *"Lunacy Law of the State of California."* Sacramento, CA: Friend Wm. Richardson.

Stephenson, C., & Halpern, D. (2013). Improved matrix reasoning is limited to training on tasks with a visuospatial component. *Intelligence*, *41*(5), 341-357. http://dx.doi.org/10.1016/j.intell.2013.05.006

Sternberg, R. (1977). *Intelligence, information processing, and analogical reasoning*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Sternberg, R. (2000). COGNITION: The Holey Grail of General Intelligence. *Science*, *289*(5478), 399-401. http://dx.doi.org/10.1126/science.289.5478.399

Sternberg, R. (2008). Increasing fluid intelligence is possible after all. *Proceedings Of The National Academy Of Sciences*, *105*(19), 6791-6792. http://dx.doi.org/10.1073/pnas.0803396105

Sternberg, R., & Detterman, D. (1987). What Is Intelligence? Contemporary Viewpoints on Its Nature and Definition. *The American Journal Of Psychology*, *100*(1), 141. http://dx.doi.org/10.2307/1422652

Sternberg, R., & Grigorenko, E. (2001). Unified psychology. *American Psychologist*, *56*(12), 1069-1079. http://dx.doi.org/10.1037/0003-066x.56.12.1069

Sternberg, R., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal Of Experimental Child Psychology*, *27*(2), 195-232. http://dx.doi.org/10.1016/0022-0965(79)90044-4

Stevenson, H., & Chen, C. (1989). Schooling and achievement: A study of Peruvian children. *International Journal Of Educational Research*, *13*(8), 883-894. http://dx.doi.org/10.1016/0883-0355(89)90071-2

Stewart, I., & McElwee, J. (2009) Relational Responding and Conditional Discrimination Procedures: An Apparent Inconsistency and Clarification. *Behavior Analyst, 32,* 309-317.

Stewart, I., & Roche, B. (2013). Relational Frame Theory: An overview. In S. Dymond, & B. Roche (Eds.) *Advances in Relational Frame Theory & Contextual Behavioral Science: Research & Application* (pp. 124-147). Oakland, CA: New Harbinger Press.

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 & Application* (pp. 178-198). Oakland, CA: New Harbinger.

Sugarman, S. (1982). Developmental change in early representational intelligence: Evidence from spatial classification strategies and related verbal expressions. *Cognitive Psychology, 14,* 410-449.

Terenzini, P., & Pascarella, E. (1991). Twenty years of research on college students: Lessons for future research. *Research In Higher Education*, *32*(1), 83-92. http://dx.doi.org/10.1007/bf00992835

Terman, L. M. (1916). *The measurement of intelligence.* Boston: Houghton Mifflin.

Thorndike, E. L. (1949) *Selected Writings from a Connectionist's Psychology*. Greenwood Press: New York.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale: Fourth Edition*. Itasca, IL: Riverside Publishing.

Trochim, W. (2000). *The Research Methods Knowledge Base*. Cincinatti, OH: Atomic Dog Publishing.

Tulsky, D., Zhu, J., & Ledbetter, M.F. (1997) WAIS-III/WMS-III technical manual. San Antonio, TX: The Psychological Corporation.

Turner, M.A. (1999). Generating novel ideas: Fluency performance in high-functioning and learning- disabled individuals with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 189-201.

U.S. Department of Health and Human Services. (1999). *"Mental Health: A Report of the Surgeon General -- Executive Summary".* Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

van der Maas, H., Kan, K., & Borsboom, D. (2014). Intelligence Is What the Intelligence Test Measures. Seriously. *J. Intell.*, *2*(1), 12-15. http://dx.doi.org/10.3390/jintelligence2010012

Vetterli, C., & Furedy, J. (1997). Correlates of intelligence in computer measured aspects of prose vocabulary: word length, diversity, and rarity. *Personality And Individual Differences*, *22*(6), 933-935. http://dx.doi.org/10.1016/s0191-8869(96)00265-6

Vernon P. A. (1983a) Speed of information processing and general intelligence. *Intelligence, 7*, 53-70.

Vial, M., & Hugon, M. A. (1998). *La commission Bourgeois (1904–1905).* Paris:

CTNERHI.

