Relational Frame Theory and Human Intelligence: A Conceptual and Empirical Analysis

Sarah Cassidy

Thesis submitted for Ph.D to:
The Department of Psychology,
Faculty of Science,
The National University of Ireland, Maynooth.
October 2008

Supervised by Dr. Bryan Roche

Acknowledgements

My sincere gratitude to:

Kevin—for knowing everything there is to know about everything, for constantly surprising me with just how amazing you are, for ten million breakfasts in bed, for generosity and patience, for cultural education and culchie enlightenment. I don't know what I'd do without you.

Patrick—for being so effortlessly perfect. You are, quite literally, the best child in the whole world. I never stop being amazed by you, or being thankful for you.

Mom, Dad, Donna, Sean, Vimla, Brenda, Karen, Matthew, Alfie and Malachy—for a lifetime of having faith in me that often exceeded the bounds of logic and reason.

Meg, Anne and Fodhla—for being the best friends a person could ask for.

The staff in the Psychology Department at NUIM—for hours of technical and administrative support.

The generous participants—for giving so willingly of your time and your family's time so that this research could be completed. A special thanks to Séamus Ó Muirithe, Proinsias Logan and the staff at Scoil Ui Riada, Cill Choca, for forgiving so many disruptions to your daily schedules, for providing the opportunity and the facility for much of this work to be conducted. None of this would have been possible without your generosity and flexibility.

Bryan Roche—for your enthusiasm and energy, for dragging me up this very steep learning curve. This has been an exercise in eduction at its finest.

Table of Contents

Abstract

The current thesis builds upon developments in the field of Relational Frame Theory (RFT), which has proposed a behavioural re-examination of the widely used concept of intelligence in terms of derived relational responding (DRR). In the first chapter the concept of intelligence is explored theoretically from a RFT perspective. A framework for the construction of interventions to raise intelligence quotients as calculated by standardised IQ tests is also provided. Specifically, the current thesis proposes that training skills in DRR by utilising multiple exemplar training (MET) can improve intellectual skills. In Experiments 1 and 2, it is shown that, by employing a MET intervention for symmetry and transitivity, modest rises in full scale IQ on the WISC-III^{UK} were generated for normally functioning children.

In Experiment 3, the MET intervention methodology is further developed across a group of both children and adults to specifically improve the relational skills which appear to underlie intelligence as a behavioural repertoire. The newer methodology is shown to generate repertoires of Same and Opposite relational responding for experimental groups of adults and children, where these repertoires were previously weak or absent.

In Experiments 4 and 5, the stimulus control of the intervention is further improved. Experiments 6 and 7 involve the addition of intervention protocols for More-Than/Less-Than relational responding. MET is again shown to facilitate the emergence of DRR for Same and Opposite (Experiments 4 and 5), and also for More-Than/Less-Than (Experiments 6 and 7) with both child and adult groups. However, Experiments 6 and 7 failed to clearly establish the functional dependence of More-Than/Less-Than

responding on Opposite relational responding alone. In Experiment 8, participants with an extended history of MET across symmetry, transitivity, Same and Opposite showed rapid acquisition of More-Than/Less-Than DRR.

Experiment 9 measured considerable rises in WISC-III^{UK} scores across an extended MET intervention for four children. Importantly, similar rises were not seen for a matched control group who had no access to the intervention. In Experiment 10a, a relational abilities index (RAI) is developed for use as a baseline relational skills index. This baseline measure is then correlated with the WISC-IV^{UK} and its subtests for a group of children with learning difficulties (Experiment 10b). Several interesting correlations between relational skills and intelligence are observed in the resulting analysis, although many theoretical and conceptual issues are also suggested by the data.

In Experiment 11, a complete MET battery is administered to an educationally challenged child group. Both RAI and full scale IQ scores rise from pre to post intervention. In the closing chapter, the implications of these rises for intellectual disability, educational psychology and behaviour analysis are discussed.

Chapter 1

Relational Frame Theory and Human Intelligence

General Introduction

Behaviour analysts have traditionally rejected hypothetical constructs (rather than private events) such as intelligence (see Skinner, 1974) in their explanations of behaviour. In contrast, most educators adopt the mainstream view that there exists a common underlying factor that ties all intellectual skills together. Spearman (1904) called this factor ‘g’, for general intelligence. However, from a behavioural perspective, there are epistemological problems with the reification of the concept of intelligence as an extant entity as well as its use in circular reasoning to explain intelligent behaviour, which it originally merely described (see Schlinger, 2003).

While interpretive behavioural analyses of intelligence measures have been provided (e.g., Schlinger, 2003), they have fallen short of suggesting immediate research programmes designed to develop interventions to increase the fluency of behaviours assessed by popular IQ tests. Of course, given the mentalistic status of intelligence as a concept, there is no reason why behaviour analysts should be concerned with interventions to raise IQ scores per se. However, we often find ourselves intervening in applied settings to increase the fluency of those very behaviours widely assessed in IQ tests (e.g., verbal and computational ability). In effect, the ability to raise IQ scores would clearly demonstrate the behaviour analyst’s ability to analyse and influence those behaviours widely referred to as “intelligent” in purely functional terms. Indeed, were such an endeavour to be successful, we could hasten the abandonment of the concept of intelligence as anything other than a mentalistic summary term for a repertoire of skills, now fully understood in behavioural terms.

Unfortunately, the field of behaviour analysis has largely ignored the determinants of consistent differences in the level of performance among individuals, despite its goal of understanding individual behaviour (see Williams, Myerson, & Hale, 2008). Williams et al. (2008) have called for researchers in the field of behaviour analysis to determine which learning tasks predict individual differences in intelligence and which do not, and then to identify the specific characteristics of these tasks that make such prediction possible. Of course, this would be a difficult endeavour, especially given that intelligence has so many different meanings. In fact, it was for this reason that one of the most prominent researchers in the field of intelligence called for the term to be abandoned (see Jensen, 1998; see also Williams et al., 2008). While, that may be a practical suggestion, the current chapter will nevertheless briefly summarise some of the major theories of intelligence. It may be clear from this summary that they all produce the same epistemological problems for behaviour analysts, due to their common mentalistic features.

For the purpose of the brief review we will rely on Lohman's (1989) classification of intelligence theories into three main types; trait theories, information processing theories and theories of thinking.

Trait Theories: Trait theories generally focus on "fluid" and "crystallized abilities" (see Horn, 1985; Snow, 1981; see also Lohman, 1989). In 1943, Cattell argued that fluid ability was a purely general ability to discriminate and perceive relations between any fundaments, new or old (p. 178). Furthermore, fluid intelligence was thought to be the source of the general factor found among ability tests administered to children and among the "speeded or adaptation-requiring" (p. 178) tests administered to adults. On the other hand, crystallized intelligence was thought to consist of "discriminatory habits long established in a particular field" (p. 178) that were originally acquired through the operation of fluid ability that no longer required "insightful perception" (Cattell, 1943). The important psychological distinction being made in these theories is between process (fluid) and product (crystallized) intelligence. For instance, in Cattell's 1963 formulation, fluid intelligence was hypothesised to "reflect the physiological integrity of the organism useful for adapting to novel situations. When invested in particular learning experiences, this fluid intelligence was thought to produce crystallized intelligence" (Lohman, 1989, p.339). This crystallized intelligence, in turn, was thought to be a product of environmentally varying, experientially determined investments of fluid intelligence (Cattell, 1963, p. 4).

It would seem that both the concepts of fluid and crystallized intelligence involve a significant degree of environmental influence, and as such, are somewhat parallel to a behavioural view of intellectual skills. On the other hand, the concepts of fluid and crystallized intelligence remain mentalistic concepts because the skills they describe are located inside the organism and are thought to have psychometrically measurable qualities. In effect, the psychologist who uses these concepts is more interested in invariant mental traits than in behaviour situated in terms of its antecedents and consequences.

Information Processing Theories: Those who favour the information processing approach have attempted to build process theories of the major ability factors identified in Horn's (1985) ability model. These are; mental speed, verbal-crystallized ability, fluid-reasoning abilities, and spatial-visualization abilities. Examples of the information processing approach can be found in the work of Jensen (1982) and Eysenck (1982) on mental speed; in Hunt (1985) and Fredericksen (1982) on verbal-crystallized abilities; in Sternberg (1977) on fluid-reasoning abilities; in Pellegrino and Kail (1982); and in Lohman (1988).

The information processing approach is clearly cognitive in terminology. This approach may be considered less environmentally oriented, and involving a greater appeal to mentalism than the trait theories of intelligence. It will be obvious to many readers that behaviour analysts find these mentalistic explanations of intellectual ability in terms of hypothetical mental traits to be oversimplified and epistemologically confused (e.g., see Skinner, 1974). According to Skinner, mentalistic modes of explanation persist within psychology because “they seem so much simpler than the facts they are said to explain” (Skinner, 1974, p. 254; see also Lattal & Chase, 2003; see also Zuriff, 2005, for a review of Lattal & Chase, 2003).

The behavioural objection to the appeal to cognitive processes in explaining behaviour is a well-established one. A behavioural view requires that no explanatory concept can be invoked which has not been analysed in the laboratory under experimental control, and the terms must interact according to empirical principles (Palmer, p. 178, in Lattal and Chase, 2003). For instance, Zuriff (2005) provides the following example of a behavioural translation of a cognitive process that might be inferred in explaining the appropriate response to a test item of the type commonly found on IQ tests. Specifically, Zuriff describes the response to the question; “What is the tenth letter after F” as “covert mediating verbal behaviour” (p. 318). This behavioural explanation posits the response to the question as a product of plausible history and employs familiar behavioural principles, such as reinforcement, generalization and chaining. In contrast, a cognitive explanation might appeal to constructs such as representations, memory storages, and control processes. These theoretical concepts are not constrained by the requirement to be of independent verification in laboratory studies and therefore cannot carry any explanatory weight to the behaviour analyst (see also Baum, 2004).

Theories of Thinking: Sternberg (1985) attempted to develop a more comprehensive theory of intelligence from general theories of thinking in cognitive psychology and artificial intelligence (see Lohman 1989; Snow & Lohman, 1989; Sternberg, 1985). Sternberg's "Triarchic Theory" of human intelligence contains three subtheories; a contextual subtheory, an experiential subtheory and a componential subtheory. The contextual subtheory attempts to specify those behaviours that would be considered intelligent in a particular culture. However, even if a particular task is thought to require intelligence, contextually appropriate behaviour is not equally "intelligent" at all points along the continuum of experience with the class of tasks. According to the experiential subtheory, intelligence is best demonstrated when the task or situation is relatively novel or when learners are practicing their responses to the task so that they can respond automatically and effortlessly. Sternberg goes on to claim that the ability to automatise processing is also a good indicator of intelligence.

In Sternberg's componential subtheory, he attempts to specify the cognitive structures and processes that underlie all intelligent behaviour. Three types of processes are hypothesised; metacomponents, which control processing and enable one to monitor and evaluate it; performance components, which execute plans assembled by the metacomponents; and knowledge acquisition components, which selectively encode and combine new information and selectively compare new information to old information. These three processes in Sternberg's componential subtheory are clearly hypothetical and support for them can only be obtained through inferential statistical analysis of test outcomes that are themselves calculated based on the usual psychometric assumptions concerning the stability of intelligence traits. Thus, while Sternberg has succeeded in theoretically unifying diverse, and even antagonistic, traditions in research on intelligence, it is unclear whether or not his efforts to develop new measures of practical intelligence have been of use (Lohman, 1989).

Another *thinking theory* of intelligence is provided by David Wechsler. Wechsler (1944, p. 3) viewed intelligence not as "a particular ability, but as an aggregate and global entity, the capacity of the individual to act purposefully, to think rationally and to deal effectively with his or her environment. However, some would argue that intelligence as measured in the tradition of Wechsler (and Binet), is better construed as scholastic aptitude rather than "intelligence" per se (Lohman, 1989). To this extent, the Wechsler test is not a quotient measure of some invariant trait at all, but rather an index of scholastic ability, subject to environmental influence.

It is not practical, or even necessary, for the behaviour analyst to integrate the various foregoing concepts of intelligence in order to arrive at some final behavioural formulation that speaks to each of these psychometric views. Rather, the behaviour analyst will start from the point of view of trying to theoretically describe, and then predict and influence, the types of skills that are examined in widely used intelligence tests. In effect, behaviour analysts will proceed by viewing intelligence as nothing more than a label for various behaviours occurring in specific contexts. More specifically, for behaviour analysts, a truly scientific understanding of intelligence can only come from a functional analysis of intelligent adaptive behaviours in the contexts in which they are observed (Schlinger, 2003). Of course behaviour analysts have no fundamental problem with the idea that intelligent behaviours may be functionally related, an observation which may support the description of wide repertoires of behaviour akin to intelligence factors. Nevertheless, the problem with the concept of general intelligence is that it remains elusive, hypothetical and not amenable to scientific manipulation or intervention. Indeed, this is perhaps just one of the reasons why the research endeavour in the current thesis may speak directly to suggestions by Williams et al. (2008) for researchers to determine which learning tasks predict individual differences in intelligence, and which do not, and then to identify the specific characteristics of those differences that make such prediction possible.

Given that the study of individual differences has largely been left out of experimental analyses of human behaviour, embarking on a programme of research to raise IQ is an ambitious endeavour, but the theoretical and technical impetus for it has already been provided by Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). This theory provides both a coherent account of human intelligence from a behavioural perspective, and also suggests empirically testable interventions to raise intellectual skills, as measured by IQ tests. The current thesis will outline the RFT account of human intelligence and describe the means by which interventions to raise IQ can be immediately generated. Firstly, however, it is necessary to provide an outline of the central features of this theoretical account of complex behaviour.

Relational Frame Theory

RFT provides a technical analysis of language and cognitive abilities and, more importantly, suggests procedures for making conceptual and empirical inroads to the problem of understanding the origins of and the functional nature of intelligence. In particular, RFT claims that the foundational skill for most intellectual abilities is *Derived Relational Responding* (DRR). Derived Relational Responding is the skill of relating objects to each other in accordance with a small family of mathematical relationships (e.g., *symmetry, transitivity, equivalence, opposition, difference, comparison*, etc).

For those unfamiliar with the concept of DRR, the formation of a stimulus equivalence relation under laboratory conditions will be outlined. Typically, across at least two conditional discrimination tasks, a participant is taught to discriminate between two comparison stimuli given a sample stimulus. The first conditional discrimination involves teaching the participant to choose between two comparison stimuli, labeled for convenience as B1 and B2, conditional upon the presentation of A1 or A2, respectively, as a sample. The second conditional discrimination task involves choosing between two further comparisons, C1 and C2, conditional upon the presentation of B1 or B2, respectively, as a sample. In effect, participants are taught to choose B1 given A1 and C1 given B1 (A1-B1-C1) and to choose B2 given A2 and C2 given B2 (A2-B2-C2). When provided with this (linear protocol) training, most verbally able participants will match each stimulus with itself in the absence of reinforcement. For instance, given A1 as a sample, and A1 and B1 as comparisons, verbally able participants will choose the A1 comparison. This behavioural outcome is referred to as *reflexivity*. Furthermore, participants will derive symmetrical relations between the stimuli without feedback or reinforcement. *Symmetry* involves the spontaneous transfer of stimulus control from the sample stimulus to the comparison stimulus in a matching-to-sample preparation. Thus, given the above training, a participant will be able to pick A1 from an array when given B1 as a sample, and B1 from an array when given C1 as a sample. Finally, participants will display *transitivity* in the absence of reinforcement. This refers to the spontaneous combining of trained relations and the emergence of stimulus control for comparison stimuli not directly associated with the original sample stimulus. For instance, if a participant is trained to pick B1 from an array given A1 as a sample, and C1 given B1, the stimulus C1 will now be chosen given A1 (i.e., the response functions of B1 have transferred to A1). When all three features have been observed, a stimulus equivalence relation is said to have formed among the relata (see Fields, Adams, & Verhave, 1993; Fields, Adams, Verhave, & Newman, 1990; Sidman, 1971, 1986).

RFT adopts a somewhat different nomenclature for discussing the foregoing effects. According to RFT, derived relations involve the properties of *mutual entailment*, *combinatorial entailment*, and the *transformation of function*.

Mutual entailment: If a stimulus A bears a relationship to another stimulus B, then a further derived relation between B and A is mutually entailed. The type of relation mutually entailed depends upon the nature of the relation between A and B (Hayes, 1994). For instance, if the stimulus A bears an *equivalence* or “coordination” relation to the stimulus B, then the relation "B is the *same as* A" is mutually entailed. However, if the stimulus A bears a relation of *comparison* to the stimulus B (e.g., A is *more than* B), then the relation "B is *less than* A" is mutually entailed.

Combinatorial entailment: If a stimulus A bears a relation to B, and B bears a further relation to C, then a relation between A and C is combinatorially entailed. The nature of the combinatorially entailed relation depends on the nature of the trained relations. For example, if “A is *more than* B” and “B is *more than* C”, then a "*more than*" relation between A and C is derived by combinatorial entailment (i.e., A is *more than* C) and a "*less than*" relation is entailed between C and A (i.e., C is *less than* A).

Transformation of function: If a stimulus A is related to another stimulus B, and A acquires a psychological function, then in the appropriate context the stimulus functions of B will be transformed in accordance with the A-B relation. For example, if “A is *more than* B”, and A elicits fear, then B will produce less fear than A.

RFT refers to equivalence relations as just one type of *relational frame*. A relational frame is conceptualised as a three-term contingency, wherein the contextual cue is the third term, the relational response (e.g., responding to stimulus B in terms of A and responding to A in terms of B) is the second term, and a history of differential reinforcement correlated with the contextual cue is the first term in the contingency. Thus, DRR is viewed as a single response unit controlled by a relevant contextual cue. Put simply, a relational frame is an "over-arching" or generalised relational operant with an infinite range of topographies (see Barnes, 1994; D. Barnes-Holmes & Y. Barnes-Holmes, 2000; Hayes, 1992; see also Galizio, 2003; Malott, 2003, McIlvane, 2003 for critiques and commentaries).

RFT suggests that the ability to derive relations is itself a learned operant skill. This sets the theory apart from Sidman's view of derived relational responding as a basic stimulus function (e.g., Sidman, 2000; see Hayes & Barnes, 1997). In simple terms, RFT suggests that the ability to derive relations is itself established by caregivers at an early stage, across multiple exemplars, often without the caregiver even being aware. At a later stage, familiar relations (e.g., equivalence) can be derived using novel stimuli, but the skill itself is far from novel. For example, suppose a mother tells her child that a certain piece of fruit is an 'orange' and reinforces appropriate echoing of that word in the presence of the object. This practice establishes the object-word relation. Now, suppose the mother also asks the child to "show me the orange" and reinforces the appropriate orienting response towards the object (e.g., by pointing). This establishes the word-object relation. Across thousands of such exemplars with different objects, the mother is wittingly or unwittingly teaching the child that all word-object and object-word relations are reversible (Hayes, Fox, Gifford, Wilson, Barnes-Holmes, & Healy, 2001, p. 26-27).

By providing a child with multiple exemplars of the foregoing kind, a caregiver is teaching a child skills that become more and more abstracted over time. Moreover, RFT suggests that these skills may be foundational to language itself.

Specifically, from a RFT perspective, all language for humans involves being able to respond to the arbitrary relations between words and events. Thus, language involves using and responding to words whose meaning has been derived through *equivalence* (Barnes, McCullagh, & Keenan, 1990; Berens & Hayes 2007, Devany, Hayes, & Nelson, 1986; Hayes, Fox et al., 2001).

RFT takes a particular interest in relations other than stimulus equivalence. These relations are perhaps most easily conceived of here as contextually controlled conditional discriminations. For instance, it is possible to train participants to respond to stimuli that are physically similar or different to each other in the presence of arbitrary cues, and thereby establish those cues as contextual stimuli that control "*same*" or "*opposite*" responding (see Steele & Hayes, 1991). These cues can then be employed during equivalence training to effectively establish relations of stimuli that are arbitrarily the "*same*" or "*opposite*" as or to each other (see Roche & Barnes, 1997; Steele & Hayes, 1991). Indeed, any cue that controls relational responding towards formally related stimuli (e.g., *smaller than, opposite to, before*, et cetera.) can be employed to create arbitrary relations of the same kind. RFT refers to these latter relations as *arbitrarily applicable.*

One important feature of RFT in the current context is its suggestion that a small variety of these arbitrary relations (e.g., *same, opposite, more-than, less-than*) may be sufficient to yield the full gamut of cognitive skills (e.g., deductive reasoning, inductive reasoning, analogy, et cetera.) associated with high intelligence. Such a perspective leads to a whole host of empirical questions regarding the generalization of relational skills. Researchers have wondered, for instance, whether *equivalence* responding requires explicit training in mutual and combinatorial entailment, or in mutual entailment alone (e.g., Boelens, 1994; Horne & Lowe, 1996). Exactly, how much and what kind of training is needed for the generalisation of relational skills is seen as an empirical matter. RFT suggests that at least some direct training in deriving relations is necessary to create a sufficient repertoire of relational responding to constitute intelligent behaviour (see Hayes & Wilson, 1996). Thus, the effective use of the RFT approach in applied settings will require research that will identify the nature and number of multiple exemplars that are needed to establish particular repertoires of relational responding. This research will need to functionally map the development of specific repertoires of relational skills in terms of their impact on specific aspects of cognitive abilities. In effect, such an endeavour would allow behaviour analysts to speak more directly than ever before to the concept of intelligence as interpreted and measured by widely employed psychometric tests. To illustrate this point the following section considers some specific dimensions of intelligence as traditionally conceived by intelligence tests and illustrates how RFT provides a conceptual framework for the functional analysis of these behavioural skills.

A Relational Frame Analysis of IQ

An important component of the modern conceptualisation of psychometric intelligence is the concept of verbal intelligence. According to RFT, most relational skills likely emerge in the context of language acquisition. That is, most children learn to speak and to use words symbolically before they learn to apply abstract relations in a mathematical context. Thus, individuals with well-developed verbal repertoires (e.g., vocabulary) might also be expected to a have a rich repertoire of relational skills. Consequently, RFT would expect to find that individuals with high scores in verbal intelligence on standard IQ tests would also show evidence of robust relational skills if assessed in a laboratory context. More specifically, an individual with a high verbal IQ can be described in relational terms as an individual who is able to "elaborate entire networks of stimulus relations quickly, to bring them under increasingly subtle forms of contextual control, to transform stimulus functions through entire networks, and to abstract features of the natural environment that will support and sustain relational responding" (Y. Barnes-Holmes, D. Barnes-Holmes, Roche, et al., 2001, p. 161).