Vogel, R., & Annau, Z. (1973). An operant discrimination task allowing variability of reinforced response patterning1. *J Exp Anal Behav*, *20*(1), 1-6. http://dx.doi.org/10.1901/jeab.1973.20-1

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Waller, N., Tellegen, A., McDonald, R., & Lykken, D. (1996). Exploring Nonlinear Models in Personality Assessment: Development and Preliminary Validation of a Negative Emotionality Scale. *Journal of Personality*, *64*(3), 545-576. http://dx.doi.org/10.1111/j.1467-6494.1996.tb00521.x

Wang, J., & Kaufman, A. (1993). Changes in Fluid and Crystallized Intelligence Across the 20- to 90-Year Age Range on the K-Bit. *Journal Of Psychoeducational Assessment*, *11*(1), 29-37. http://dx.doi.org/10.1177/073428299301100104

Ward, L.C. (1990). Prediction of Verbal, Performance, and Full Scale IQs from seven subtests of the WAIS-R. *Journal of Clinical Psychology, 46*, 436–440.

Watkins, M., & Smith, L. (2013). Long-term stability of the Wechsler Intelligence Scale for Childrenâ€"Fourth Edition. *Psychological Assessment*, *25*(2), 477-483. http://dx.doi.org/10.1037/a0031653

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York, NY: Psychological Corporation.

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale manual*. New York, NY: Psychological Corporation.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins.

Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children–Revised.* New York: Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–Revised manual.* New York, NY: Psychological Corporation.

Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence–Revised, manual.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children, Third Edition, manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *WAIS-III WMS-III technical manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence, Third Edition, technical manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale, Fourth Edition technical manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition, manual*. San Antonio, TX: Psychological Corporation.

Weinert, F. E., & Schneider, W. (1999). *Individual development from 3 to 12: Findings from the Munich Longitudinal Study.* Cambridge: Cambridge University Press.

Weinert, F. E., & Hany, E. A. (2003). The stability of individual differences in intellectual development: Empirical evidence, theoretical problems, and new research questions. In R. J. Sternberg & J. Lautrey (Eds.), *Models of* *intelligence: International perspectives* (pp. 169- 181). Washington, DC: American Psychological Association.

Wilson, K., & Hayes, S. (1996). Resurgence of derived stimulus relations. *J Exp Anal Behav*, *66*(3), 267-281. http://dx.doi.org/10.1901/jeab.1996.66-267

Wolfle, D., & Oxtoby, T. (1952). Distributions of Ability of Students Specializing in Different Fields. *Science*, *116*(3013), 311-314. http://dx.doi.org/10.1126/science.116.3013.311

Wright, A., & Delius, J. (1994). Scratch and match: Pigeons learn matching and oddity with gravel stimuli. *Journal Of Experimental Psychology: Animal Behavior Processes*, *20*(1), 108-112. http://dx.doi.org/10.1037/0097-7403.20.1.108

Wulfert, E., Greenway, D., Farkas, P., Hayes, S., & Dougher, M. (1994). Correlation between self-reported rigidity and rule-governed insensitivity to operant contingencies. *J Appl Behav Anal*, *27*(4), 659-671. http://dx.doi.org/10.1901/jaba.1994.27-659

Yamamoto, J., & Mochizuki, A. (1988). Acquisition and functional analysis of manding with autistic students. *J Appl Behav Anal*, *21*(1), 57-64. http://dx.doi.org/10.1901/jaba.1988.21-57

Younger, B. (1993). Understanding category members as “the same sort of thing”: Explicit categorization in ten month infants. *Child Development, 64*, 309–320.

Zheng, Y., Myerson, J., & Hale, S. (2000). Age and individual differences in visuospatial processing speed: Testing the magnification hypothesis.

Psychonomic Bulletin & Review, *7*(1), 113-120.
http://dx.doi.org/10.3758/bf03210729

Zimmerman, L.L. & Woo-Sam, J.M. (1973). *Clinical interpretations of the Wechsler Adult Intelligence Scale.* New York: Grune and Stratton.

Appendix A

Information Sheet for Potential Research Participants

The current research project is designed to test the effectiveness of a newly developed technique for raising intellectual ability. The particular method used in this study is based on a psychological theory known as Relational Frame Theory and is called SMART training (Strengthening Mental Abilities with Relational Training).

If at any point you have attended a school of special education outside of the mainstream school system due to learning difficulties, or if you suffer with any intellectual problems that you know or feel constitute an intellectual disability then you may not be of use to us in this study and you should not volunteer to participate.

SMART training is based on the finding that most of what psychologists consider intelligent behaviour involves relating things to each other in a variety of ways. That is, intelligent people have a highly developed understanding of concepts such as before, after, more, less, opposite, different, same, here, there, and so on. When we teach these skills, intelligence, as measured by IQ tests, appears to rise. However, more research is needed to confirm that this is the case.

The tool we are using to train “relational skills” in this study has already been developed in previous research at Maynooth University. The main findings of that research have been published, and a web-based tool developed as a commercial enterprise within Maynooth University has also now been made publicly available. It is the utility of this online training tool, available at RaiseYourIQ.com, that we wish to test. As such, there is no guarantee that volunteers will benefit intellectually from the training.

This research is being conducted as part of a Masters level research programme by Mr. Dylan Colbert under the supervision of Dr. Bryan Roche of Maynooth University. SMART training involves an extended period of training on a personal computer or other internet connected device. Sessions will be at the users’ convenience, but all users will be asked to train for around 30 minutes, three to four times per week. The training usually requires approximately three months to complete. The training involves a quick intellectual assessment at the outset, which is then repeated at the end of the training. All of this is automated and will be delivered entirely online.

Training consists of solving a very large number of logical puzzles, followed by feedback from the computer in some cases but not in others. For example, users may be asked “If A is more than B and C is less than B, is A more than C?”. Users indicate their answer in all cases by clicking on the words “Yes” or “No” on the computer screen. Puzzles get increasingly more difficult but progression is guided by the computer, which delivers feedback to help the user improve their accuracy at these types of tasks.

Not all volunteers will be given this type of online training as part of the study. Half of the volunteers will be assigned to a “control group”, who will not receive any training during the study. However, all control group volunteers will get free access to the brain training at the end of the study.

The training will all take place at the user's home in privacy and at their convenience, although the researcher will be able to remotely see the frequency of logins by the user as well as their progress and may contact them from time to time just to ask how they are getting on and to check if the volunteers are happy to continue with their participation.

As part of this research it is also necessary to administer a full intelligence (IQ) test twice; once at the beginning and one at the end of the research programme (approximately three months apart). The test used will be the WAIS III IQ test. The researcher is currently being supervised by a qualified psychologist (Dr. Sarah Cassidy, Maynooth University) who will also take responsibility for seeing that these tests are administered appropriately and that data is processed and stored correctly and in accordance with normal data protection procedures. All IQ scores will be associated with user names using an encryption technique so that confidentiality is assured.

While the IQ scores from both assessments will be provided to the user at the end of their participation, a detailed psychological report will not be provided because IQ is being measured for research purposes only. If users have any concerns about their scores, however, they will be referred to the research team's educational Psychologist Dr. Sarah Cassidy free of charge. She will address any concerns participants have about their IQ score, and will help them to understand the meaning of their IQ test performance.

Finally, as part of this study a short self-esteem scale will be administered both at the outset and at the end of the experiment. The responses on this short questionnaire will also be treated confidentially and stored under code names only.

Volunteers can withdraw from the study at any stage even after giving consent, and may also withdraw their data at the conclusion of the study if they still have concerns.

Before you volunteer to participate, you should be made aware that in some circumstances, confidentiality of research data and records may be overridden by courts in the event of litigation or in the course of investigation by lawful authority. In such circumstances the University will take all reasonable steps within law to ensure that confidentiality is maintained to the greatest possible extent.

Mr. Dylan Colbert can be reached by email at DYLAN.COLBERT.2011@nuim.ie. Dr. Bryan Roche can be reached by email at the Department of Psychology, NUI Maynooth at Bryan.T.Roche@nuim.ie or by telephone at (01) 7086026. Dr. Sarah Cassidy can be reached at soconnor00@hotmail.com

Appendix B

Consent Form for Control Participants

• In agreeing to participate in the research project I understand the following:

• Mr. Dylan Colbert and Dr. Bryan Roche, of the Department of Psychology, Maynooth University are conducting this research. Mr. Colbert is the principal investigator and is a psychology graduate, currently gathering data for his postgraduate studies at Maynooth University.