Some examples of items in the Vocabulary subtest of the Wechsler Intelligence Scale for Children, third edition, UK (WISC-III^{UK}) include questions that appear to assess prima facia relational skills. Specifically, the WISC-III^{UK} contains questions like, "What is an umbrella?" and "What does brave mean?". While these items clearly test for object-word equivalence relations and word-word equivalence relations, respectively, their intention is to examine the extent of a person's vocabulary rather than the culturally specific arrangements of language categories. Questions such as, "What does dilatory mean?" or "What does imminent mean?" are further examples of word-word relations, while "What does aberration mean?", "What is an amendment?", and "What is an affliction?" are further examples of probes for word-object equivalence relations. From the RFT perspective, a vocabulary test, while relatively rudimentary as a test of foundational language skills, likely makes a satisfactory approximation of a test for relational skill because the two skills should correlate very highly for a verbally able individual.

The Picture Concepts subtest on the Wechsler Intelligence Scale for Children, fourth edition, UK (WISC-IV^{UK}), while it is not classified by the test manufacturers as a verbal test, is clearly a test for equivalence relations, or, in RFT terms, frames of *coordination*. Specifically, in this subtest, a child is presented with two or three rows of pictures, and must choose one picture from each row to form a group with a common characteristic (see Wechsler, 2004). For example, in one row a child will see a piece of cheese, a butterfly, flowers and weighing scales. In the next row, a child will see a map, a palette of paint, a lamp and a paintbrush. In the third row, the child can see a newspaper, an ice-cream cone, a different bunch of flowers and a postage stamp emblazoned with a flower. The child must choose the scales, the map and the newspaper as having common characteristics (i.e., because one can "read" all of these items to gain information). Of course, the formal features of these stimulus items are dissimilar, requiring that the commonalities be based on the participation of the relevant stimuli in common derived frames of coordination, rather than on their formal features. Thus, the verbal skills assessed in this task are over-arching skills applied across many domains, some of which may be traditionally referred to as verbal, others as spatial, and others as computational. It may well emerge from future research therefore, that all intelligence test items will come to be viewed by behaviour analysts as tests for various forms of DRR and the transfer of discriminative control among related stimuli. This is the view taken by RFT. Indeed, it is precisely because of the expected dependence of overall intelligence quotients on the ability to derive relations among stimuli that the dramatic correlations between early language learning environments and later educational ability (e.g., Hart & Risley, 1995) are of no surprise to the Relational Frame theorist.

Research has already begun to support the RFT prediction that correlations should be observed between verbal abilities on standardized IQ tests and relational skills measured in laboratory settings. In one study, O'Hora, Pelaez and Barnes-Holmes (2005) sought to determine if performance on relational tasks would predict performance on verbal or performance subtests of the Wechsler Adult Intelligence Scale III (WAIS III). In this research 75 subjects were broken into two groups based on their performance on a complex relational task involving pre-training and testing for *before/after* and *same/different* relations, a test for instructional control and a test for generalisation of instructional control using novel stimuli. These subjects were then exposed to three subtests of WAIS III (Vocabulary, Arithmetic, and Digit-symbol coding). Subjects who successfully completed the complex relational task (n=31) performed significantly better on the Vocabulary and Arithmetic subtests than those subjects (n=44) who failed to do so. No significant differences in relational task performances were found between groups for the Digit-symbol coding subtest. These findings support the position that derived relational performances may be related to human language abilities, as measured by subtests of the widely utilised WAIS III.

More recently, a research programme conducted by O'Toole and Barnes-Holmes (in press) involved presenting participants with before/after and similar/different relational tasks using the Implicit Relational Assessment Procedure (IRAP). Participants were then administered the Kaufman Brief Intelligence Test (K-BIT). Response latencies on consistent and inconsistent trials were measured. In addition, difference scores were calculated by subtracting consistent response latencies from inconsistent response latencies. The response latencies and difference scores were considered to provide measures of relational flexibility. Results of this research programme showed that faster responding on the IRAP and smaller difference scores predicted IQ. Findings suggested that response speed on trials that were inconsistent might provide a possibly useful measure of relational or cognitive flexibility. O'Toole et al. (in press) note that assessing relational flexibility may be particularly advantageous because flexibility is widely regarded to be an important component of human cognitive abilities (e.g., Cattell, 1971; Kyllonen, Lohman, &

Woltz, 1984; Premack, 2004). For these reasons, O'Toole et al. (in press) suggest relational flexibility as something to be targeted in educational settings.

The foregoing studies show promising empirical results that derived relations are important to an understanding of intelligence, particularly those aspects of intelligence that might be considered verbal. In fact, the idea that DRR and language skills are functionally related would now appear to be well supported in the literature. Specifically, Devany, Hayes and Nelson (1986), and Barnes, McCullagh and Keenan (1990), have shown that the ability to derive equivalence relations is absent in language-disabled children. Dugdale and Lowe (1990) and Hayes (1992) have argued that despite the capacity of most vertebrate species to acquire the basic trained relations, only verbally-able human subjects display the spontaneous emergence of novel relations satisfying criteria for equivalence, with a few possible exceptions (e.g., C. R. Kastak & Schusterman, 2002; C. R. Kastak, Schusterman, & D. Kastak, 2001; D. Kastak & Schusterman, 1994).

One study in particular is often cited as support for the idea that language ability and DRR are functionally related. This longitudinal study, conducted by Lipkens, S. C. Hayes and L. J. Hayes (1993) found that the ability to derive relations occurs at the same time as the language explosion. In the authors' words; "…by 23 months the child would mutually relate novel names and objects based on a relation of difference with a known object" (p. 41 in Hayes, Barnes-Holmes, & Roche, eds., 2001).

Neurophysiological evidence has also been published that points to common patterns of brain activation associated with both language and the performance of stimulus equivalence. Specifically, Dickens, Singh, Roberts, Burns, Downes, Jimmieson and Bentall (2000) reported that brain activation patterns as measured by fMRI during equivalence tasks closely resembled those involved in semantic processing associated with language. Moreover, the activation did not appear without involving regions concerned with sub-vocal articulation of stimulus names.

Mathematical skills as assessed in standard IQ tests also represent what appears to be an index of the ability to derive and apply abstract relations. For example, in the Arithmetic subtest of the WISC-III^{UK}, a child is presented with the following problem; “Joseph has 5 cakes. He gives 1 to Sam and 1 to Alice. How many does Joseph have left?” Another problem is as follows; “Phil earned £36; he was paid £4 an hour. How many hours did he work?” Questions like these are highly abstract and novel, but from a RFT perspective the skills involved in responding correctly to these test items may not be so novel. Specifically, answering an infinite range of such questions correctly requires a highly topographically flexible repertoire of relational skills. The infinite variety of possible questions of this kind precludes the possibility of learning each one individually (i.e., producing a relationally inflexible topographically constrained response to pre-set questions). For instance, a child who responds correctly to the questions above should also be able to respond correctly if Joseph gave an additional cake to Michael or if Phil earned £5 an hour. The reason for this is that with mathematical skills, a teacher does not only teach computation by rote, but also teaches the relative relations between numbers such that a child should be able to respond to the relations 5-2 and 8-5 as being the same (i.e., 3). For instance, if presented with the numbers; “1, 5, 9, 13, 17 …”, most verbally able adults would have little trouble correctly providing the next number in the series (i.e., 21). This is a relational problem and is solved by responding to the single relation that consistently obtains between subsequent items in the series and applying that relation arbitrarily to the last number in the series. In the above case, the relation between subsequent items in the series might be called "*plus 4*" (see Y. Barnes-Holmes, D. Barnes-Holmes, Roche, et al., 2001, p. 162). Thus, the skill that is being taught has little to do with the fact that "17+4=21", but everything to do with the ability to generalise the skill of "*adding 4*" to any given arbitrary number or sequence of numbers.

Complex mathematical problems often involve increasingly more subtle contextual control over DRR. For instance, in a problem involving calculating the distance traveled by a train between two points in a given time under a range of different conditions (e.g., varying speed) there may be multiple sources of contextual control that come together to produce the correct response. More specifically, the problem may not be correctly solved by bringing relational responding under the control of one specific contextual cue for responding in accordance with an addition or a multiplication relation. Rather the solution may involve responding to both relations simultaneously or in a specific sequence. The history of exemplar training required to produce these highly subtle forms of contextual control over relational responding needs to be considerably extended. Indeed, the ability to solve such problems at a high level of fluency may not be routinely established for many verbally able adults by our educational systems.

Other relations of interest in the development of a full and rich repertoire of relational responding (e.g., *opposition, comparison,* etc.) are clearly applied in correctly responding to many items on standard IQ tests. In the WISC-III^{UK}, for example, evidence of tests for temporal relations can be found in the Information subtest. This contains questions such as; "Which month comes next after April?" and "What is the day that comes after Friday?" Further examples of temporal relations can be seen in the Picture Arrangement subtest, in which a child is presented with several cards that depict a short story in a comic-strip format. The task requires that the child arranges the cards so that they tell a story that makes sense in real time. For example, a girl must take money out of her wallet before she can put it in the vending machine, she must then put money in a vending machine before she can choose a chocolate bar. Finally, she must choose the chocolate bar before the vending machine will dispense it (see Wechsler, 1992). If a child does not have a previously established history of flexible and richly contextually controlled temporal responding, they will not be able to complete these tasks using novel stimuli.

Examples of the relation of hierarchy, or what we might call "*containment*" can be found in the questions; "What is water made of?" and "What is the main material used to make glass?" These test items require participants to respond to the arrangement of substances in relation to each other in the context of water. So for example, the answer "molecules" is insufficient because all objects are ultimately made of molecules. Instead, what is required is to organise the levels of object structure so that the next lowest level of object structure beneath water as a chemical compound is named correctly. This requires responding to the materials inherent in water in the correct hierarchical order.

The Similarities subtest of the WISC- III^{UK} presents examples of relational tests for frames of coordination (or stimulus equivalence) that are often quite abstracted (i.e., arbitrarily applicable). Specifically, one question in this subtest is; "In what way are a piano and a guitar alike?" This question is clearly analogical. That is, the question involves responding to one stimulus item as equivalent to another in terms of a further set of topographical or arbitrary features. In this example, that further set of features happens to be topographical (i.e., both are musical instruments with steel strings). In answering correctly, the individual is responding to the common classification of both stimuli as musical instruments. In other words, the stimuli share a common *equivalence* relation with the term “musical instrument”. In fact they are even defined as such. Thus, the task is examining two very frequently encountered verbal relations and the subject’s ability to respond to these two relations as involving a common member. Responding in this way requires a rich history of responding to the test items in a variety of different contexts including both the relationships among the stimuli and the functions of the stimuli.

Further examples of simple analogical tasks can be seen in the WISC-III^{UK} within the Similarities subtest. For instance, in that subtest the question is asked, “In what way are a painting and statue alike? A painting and a statue are both members of equivalence relations with the term “art”. In other words, they bear the same relationship to art; this is what they have in common. Another question asks; “In what way are rubber and paper the same?” In this example, the commonality is that both rubber and paper are manufactured from trees. This example is based on a more abstracted commonality among stimulus items than seen in simpler analogies. This task requires that a child can identify a commonality between items that would not usually be thought of as similar. Indeed, in one test item commonalities must be discriminated between items normally responded to as the opposite of each other (i.e., “How are anger and sadness the same? and “In what ways are first and last alike?”).

The foregoing types of relational tasks can also be seen in other tests of intelligence, such as the Cognitive Abilities Test (CAT), and the AH4. In this section I will briefly outline these tests and provide examples of relational skills assessed by each.

The CAT is a group-administered test intended to provide a set of measures of an individual’s ability to use and manipulate abstract and symbolic relationships

(Thorndike, Hagen, & France, 1986). Thorndike et al. (1986) have explicitly described the test items as providing an index of "relational thinking" (p. 1). They define relational thinking as the "perceiving of relationships among abstract elements in a variety of media and settings" (p. 1). The CAT is composed of three batteries: a verbal, a quantitative and a nonverbal battery.

The verbal battery of the CAT is designed to appraise "relational thinking" when the relationships are formed in verbal terms (p. 1). From the RFT perspective, this test is clearly composed of probes for equivalence relations among stimuli. An example of one test item involves presenting a child with a word, such as "change", in bold print, and asking the child to pick the word that has the same meaning from a further list of words, such as; "leave, loose, coins, fasten, noise". Another test in this section presents the student with the following incomplete statement; "Jack, Jim and Charles are ______." The child must choose the best answer from the following list: "sisters, daughters, mothers, brothers, grandmothers". This item clearly requires the child to tact the equivalence relation that obtains between the stimuli presented in the prompt. In another subtest of the verbal battery, a child is presented with a word list and informed that the words are alike in some way. For example, they may be presented with; "gaze, glance, stare" and asked to choose which of the following words belongs in the foregoing list; "wonder, dream, notice, study, look". This probe for effective knowledge of synonyms would appear to represent a clear example of a test for stimulus equivalence among verbal stimuli.

The quantitative battery on the CAT is designed to appraise a pupil's perception of relationships among concepts. From a RFT perspective, these tasks assess more-than and less-than relations between numerical stimuli, which are themselves products of relational responding to pairs of items. For example, the student is presented with two columns of items and asked to mark A if column I is more than column II, to mark B if column I is less than column II and to mark C if column I is equal to column II. In this exercise, column I might consist of something like; “25% of 200” and column II might consist of “50% of 100”. In this case, calculating a percentage requires the student to respond analogically to each stimulus pair. More specifically, the student must respond to the relation between the numbers presented in the first stimulus pair (e.g., 25 and 200) in terms of another stimulus relation not present. That is, the student must respond to the relation between 25 and 100 (i.e., the first relational response required in order to respond correctly to percentage problems; in this case the former is one quarter of the latter) and apply this relation arbitrarily to 200. When one quarter of 200 is responded to as 50, the student has identified that the relation between 25 and 100 *is the same as* the relation between 50 and 200. Thus, the first behavioural product is 50. The second behavioural product (50% of 100) can now be calculated in the same way (the answer is also 50). The relation that obtains between these items presented in a given sequence (i.e., equivalence) can now be tacted by a student exposed to a sufficient number of more-than, less-than, and equivalence exemplar tasks.

In another version of the problem, column I might consist of the “Cost of ten lemons at 3 for 13p” and column II might consist of the “Cost of ten lemons at 4 for 15p”. Thus, the subject is being asked again to tact the relationships among the complex verbal stimuli, which essentially produce the same behavioural product (e.g., variously described amounts of money have equal reinforcing value) or share the same behavioural functions despite obvious topographical differences.

Finally, the non-verbal battery of the CAT tests identification of, and flexibility in manipulating relationships which are expressed as figural symbols or patterns (Thorndike et. al., 1986, p. 1). For example, in one item the student is presented with a small white circle on top of a small white circle, a small white diamond on top of a small white diamond and a small white triangle on top of a small white triangle. The student is then asked to choose a drawing that goes with the first three from a sample of; a large white diamond, a small white semi-circle on top of a small white semi-circle, a large light shaded rectangle, a small white sideways triangle on top of a black sideways triangle and a large white semi-circle beside a large white semi-circle, where the semi-circles are facing in opposite directions.

From a RFT perspective, the foregoing is a test for analogical reasoning. According to RFT, an analogy is established when the trained or derived relations in one network of relations are placed in a frame of *coordination* with the trained or derived relations in another network of relations (Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001; see below for more detail on the issue of analogy). Thus, the foregoing test item assesses a participant's ability to tact the common relation between sets of relations. In the above example, the relations characterised by the geometric shape pairs are all of *equivalence* (i.e., the shapes appearing as pairs are all the *same as* each other). However, in the comparison set of stimuli, only one of these pairs contains geometric shapes that are the *same as* each other (i.e., a small white semi-circle on top of a small white semi-circle). Thus, the participant taking the test must choose which of the comparison pairs is characterised by the same stimulus relation (e.g., *same, opposite,* et cetera.) as that characterised by all the stimulus pairs in the sample set. RFT views this sort of analogical responding as a higher order level of relational ability.

The AH4, developed by Heim, Watts and Simmonds, is designed as a group test of general intelligence for use with an adult population (Heim, Watts, &

Simmonds, 1968, 1975). In the AH4, there are also many examples of relational skill tests. Probes for derived relations of opposite can be seen in questions like; “Up means the opposite of; 1) short, 2) small, 3) low, 4) down, 5) young”, and; “Near means the opposite of; 1) close, 2) road, 3) speed, 4) far, 5) distance”. An example of a synonym test item that requires responding to equivalence relations between words is; “Ill means the same as; 1) health, 2) fever, 3) dirty, 4) mumps, 5) sick”. A further example is; “Portion means the same as; 1) some, 2) whole, 3) part, 4) any, 5) cake”.

Examples of test items that require responding to relations between relations (i.e., analogy) can be seen in questions such as; “Army is to navy as soldiers is to; 1) airman, 2) sea, 3) service, 4) sailor, 5) uniform”. Finally, clear examples of tests for *larger than/smaller than*, *before/after*, *if/then* relations and number series problems can be seen in questions such as; “If a castle is bigger than a cottage, write down the second number of these figures: 1, 2, 3, 4, 5, 6, 7, 8, 9. “If it is not, write down the sixth.” In these test items the student is asked to tact the increasingly complex and abstract relations among stimuli and among relations among relations.

Other Types of Relational Skills

So far it has been illustrated the relational features of many test items across a range of commonly used tests of intelligence and general cognitive ability. However, it is important to understand that not all test items require responding to merely one type of derived relation specified by one contextual cue (e.g., the phrase “choose the word that has the *same* meaning as…”). Some test items require responding to multiple stimulus relations simultaneously, or in a particular sequence. Tests for analogy represent a good example of such task types. Thus, analogy represents a higher-order level of relational responding insofar as it involves multiple stimulus relations, but is nevertheless applicable to a wide variety of stimulus topographies.

According to RFT, analogical relational responding, or responding to relations among relations, is a particularly important form of relational responding as it underlies many forms of problem solving. Indeed, analogical abilities are commonly considered a measure of intelligent behaviour in their own right (e.g., Sternberg, 1977) and are frequently used to predict academic success (e.g., in the Graduate Record Exam or the Miller Analogies Test).

Stewart, Barnes-Holmes, Hayes, and Lipkens (2001) described analogies as equivalence relations between two relational networks. Thus, analogy might be construed as an example of relations among relations. For example, if a person is directly trained to relate A to B and C to D in the presence of a contextual cue for a frame of coordination, then an analogous relation is entailed between the trained relations (see also Stewart & Barnes-Holmes, 2001; Stewart & Barnes-Holmes, 2004; Stewart, Barnes-Holmes, & Roche, 2004; Stewart, Barnes-Holmes, Roche, & Smeets, 2001).

The Stewart et al. (2001) model of analogy also allows us to make sense of metaphor in relational terms. Specifically, the model defines a metaphor as the transfer of stimulus function based on shared properties. In other words, metaphor requires that relations between relations be based on common psychological functions transforming in accordance with stimulus relations in two separate relational networks. Stewart et al. provided the following example; *"Struggling with anxiety is like struggling in quicksand"*. In this metaphor, the reader experiences the psychological functions of both struggling with quicksand and anxiety (i.e., because of the participation of these words in further equivalence relations with other words and stimuli). However, they are said to "understand" the metaphor only when they can tact the formal similarity of the functions transforming in each case. That is, the reader is first required to covertly experience the psychological effects of sinking in quicksand present for the word "struggle". In effect, the functions of the word "struggle" are transformed by the word "quicksand" (and vice-versa) such that a resulting sense of hopelessness is experienced by the reader. Similarly, the response functions of the word "struggle" are also transformed by the functions of the word "anxiety" (and vice-versa) such that a similar experience of hopelessness is experienced in the presence of the word struggle. If the reader has never experienced the hopelessness of struggling with anxiety they will likely not be able to make the appropriate tact response and will therefore be said to not understand the metaphor. If the reader does make the appropriate tact response, their future response to anxiety may change, such that he or she no longer attempts to resist his or her anxiety. In this case, the appropriate response functions of quicksand (i.e., do not struggle) have transformed the functions of anxiety (i.e., do not struggle) because of the analogical relationship that obtains between the two sets of relations involved.

Analogical and metaphorical responding is often required by intelligence test items and a modern analysis of these response repertoires is now provided by RFT that has already undergone empirical investigation (e.g., Stewart & Barnes-Holmes, 2001; Stewart & Barnes-Holmes, 2004; Stewart, et al., 2004; Stewart, et al., 2001).

Another obvious skill necessary for performing well on intelligence tests is the ability to "remember". Memory is a term that behaviour analysts have attempted to approach functionally in the past. For instance, in *About Behaviorism,* Skinner (1974) suggested that being "reminded" is simply "being made likely to respond, possibly perceptually" (p. 171). In other words, memory need not be assumed to be an active system of information retrieval. Instead, from a functional-analytic perspective, remembering involves behaving as if a stimulus were present, or as we behaved in the presence of that stimulus on some earlier occasion. RFT extends this Skinnerian paradigm by introducing the concept of relational networks. From a RFT point of view, memory can be considered an elaborated relational network according to which the individual responds. The greater the number of stimuli and relations involved in a network, and the greater the amount of contextual control that can be exerted over relational responding, the greater an individual's memory is said to be. More specifically, the more relations that obtain between stimuli in a network, the easier it will be to elicit an appropriate response to any given stimulus, because not only do multiple relations obtain between each stimulus and several others, but the transformation of response functions by some stimuli over others is more likely.

To illustrate this idea, imagine presenting a picture of a dog to a young child who has been taught only one name for this item, and therefore can be said to have a poor repertoire of relational responding. In this case, the one and only possible answer to the question; "What is this?" is the vocal utterance of the word "dog". Imagine now, however, that for a second child, the picture of a dog participates in a rich relational network with several possible word candidates (e.g., chien [French], madra [Irish], and dog [English]). In terms of the current framework, these words participate in derived relations with each other (i.e., they are multilingual synonyms). In this case, several answers to the foregoing question are both possible and correct. In a very literal sense, when relational networks are rich there are more complex psychological responses occurring than when networks are barren. Therefore, the formal properties of the various words may be produced by the presentation of the picture (i.e., the child may vocally produce more than one word). In addition, the presentation of any of the synonyms produces the functions of the picture (e.g., the child may "see" the dog in its absence, or "remember" it in Skinner's terms). In effect, the richer the relational network, the more psychological activity that occurs upon the presentation of any of the relevant stimuli.