• The purpose of this psychological research is to examine the effectiveness of a form of online intellectual skills training for increasing intellectual ability or "IQ".

• I understand that I will be asked to complete a full scale IQ test before and after a three month interval.

• I understand that I will not be given access to the online training software until after the completion of the study and my second intelligence assessment.

• All persons participating in this study will remain anonymous from each other and will not be referred to by name in any publication or document. The data will remain confidential at all times and will be referred to by code names only. The data collected will be used only by the researchers. This data will be available to each participant should they request it.

• The researchers will conduct all parts of this study in line with the ethical code of conduct laid down by the Psychological Society of Ireland and the Ethics Committee of Maynooth University.

• I understand that I may withdraw from the study at any stage even after giving my consent. I may also withdraw my data at the conclusion of my participation if I still have concerns.

• I understand that I may contact Dr. Bryan Roche at Bryan.t.roche@nuim.ie of Maynooth University as supervisor of this research if I have any concerns and can access a private clinical consultation with Dr. Sarah Cassidy (private practice and Maynooth University) if I have concerns about my IQ score.

• I understand that the training I will receive at the end of this study, is experimental and not clinical in nature and that I will not receive a full psychological report along with my IQ scores at the end of my participation.

• I have been informed as to the general nature of the study. I have read the research information sheet. I understand that at the conclusion of my participation, any further questions or concerns I have will be fully addressed.

• I acknowledge that in some circumstances, confidentiality of research data and records may be overridden by courts in the event of litigation or in the course of investigation by lawful authority. In such circumstances the University will take all reasonable steps within law to ensure that confidentiality is maintained to the greatest possible extent.

• I am over 18 years of age.

Signed in duplicate:

______________________________ Participant

______________________________ Researcher

______________________________ Date

Appendix C

Consent Form for Experimental Participants

• In agreeing to participate in the research project I understand the following:
• Mr. Dylan Colbert and Dr. Bryan Roche, of the Department of Psychology, Maynooth University are conducting this research. Mr. Colbert is the principal investigator and is a psychology graduate, currently gathering data for his postgraduate studies at Maynooth University.
• The purpose of this psychological research is to examine the effectiveness of a form of online intellectual skills training for increasing intellectual ability or "IQ".
• I understand that I will be asked to complete a full scale IQ test before and after my online training. These will each take around 90 minutes to complete. I will also complete a short assessment of my "relational skills" before and after training which will take the form of a test for logical reasoning, not unlike an algebra test.
• My training will take place online at my own convenience but I am requested to train for around 30 minutes 3-4 times per week until I have completed all stages of the training. This will take around three months.
• All persons participating in this study will remain anonymous from each other and will not be referred to by name in any publication or document. The data will remain confidential at all times and will be referred to by code names only. The data collected will be used only by the researchers. This data will be available to each participant should they request it.
• The researchers will conduct all parts of this study in line with the ethical code of conduct laid down by the Psychological Society of Ireland and the Ethics Committee of Maynooth University.
• I understand that I may withdraw from the study at any stage even after giving my consent. I may also withdraw my data at the conclusion of my participation if I still have concerns.
• I understand that I may contact Dr. Bryan Roche at Bryan.t.roche@nuim.ie of Maynooth University as supervisor of this research if I have any concerns and can access a private clincal consultation with Dr. Sarah Cassidy (private practice and Maynooth University) if I have concerns about my IQ score.
• I understand that this experiment does not constitute any kind of validated intervention and that intellectual gains are not guaranteed.
• I understand that the training I will undergo is experimental and not clinical in nature and that I will not receive a full psychological report along with my IQ scores at the end of my participation.
• I have been informed as to the general nature of the study. I have read the research information sheet. I understand that at the conclusion of my participation, any further questions or concerns I have will be fully addressed.
• I acknowledge that in some circumstances, confidentiality of research data and records may be overridden by courts in the event of litigation or in the course of investigation by lawful authority. In such circumstances the University will take all reasonable steps within law to ensure that confidentiality is maintained to the greatest possible extent.
• I am over 18 years of age.

Signed in duplicate:
______________________________ Participant
______________________________ Researcher
______________________________ Date