The establishment of rich relational networks not only expands the repertoire relevant to responding to standard IQ test items but it also raises the chances of appropriate responses being made in the presence of stimuli in the real world. Thus, rich relational networks, or memory, may be an important factor in determining one's responses to IQ test items and therefore in determining one's overall score. Indeed, Bors & Forrin (1995) found that performance on a free recall task had a substantial independent effect on a common measure of fluid intelligence, the Raven Advanced Progressive Matrices (RAPM), when age and latency were held constant. Given all of the foregoing, if levels of fluency in responding to analogies and responding to increasingly large and complex relational networks can be enhanced, it should be possible to create measurable improvement to intelligence test scores.

Developing Relational Skills Interventions to Raise IQ.

The unique contribution of RFT to understanding intellectual development stems from the fact that it suggests improvements for educational technologies that are traditionally concerned with content delivery rather than behavioural process. RFT identifies some of the core processes that may be involved in the acquisition of a relational repertoire and thereby presents the potential of an immediate empirical analysis of intelligent behaviour as an over-arching set of relational skills. Moreover, it suggests specific testable techniques (e.g., Multiple Exemplar Training: MET) to examine the expansion of verbal relations and their increasing sensitivity to contextual control. Skinner (1959) supplied a functional-analytic framework for this work and RFT extends this approach by adding the concept of the derived stimulus relation (see D. Barnes-Holmes, Y. Barnes-Holmes, & Cullinan, 2000).

The RFT account of intelligent behaviour suggests that it should be possible to raise the intelligence quotient (as measured by commonly used IQ tests) of an individual if a sufficiently comprehensive and appropriate relational responding intervention is employed. RFT suggests a Multiple Exemplar Training (MET) intervention that would be appropriate to this end. Put simply, this intervention involves training children in the core relational skills, such as deriving relations in accordance with a wide variety of relational frames and across a very large number of exemplars. Once such component relational skills are established and sufficiently generalised across novel stimulus sets, a child should be able to respond appropriately to an almost infinite number of other similar relational tasks. Consequently, their ability to respond appropriately to the relational tasks presented in common IQ tests should be enhanced.

Traditionally, psychometricians view intelligence as an invariant trait that is more or less normally distributed across the population. The idea that test scores are constrained by stable innate abilities is supported by the fact that quotient scores change little across the lifetime. For instance, several studies have been conducted that provide evidence for the concurrent validity and reliability of the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1992). These studies provide strong evidence for the longitudinal stability of global IQ estimates derived from the Wechsler scales (Sattler, 1988). Other studies have shown acceptable levels of predictive validity for black and white children (Hartlage & Steele, 1977; Juliano, Haddad, & Carroll, 1988; D. J. Reschly & J. E. Reschly, 1979) and comparable validity among males and females (Reynolds, Gutkin, Dappen, & Wright, 1979). In effect, it is generally accepted among psychometricians that the construct of IQ is relatively stable across a lifetime and across the population. In effect, the stability of intelligence quotient scores (as opposed to raw scores, which change significantly across the lifespan) is used as evidence that the underlying trait is itself stable (Gardner, 1993). Given these traditional views, it would seem that intelligence quotients cannot in principle be raised (see Gardner, 1993, for criticisms of this approach).

Psychometric tests are constructed carefully so that they are not prone to practice effects. Arguably, it is the ability of scientists to create reliable tests that has made the idea of an invariant construct, such as general intelligence, so seductive. In fact, raw IQ scores do typically rise by a considerable amount across a lifetime and measurably so from year to year, and even from quarter to quarter. This effect is called the "IQ drift" and psychometricians compensate for its disruptive effect on the stability and distribution of IQ scores by; a) adjusting for chronological age in calculating IQ scores, and; b) revising IQ tests every decade or so. Even though these practices may seem suspect to the behaviour analyst, it makes sense from a psychometric perspective to adjust IQ tests to better measure a 'known' construct with known statistical properties.

The statistically generated normal distribution of IQ test scores is employed to provide each individual taking an IQ test with a score relative to the general population or a relevant group of peers. This relative score is calculated based on demographic characteristics such as age and gender. These factors are used to alter the score proportionate to their known impact on the raw test score, which changes as a function of these variables. In effect, the relative rarity of rises in IQ is attributable to the fact that IQ scores are corrected statistically by precisely that factor required to keep scores constant given the known effects on scores due to increasing age and practice. Thus, a very large improvement in raw IQ score would be required in order for a significant change in IQ (e.g., one standard deviation or a move from one diagnostic range to another) to be recorded. With the foregoing in mind, a behavioural approach to raising IQ scores may not appear to be feasible. Specifically, what is required is an intervention to raise intellectual skills, sufficient to move raw IQ test or subtest scores (i.e., before normalisation techniques are applied) more than they typically do in a given period of time.

Previous behaviour-analytic studies have included IQ test measures as part of interventions for severe disability. For example, Lovaas (1987) reported IQ gains as large as 30 points from the outset of a three year intensive ABA intervention for autism. Just under half of the children that took part in that study appeared to "recover" from autism, in that they were not noticeably different from normally functioning children after three years (Reed, Osbourne, & Corness, 2005). Unconvinced of the reliability of the reported IQ rises, Reed et al. (2005) raised concerns regarding the internal and external validity of the study (see also Connor, 1998; Gresham & MacMillan, 1997). Magiati and Howlin (2001) also criticised the study on the grounds that different IQ tests were often used at baseline and at follow up, thereby reducing the reliability of the measurement. In addition, these researchers pointed to a series of serious methodological flaws regarding subject selection, treatment condition assignment, differing treatment periods across the experimental and control groups, and the already high-functioning intellectual ability of the treatment group. Nevertheless, Sallows and Graupner (2005) also recorded significant IQ rises in a more recent replication of the Lovaas (1987) study.

In a further study, Smith, Eikeseth, Klevstrang and Lovaas (1997) studied IQ, expressive speech and adaptive behaviour improvements among severely mentally retarded children with autistic features during an ABA intervention. Children exposed to the treatment condition displayed a higher mean IQ at follow-up and evinced more expressive speech than did those in the comparison group. Behavioural problems diminished in both groups. While the work of Smith et al. (1997) and the other studies outlined here have shown promise that behavioural interventions may lead to IQ rises, these research programmes were concerned with IQ only as one part of a larger range of dependent measures in wide–ranging and multifaceted studies. These studies typically involved interventions to improve the autistic condition and/or other pervasive developmental difficulties. What is required, however, is a focused approach to understanding what we mean by intelligence from a behavioural perspective, and a targeted programme of research and intervention to illustrate that intellectual skills can be brought under operant control.

The RFT approach suggests the possibility of a dedicated programme of research and intervention to raise IQ scores among both normally developing and developmentally delayed individuals. Specifically, RFT offers the advantage of a well worked out nomenclature for examining the types of test items seen on IQ tests. This nomenclature allows the RFT researcher to distinguish between test items and to functionally organize ranges of test items in ways useful to the behaviour analyst. For instance, test items can be classified in terms of the types of relational skills involved so that appropriate MET can be administered for those items. This type of functional analysis of test content is important because a behavioural approach would require that we understand the functional relations that obtain between relational repertoires of different types. By applying a well worked out taxonomy of relational skills, the RFT approach should allow researchers and therapists to more accurately identify with precision which aspects of an IQ test pose a problem for a particular client.

There is no published study to date that examines the possibility of raising IQ using a RFT intervention. However, a small number of published studies to date have systematically investigated the impact of a derived relations intervention on relational skills repertoires. These studies will be reviewed here to provide the reader with detailed examples of the types of interventions that may be of use in a behavioural programme to enhance general intellectual skills.

In the first of these studies, Y. Barnes-Holmes, D. Barnes-Holmes, and Roche (2001) found that explicit exemplar training is a reliable means by which to facilitate generalisation of a relational skill in accordance with *symmetry*. In this study, the authors employed sixteen children (aged four-five) across four experiments (i.e., four children in each experiment). In the first experiment, participants were first trained to name two actions and two objects by demonstrating listening, echoic, and tacting behaviours. Participants were then trained in an action-object conditional discrimination using the previously named actions and objects. Participants were then re-exposed to the name training, before being exposed to a test for derived object-action symmetry relations. Across subsequent sessions, a multiple-baseline design was used to introduce MET (i.e., explicit symmetry training) for those participants who failed the symmetry test. The second experiment replicated the first experiment, except that the name re-training (between the conditional discrimination training and symmetry test) was removed. The third experiment also replicated the first experiment, except that participants were trained to tact all of the actions and objects during conditional discrimination training and symmetry testing. The fourth experiment also replicated the first experiment, but with a reversal of the trained and tested relations. Across the four experiments, 13 out of 16 participants failed to show derived object-action (Experiments 1-3) or action-object (Experiment 4) symmetry until they received explicit symmetry training.

The foregoing effect was also demonstrated in a second study by Y. Barnes-Holmes, D. Barnes-Holmes, Roche, and Smeets (2001). The main purpose of this follow-up study was to determine whether participants would demonstrate the transformation of function in accordance with symmetry relations if provided with exemplar training in symmetry, but not provided with explicit name training. This study also examined whether pre-training that was formally similar to the symmetry test, but that did not reinforce symmetry relations, would have the same facilitative effect as did MET in the original study. Sixteen children (between the ages of four and five years) were employed across three experiments (i.e., four children each in Experiments 1 and 2, and eight children in Experiment 3). Results indicated that across Experiments 1 and 2, none of the eight participants showed derived object-action (Experiment 1) or action-object (Experiment 2) symmetry until they received explicit symmetry training. Pre-training object-action responding in Experiment 3 appeared to facilitate symmetry, but only for four of the eight participants. For the four participants who failed, symmetry emerged following exposure to MET.

A more recent study made the first attempt to generate repertoires of relational responding, in accordance with *more-than* and *less-than*, using MET when these skills were absent in young children (Y. Barnes-Holmes, D. Barnes-Holmes, Smeets, Strand, & Friman, 2004). Using interventions suggested by RFT, three children were exposed to a basic problem-solving task that involved two or three identically sized paper coins in an attempt to test and train patterns of relational responding in accordance with more-than and less-than relations. On each trial, the experimenter described how the coins compared to one another in terms of their value, and the child was then asked to pick the coin that would "buy as many sweets as possible". All three participants failed to pass baseline tests for specific patterns of derived more-than and less-than responding. Interventions suggested by RFT, including training and testing across stimulus sets, were then successfully used to establish increasingly complex patterns of relational responding in all three children. Generalisation tests demonstrated that the relational responding successfully generalised to novel stimuli and to a novel experimenter.

In a further study, (Y. Barnes-Holmes, D. Barnes-Holmes, & Smeets, 2004), children were trained to relate stimuli in accordance with relations of opposition and then to derive novel *same* and *opposite* relations across several sets (e.g., the opposite of an opposite is a same, but the opposite of an opposite of an opposite is an opposite). In effect, participants were presented with a sample derived relations problem and then re-presented with the same problem involving different stimuli. Initially, all three participants failed to pass baseline tests for specific patterns of relational responding in accordance with opposite relations. Various interventions, including training and testing across different stimulus sets and across different numbers of sets, were then successfully used with all participants to establish these relational responses as well as increasingly complex patterns of opposite responding. Generalisation tests also demonstrated that the relational responding generalised to novel stimuli and to a novel experimenter.

In a recent study, Berens and Hayes (2007) systematically tested the impact of each of several phases of MET on the derivation of the entire frame of comparison. Their participants included four female participants, ages four-five years old, all of whom could not perform a series of problem solving tasks involving arbitrary *more-than* and *less-than* relations. Each child was first administered the Vineland Adaptive

Behavior Scale to get a general picture of their individual ability levels. Stimuli included three sets of uniquely colored paper pictures. Each session began with the experimenter telling the child that they were going to play a game and that the child's job was to pick the picture that would buy them the most candy. Trials consisted of linear relations (A >B or A<B) and mixed non-linear trials (A>B>C and A<B and C<B). Responses were followed by contingent feedback. Reaching accuracy goals were reinforced with prizes chosen by participants' at the beginning of each session when the goals were set. Non-contingent reinforcement was provided during baseline and probe conditions due to the considerable length of these conditions. A multiple probe across stimulus sets was employed to evaluate the degree to which reinforced responding with the targeted stimulus set generalised to untrained stimulus sets. A multiple baseline across participants was employed to control for maturation and extra-experimental conditions. The study found that reinforced MET facilitated the development of arbitrary comparative relations, and that these skills generalised across stimuli and trial types.

Finally, and most recently, Gomez, Lopez, Martin, Y. Barnes-Holmes, & D. Barnes-Holmes (2007) partially extended the research of Y. Barnes-Holmes, D. Barnes-Holmes, Smeets, Strand, & Friman (2004). These researchers employed an almost identical methodology in which four 4-5 year old children were trained to name two actions and two objects by demonstrating listening, echoic and tacting behaviours. These children were then trained in action-object conditional discriminations using the previously named actions and objects. Children were then exposed to a symmetry test for derived object action relations. If they failed the symmetry test, a multiple baseline design was used to introduce exemplar training. A further part of the study involved training in an action-object conditional discrimination and then training to establish naming of two novel actions demonstrating listening, echoic and tacting behaviours. The children were then trained in a second action-object conditional discrimination using two novel actions just learned and the same objects. Subsequently, they were exposed to a symmetry object-action test using novel actions learned. If participants passed the symmetry test, they were exposed to an equivalence action-action test. If they failed equivalence across subsequent sessions, a multiple baseline design was used to introduce exemplar training for those children who failed. The researchers concluded that MET training is an effective means of establishing symmetry and equivalence when found to be absent in young children. They also suggested that RFT can make an important contribution to the development of interventions that may prove effective when important relational skills are deficient or absent.

Conclusion

If there is a functional relationship between DRR and language skills, as suggested here, an improvement in DRR repertoires may well produce a measurable change in language ability (i.e., acquisition rates, fluency, and extent of vocabulary). This is a remarkably exciting possibility for behaviour analysts given the already healthy state of the research on multiple exemplar technology. These improvements in relational ability may also lead to modest or even dramatic rises in overall IQ scores, or on specific dimensions or subtests of IQ, as argued in this thesis. Thus, one true value of research into derived relational responding will be found in the educational programmes that might be established to produce changes in the intellectual abilities of children.

Of course, over time the relative impact of relational skills levels on overall IQ scores may emerge. It is likely that other factors, such as attentional skills, the absence of sensory deficits and other diagnosed behavioural and emotional difficulties, are likely to also play an important role. Thus, the effects of relational interventions may not be linear or easily predictable without understanding their relationship to a whole host of other important educational, social, biological, and psychological variables that have been studied by behaviour analysts for the past half-century. Only the efforts made by researchers to address these research issues will help us to determine whether or not the RFT approach to intellectual deficit will be sufficiently useful in making a real difference to the relational repertoires, educability and lives of those who most need our help. Armed with RFT as a conceptual framework and a touchstone for the development of practical interventions, behaviour analysts are poised to make what might be our most impressive contribution yet to the world of psychology; the establishment of environmental control over "invariant traits".

Chapter 2

Examining Possible Correlations Between Measured IQ and Derived Relational Responding on a Stimulus Equivalence Task

Experiments 1-2

Introduction

As argued in the previous chapter, we might expect to see a functional relationship between Derived Relational Responding (DRR) ability and general cognitive ability, as measured by commonly used IQ tests. The first obvious step in investigating these relationships is to conduct an analysis of the co-variations in the relevant variables of interest. Specifically, we need to know if IQ and DRR actually correlate, but more importantly we require a detailed analysis of the subtest scores on one or several IQ tests and their correlations with DRR. Such an analysis would shed light on the precise dimensions of IQ that correlate most highly with DRR.

One study by O'Hora, Pelaez and Barnes-Holmes (2005) strongly supports the foregoing RFT predictions that there are correlations between verbal abilities on standardised IQ tests and relational skills measured in the laboratory setting. In this study, O'Hora et al. (2005) sought to determine if performance on relational tasks would predict performance on verbal or performance subtests of the Wechsler Adult Intelligence Scale III (WAIS III). In this research 75 participants were assigned to one of two groups based on their performance on a complex relational task involving pre-training and testing for Before/After and Same/Different Relations, a test for instructional control and a test for generalisation of instructional control using novel stimuli. These participants were then exposed to three subtests of WAIS III (vocabulary, arithmetic, and digit-symbol coding). Participants who successfully completed the relational task (n=31) performed significantly better on the Vocabulary and Arithmetic subtests than those participants (n=44) who failed to do so. No significant differences in relational task performances were found between groups for the Digit-symbol encoding subtest. As stated, these findings support the position that derived relational performances may be related to human language abilities, as measured by subtests of the widely utilised WAIS III.

More recently, a research programme conducted by O'Toole and Barnes-Holmes (in press) involved presenting 62 participants with before/after and similar/different relational tasks using the Implicit Relational Assessment Procedure (IRAP). After controlling for confounding effects, 55 of the 62 participants were then administered the Kaufman Brief Intelligence Test (K-BIT). Response latencies on consistent and inconsistent trials were measured. In addition, difference scores were calculated by subtracting consistent response latencies from inconsistent response latencies. The response latencies and difference scores were considered to provide measures of relational flexibility. Results of this research programme showed that faster responding on the IRAP and smaller difference scores predicted IQ. Findings suggested that response speed on trials that were inconsistent might provide a possibly useful measure of relational or cognitive flexibility. O'Toole and colleague note that assessing relational flexibility may be particularly advantageous because flexibility is widely regarded to be an important component of human cognitive abilities (e.g., Cattell, 1971, Kyllonen, Lohman, & Woltz, 1984; Premack, 2004). For these reasons, O'Toole et al. (in press) suggest relational flexibility as something to be targeted in educational settings.

The current research programme aims, in part, to extend the O'Hora et al. (2005) research by attempting to identify correlations between basic derived relational responding skills (i.e., the ability to derive symmetry and transitivity relations) and any and all subtests of the WISC-III^{UK}. O'Hora et al. adopted the use of three subtests which have high reliability coefficients (.81 or higher). The O'Hora et al. research supports the supposition that correlations between some subtests (e.g., verbal subtests) and DRR ability would be stronger than correlations between other subtests (e.g., performance subtests) and DRR ability. Thus, although the subtests used by O'Hora et al. proved to be a good starting point, a thorough analysis of the various correlations between IQ and DRR needs to be more inductive than deductive. As such, the entire battery of subtests from an individual IQ test should be administered in such a study.

Administering an entire battery of subtests from a more comprehensive intelligence test was suggested by O'Toole et al. (in press). In this piece of research, it was reported that while the K-BIT was a complete intelligence test, it was a brief intelligence test. For this reason, O'Toole et al. suggested that one improvement for further research might include the administration of a more extensive intelligence test, such as the Wechsler Adult Intelligence Scales Revised (WAIS-R; Wechsler, 1981). O'Toole and colleague note a correlation coefficient of .75 between the K-BIT composite IQ scores and full scale IQ scores on the Wechsler Adult Intelligence Scales Revised (WAIS-R; Wechsler, 1981; see also A. S. Kaufman & N. L. Kaufman, 1990). Moreover, O'Toole et al. (in press) also report a correlation coefficient of .60 between the Vocabulary subtest of the K-BIT and the verbal IQ scale on the WAIS-R. Furthermore, a correlation coefficient of .52 is reported between the Matrices subtest on the K-BIT with the performance IQ scale on the WAIS-R (see O'Toole et al., in press; see also A. S. Kaufman & N. L. Kaufman, 1990). Thus, similar findings to those of O'Toole and colleague might be expected from IQ tests which have strong construct validity and that are correlated with each other.

Other ways that the current work extends upon previous work are noted here. Specifically the current work extends upon the research of O'Hora et al. (2005) in several important ways. Firstly, the current study administered all subtests of the

WISC-III[UK], rather than a selected set based on deductive reasoning. Secondly, the administration of the full test allowed for the calculation of separate verbal and performance IQ scores, in addition to full scale IQ scores. This is important, because RFT predicts that DRR should be highly correlated with verbal IQ. Thirdly, the current study made use of child participants with a view to establishing an intervention for use with children in need of educational interventions. Fourthly, the current study also employed a more traditional DRR task as the baseline measure of DRR skill. In the O'Hora et al. (2005) study, a complex rule following task was used as a baseline measure of DRR skill. Such a complex DRR task is inappropriate for use with young children. Moreover, it is important to begin an extended research programme with the intention of identifying the simplest possible DRR correlates of IQ. Beginning with a highly complex task would not allow us to ascertain which of the relations involved, or which combinations of the relations involved, serves as the best proxy measure of IQ. Thus, in the current study, a traditional stimulus equivalence (SE) task was employed to examine levels of DRR skill.

An additional benefit to using a standard SE task over a complex rule as measure of DRR skill is that SE, or a relational frame of sameness, is considered to be the most fundamental of the derived relations. Specifically, according to Hayes, Fox, Gifford, Wilson, Barnes-Holmes, and Healy (2001), equivalence relations are the relations with which most research into derived relational responding is concerned. More importantly, however, much of the earliest language training received by children seems to be of this kind. Thus, a relational frame of coordination is probably the first to be abstracted sufficiently that its application becomes arbitrary (Hayes, Fox, et al., 2001).

Given the foregoing, it would seem that SE may correlate more highly with IQ and the various subtests of an IQ test than will a complex rule. Of course such a suggestion is mere conjecture in the absence of data and this issue will remain empirical until research of the kind suggested here is conducted. With this in mind, the current research sought to assess possible correlations between intelligence (as measured by IQ) and its subtests, and competence on a stimulus equivalence DRR task.

Twelve normally developing children were recruited to take part in Experiment 1 of the study. This study involved administration of an IQ test, Wechsler Intelligence Scale for Children, Third Edition, UK (WISC-III^{UK}) and a SE task. A Spearman's Rho correlational analysis was then conducted to examine the relationship between IQ scores and fluency at forming the baseline conditional discriminations employed during equivalence training. Further analyses examined correlations between IQ scores and; (a) fluency at forming symmetrical relations; (b) fluency at forming transitive relations; (c) fluency at forming symmetrical and transitive relations by combining scores from symmetry and transitivity tests (i.e., stimulus equivalence).

In Experiment 2, eight of the children from Experiment 1 were re-recruited to take part in an intervention. This intervention involved providing participants with baseline conditional discrimination training as in Experiment 1, but using a novel stimulus set, and testing for symmetrical and transitive responding with corrective feedback. Following a test for symmetry and transitivity with feedback, participants were exposed to a further novel stimulus set of conditional discrimination training and then exposed to testing for symmetrical and transitive responding without feedback. This method of cycling testing with and without feedback continued until each participant had completed the intervention programme.

The aim of the multiple exemplar training intervention was to supplement the basic training and testing from Experiment 1 by increasing relational fluency. Fluency is defined here as producing derived relations on a novel set but in a minimal number of testing trials. It was expected that this might lead to rises in overall IQ scores or in one or more IQ subtests.

Experiment 1

Method

Participants

Twelve normally developing children (ages 8-12 at the start of the experiment) were recruited from a local school and from friends and family of the researcher (see Table 1 for exact ages). Ten of these children were female and two were male. Children were chosen based on availability, and also on having been identified by parents and teachers as students who were not presenting with any known or suspected learning difficulties. No child had previously been exposed to any IQ testing or to research involving derived relational responding. For these reasons, all participants were considered naïve. Research commenced following the provision of detailed descriptions of the research to the relevant authorities and to parents (see Appendix 1). All parents were also provided with a detailed consent form (see Appendix 2), and weekly consultations regarding their child's participation and progress through the research programme. The reader should note that these consultations included only generic information such as whether or not each child was enjoying the tasks and moving at a rate that meant that they could continue to participate in the study. Parents were not ever informed as to the precise details of the stimulus equivalence training and testing protocol. This was done in order to prevent confounding of the measure by parental interventions (verbal or otherwise) between sessions without the experimenter's awareness.

One of these twelve children (P5) was omitted from the research study because after four one hour sessions, she was unable to pass the required baseline conditional discrimination training. Her data will not be discussed.

Table 1. Participants' ages at start of the experiment.

Participant ID	Age
P1	12 years, 8 months
P2	10 years, 5 months
P3	10 years, 0 months
P4	8 years, 10 months
P5	9 years, 1 month
P6	10 years, 4 months
P7	10 years, 0 months
P8	10 years, 1 month
P9	10 years, 5 months
P10	9 years, 2 months
P11	10 years, 1 month
P12	12 years, 3 months

Setting and Materials

Initially, each child was administered the WISC-III^{UK} in his/her own home with a parent present in the house, but not in the same room. The WISC-III^{UK} is an individually administered clinical instrument for assessing the intellectual ability of children (Wechsler, 1992). The instrument is comprised of thirteen subtests. Twelve of these subtests were administered in this experiment and are described in Appendix 3. Wechsler (1992) divided a child's performance on these various subtests into categories which yield three composite scores (performance intelligence quotient, or PIQ, verbal intelligence quotient, or VIQ and full scale intelligence quotient, or FSIQ). The Wechsler scales have been almost universally adopted as the standard instrument for assessment of cognitive abilities in all parts of the world (Wechsler, 1992). For discussions on how the WISC-III^{UK} is widely regarded as a well developed measure of intelligence and cognitive ability and for further discussion on the development of intellectual measures see; Carroll (1989); Jensen (1980); Kaufman (1990); and Sattler (1988). In each instance, a quiet private room was used to minimise distractibility. Times for sessions were also chosen based on when there were least likely to be external distracters (such as the presence of other family members) in the house. A room was also chosen in each house based on the presence of a large table where each child could sit and work comfortably for the duration of each session. All conditional discrimination training, as well as symmetry and transitivity testing were administered on a Macintosh™ *ibook* lap-top computer. The conditional discrimination training and the symmetry and transitivity tests were controlled by software written by the author using *Psyscope*. *Psyscope* is a graphic interface and scripting language application for the creation of computer-controlled experiments in Psychology (Cohen, McWhinney, Flatt, & Provost, 1993). See also Roche & Dymond (1999). There were a total of 42 arbitrary nonsense syllables (A1, B1, C1, A2, B2, C2) employed as stimuli in the conditional discrimination training, symmetry and transitivity testing. These nonsense syllables are listed in Appendix 4.

All nonsense syllables were composed of three letters and all nonsense syllables were intended to be pronounceable.

Ethics

The design of this study received ethical approval by the NUIM Ethics Committee. The committee recommended clarification within the consent form that the research did not constitute any kind of counselling or medical treatment and that the study would not form any kind of medical diagnosis. Furthermore, the Ethics Committee recommended that parents be made aware that the study was experimental and not clinical in nature. These concerns were reflected in the consent form (see Appendix 2). Consent was obtained from each parent or set of parents of each participant. Verbal consent was also obtained from each child who participated. Information about the general nature of the study was also provided to each parent prior to commencement of the research (see Appendix 1). IQ test sessions were conducted in one sitting (approximately 1-2 hours), as is typically the case. Separate sessions for conditional discrimination training and for symmetry and transitivity testing were limited to one hour sessions or until the child requested to finish, or showed any signs of distress, such as crying, or hesitation in continuing. Each child was permitted to take as many breaks as they needed, although few of them ever took a break outside of an occasional trip to the bathroom. Parents were always in the vicinity of the experiment and were welcome to be present in the room during the experiment. It should be noted that no parent accepted this offer.

General Experimental Sequence

The experiment took place over the course of several 1-2 hour sessions, depending on the availability of the participants and their caregivers. These sessions consisted of five phases in total.

Phase 1 included the administration of the entire battery of a commonly used IQ test for children, the (WISC-III^{UK}). Phase 2 included the administration of a series of baseline conditional discriminations via a laptop computer. These baseline conditional discriminations formed the equivalence training procedure. Phase 3 included the administration of a symmetry test via laptop computer. Phase 4 included the administration of more conditional discrimination training using the same stimulus set employed in Phases 2 and 3. This training served to re-familiarise participants with the baseline conditional discriminations that they had learned during Phase 2. Finally, Phase 5 included the administration of a transitivity test employing the same stimulus set used for Phases 2, 3 and 4, also via laptop computer.

Phase 1: IQ Testing.

Each participant was seated comfortably at a large table in a quiet room. The experimenter sat opposing the participant. No one else was present in the room, although in every instance, the participants' parent(s) were either in the next room or in another room in the house. Parents were also informed that they should feel free to come in and out of the room as they chose. As mentioned previously, no parent did this. The participant was then asked if they were willing to participate in the research. If they responded, "yes" (e.g., that they were willing), then the experimenter proceeded to administer the WISC-III^{UK}. Every participant agreed to continue. The experimenter then proceeded to ask a series of questions about everyday situations, about general word usage, about basic computation and about abstract concepts, among other items. The experimenter also asked the participant to complete several timed tasks involving the manipulation of concrete objects (i.e., blocks, picture cards, jigsaw puzzle pieces, et cetera). See Wechsler (1992) for more information about the procedures for administration of the WISC-III^{UK}, as well as actual subtest items. Each IQ test administration took between one and two hours and each child was permitted to take as many breaks as they needed.

Following the IQ test administration, another appointment was set up with each child's parents for a date within one week of the IQ subtest. Appointments were always set for one-two hour sessions at the family's convenience. The second hour allowed time for set-up prior to commencement of each session and data entry, as well as packing up materials following each session. Appointments were not set at the end of a session if a child successfully met criterion on all phases.

Phase 2: Baseline Conditional Discrimination Training

During baseline conditional discrimination training, participants were again seated comfortably at a large table with the experimenter seated opposing them directly. Only one participant was seen at a time. During these sessions, participants were seated in front of a lap top computer for the duration of their training. A standard matching-to-sample procedure (e.g., Sidman 1971, 1986) was employed to train and test two three-member equivalence relations (i.e., A1-B1-C1 and A2-B2-C2), in a "linear" training protocol. This well-established methodology (see Barnes, 1994; see also Arntzen & Holth, 2000) involved the presentation of a sample stimulus in the centre of a computer screen, and the presentation of two further choice stimuli at the bottom of the screen. The participant's task was to choose which of the comparison stimuli was equivalent to, or "goes with", the sample. A participant's choices were guided by reinforcing feedback, delivered by the computer.

The two three-member equivalence relations were trained using the following tasks; A1-B1 (not B2), A1-C1 (not C2), A2-B2 (not B1) and A2-C2 (not C1). All stimuli, including A1 and A2, were presented in black size 48 font, in New York style. There were two baseline conditional discrimination training tasks, comprised of four matching-to-sample tasks. Each of the four tasks was presented four times in a block of sixteen tasks in a quasi-random order. This included no more than two consecutive exposures to any one task.

Prior to training, participants were seated in front of the *Macintosh* i-book lap-top computer. The computer first displayed a screen which presented the statement "Thank you for participating." in orange coloured, size 48, Chalkboard style font. This message was accompanied by a song by the artist Kylie Minogue, which would have been popular with the age group of the participants at the time of the experiment. The message and music faded after 9000 milliseconds. Immediately after that, a new computer screen appeared which presented the participants with brief instructions requesting them to use the computer mouse to click on the comparison stimulus they believed to be correct (see Appendix 5 for training instructions). These instructions were presented on the computer screen in orange coloured, size 24 font in Chalkboard style. Identical instructions, pre-recorded by the experimenter, in the experimenter's voice, were presented at the same time in audio format. Participants matched the comparison stimuli (e.g., Cug, Paf) to the sample (e.g., Ler, Vek) by clicking on the comparison of their choice using the computer mouse and cursor. Comparisons and samples were presented in black coloured, size 48, New York style font. All choices were followed by feedback delivered by the computer. Feedback appeared in red coloured, size 48, New York style font for 1500 milliseconds and then the screen went blank. This feedback informed participants as to whether their choice was correct or incorrect, by the word "correct" appearing on the screen or the word "incorrect" appearing on the screen. The word "correct" was accompanied by a high-pitched beep and the word "incorrect" was accompanied by a low-pitched beep.

This was an attempt to establish fluent equivalence responding with a novel set of stimuli during training.

Participants were required to reach a pre-set criterion of 100% correct responding across a block of 16 trials (i.e., four exposures to each of the four tasks) to finish training. The programme continued to deliver training blocks until criterion was reached (or up to a maximum of 1000 cycles). One participant (P5) was excused from completing the study due to not meeting criterion after four sessions.

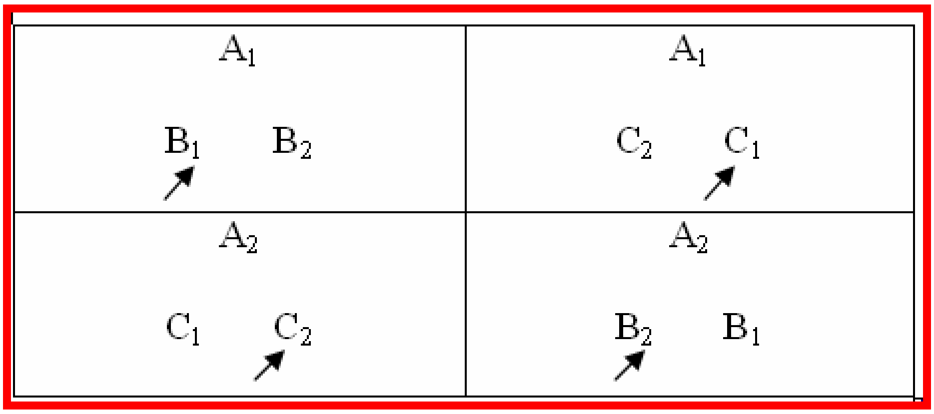


Figure 1. Conditional Discrimination training tasks. Training tasks employed during phase 2. Alphanumerics represent the nonsense syllables employed as stimuli. Arrows indicate the correct response in the presence of feedback.

Phase 3: Symmetry Testing

After reaching criterion during training, a message appeared on the computer screen. This message presented the statement, “Please feel free take a short break now. When you are ready to begin, please click the mouse”. The same popular song by the artist Kylie Minogue played softly in the background and ceased when the participant clicked the mouse. Instructions for testing can be seen in Appendix 6. Identical computer-generated audio instructions, in the experimenter’s voice were once again administered at the same time. Participants were then were exposed immediately to a block of 16 symmetry testing tasks. The symmetry testing tasks were B1-A1 (not A2), C1-A1 (not A2), B2-A2 (not A1), C2-A2 (not A1).

Instructions for testing can be seen in Appendix 6. Identical computer-generated audio instructions, in the experimenter's voice were once again administered at the same time. These four tasks were presented in a quasi-random order for 16 trials (i.e., four exposures to each of the four tasks) with no more than two successive exposures to any task. Participants were not provided with corrective feedback during this phase.

The symmetry testing block was cycled as many times as necessary to a maximum of 24 cycles, or until the participant responded with 100% accuracy on the same stimulus set. It was pre-determined that if the participant did not meet criterion after 24 cycles, they would be provided with computerised feedback, which told them that they were finished with this part of this experiment, and asked them to notify the experimenter. They were then thanked for their participation. No participant failed to meet criterion during this phase although time spent on this phase varied from participant to participant.

It was pre-determined that if a participant failed to meet criterion after 24 cycles that they would then have to be eliminated from the experiment. Participants who failed to meet criterion during a session simply continued with testing during the following session. Therefore, the total of 24 test cycles applied across sessions as well as within sessions (i.e., no participant was to be exposed to more than 24 cycles of symmetry testing, regardless of the number of sessions required to reach this limit).

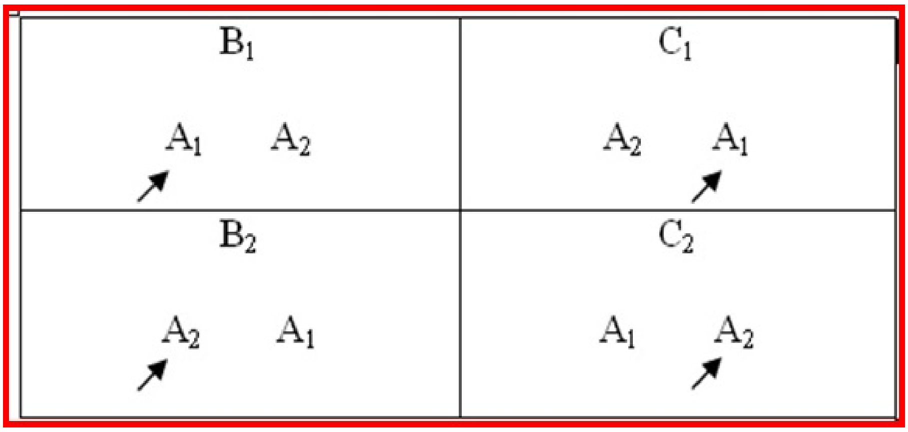


Figure 2. Symmetry Testing Tasks. Test tasks employed during phase 3. Alphanumerics represent the nonsense syllables employed as stimuli. Arrows indicate correct choices in the absence of feedback.

Phase 4: Conditional Discrimination Re-Training

The purpose of this phase was to re-familiarise participants with the baseline conditional discriminations following the symmetry test. This phase occurred before a test for transitive relations. The same procedures and criteria were employed as for Phase 2.

Phase 5: Transitivity Testing Phase

In the transitivity testing phase, there were four exposures to each of four tasks in a block. Once again, participants were required to reach a pre-set criterion of 100% correct responding across a block to finish the testing cycle. Participants were not provided with corrective feedback during this phase. As before, testing proceeded without a break, in blocks of 16 trials with these blocks cycled to a maximum of 24 cycles, or until the participant responded with 100% accuracy on the same stimulus set. If the participant did not meet criterion after 24 cycles, they were provided with computerised feedback, which told them that they were finished with this part of the experiment. The computerised feedback then asked them to notify the experimenter that they were finished and thanked them for their participation. No participant failed to meet criterion on this section although time spent on this phase varied from participant to participant. Also, identical to the symmetry testing phase, it had been previously determined that if a participant failed to meet criterion after 24 cycles in one session, then they would have to be eliminated from the experiment. Participants who failed to meet criterion during a session simply continued with testing during the following session. The total of 24 test cycles, therefore, applied across sessions as well as within sessions (i.e., no participant was exposed to more than 24 cycles of transitivity testing, regardless of the number of sessions required to reach this limit).

The transitivity testing tasks were B1-C1 (not C2), B2-C2 (not C1), C1-B1 (not B2) and C2-B2 (not B1), whereby the stimuli parentheses indicate incorrect choices. During testing, participants were expected to match B1 to C1, C1 to B1, B2 to C2, and C2 to B2, thereby demonstrating stimulus equivalence (see Barnes, 1994; Fields, Adams, Verhave, & Newman, 1990; Sidman, 1986).

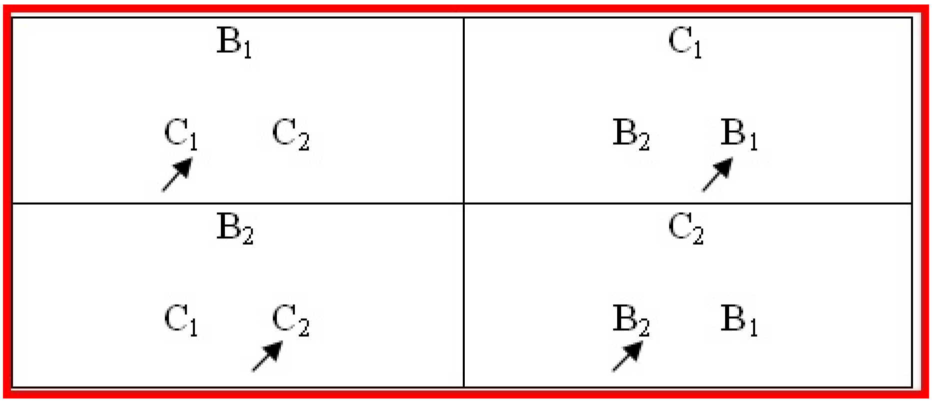


Figure 3: Transitivity Testing Tasks. Test tasks employed during phase 5. Alphanumerics represent the nonsense syllables employed as stimuli. Arrows indicate correct choices in the absence of feedback.

Results and Discussion

The data listed in this section consist of the total number of blocks required to pass the Phase 2 baseline conditional discrimination training, the total number of blocks to pass the Phase 3 symmetry test, the total number of blocks to pass the Phase 4 transitivity test, and finally, the total number of blocks to pass both the symmetry test and the transitivity test as a combined score. These various scores represent different gross measures of the fluency of the derived symmetrical and transitive responding for the participants.

One participant (P5) was unable to pass baseline conditional discrimination training after 4 one-hour sessions and will not be discussed. This participant was subsequently thanked for her participation and eliminated from the study.

Table 2. The number of training and testing blocks required by each participant to meet 100% correct criterion on baseline conditional discrimination training, re-training, symmetry testing, transitivity testing and symmetry and transitivity testing combined.

Participant Number	Session	BCD Training Blocks (A/B, A/C)	BCD Re-Training	Symmetry Testing Blocks (B/A, C/A)	BCD Re-Training	Transitivity Testing Blocks (C/B, B/C)	Symmetry/Transitivity Blocks Combined Scores (B/A, C/A, C/B, B/C)
1	1	42					
	2	23		14			
	3		2	1			
	4				26		
	5				1	2	17
2	1	32					
	2	20					
	3	3		2			
	4				1	12	
	5				2	4	
	6				1	13 (NP TB)	
	7				1	6	39
3	1	3		2			
	2		1			1	3
4	1	2		2			
	2		1			4	6
5	1	10					
	2	24					
	3	15					
	4	18		N/A			
6	1	10		2			
	2		1			2	4
7	1	13		4			
	2		1			3 (NP TB)	
	3		1			2	9
8	1	15					
	2	8					
	3	3		2			
	4		1			3	5
9	1	9		1			
	2		1			2	3
10	1	8		1			
	2		1			2	3
11	1	33					
	2	14		6			
	3		1			8	14
12	1	4		4			
	2		1			6	10

In Table 2, each individual session for each participant is presented. The number of blocks required by each participant to meet criterion on; the baseline conditional discrimination training tasks, re-training tasks (either because they started a new session or because they failed to meet criterion on a test), symmetry testing tasks, further re-training tasks, transitivity testing tasks, and combined symmetry and transitivity testing tasks are shown. The letters "NP TB" in column 7 indicates that the participant did not pass, but that the session was finished as the participant indicated that they wanted to take a break. The reader should also note that a participant could not progress to a testing phase without first meeting criterion on a training phase, even if this mean exposing a participant to training across multiple sessions.

It can be seen from Table 2 that there was little variation among the participants regarding number of blocks required to meet criterion on all measures. However, as a rule, symmetry and transitivity emerged quickly for most participants (i.e., many of them achieved perfect responding on their first attempt at a test). Some of the participants initially required a great deal of exposure to the baseline conditional discrimination training for A/B and A/C relations to meet training criterion. It is speculated that this may have been a result of the novelty of the context. However, all participants, except P5, met criterion on the baseline conditional discrimination training for A/B and A/C relations. Seven of the twelve participants met criterion for this training during their first session. All participants also met criterion for the symmetry testing task. One participant (P1) required exposure to 15 blocks to meet criterion. The lowest number of blocks required was one (P9, P10). However, most of the other participants required between two and four blocks to meet criterion. During transitivity testing (for C/B, B/C relations), one participant (P2) took 35 blocks to meet criterion. The lowest number of blocks required by any participant was one (P3). Most other participants (P1, P4, P6, P7, P8, P9, P10, P11 and P12) required fewer blocks to meet criterion (an average of 3.5 blocks) for this task. For the combined symmetry and transitivity testing measure, participants required a minimum of three blocks (P3, P9 and P10), and a maximum of

39 blocks (P2), with most participants requiring less than ten blocks altogether to meet criterion. The average number of blocks to reach criterion was ten blocks.

Participants also varied in their performances within the blocks of training and testing. The greatest variation was within the training task blocks. To illustrate these differences, participants have been divided into three categories based on their training performances (slow, medium and fast). “Slow” starters can be categorised as attaining less than half (eight) correct in a block of sixteen items by the fourth training block. These participants (P1 and P2) showed the steepest learning curves. The “medium” starters can be categorised as attaining scores of eight or more correct (out of 16) by the fourth block of training. The majority of participants (P6, P7, P8, P9, P10 and P11) fell into this category. These participants can be described as attaining better than chance level performances early in their training. The “fast” starters can be described as attaining scores of 13 or better (out of 16) by the fourth block of training. P3, P4 and P12 fell into this category. In fact, these participants all met criterion by the fourth block of training, if not before that time.

Less variance was seen for performance on symmetry and transitivity testing tasks. Using the same divisions used with training tasks, all participants, except P1, can be described as “fast” for symmetry testing. P1’s performance can be described as “medium”. Similarly, all participants, with the exception of P2, can be described as “fast” for transitivity testing. P2’s performance on transitivity testing can be described as “medium”. Thus, typically, participants attained better than chance levels on training tasks by the fourth block of training. Additionally, once participants were trained to criterion, they typically were quick to meet criterion on tests for symmetry and transitivity (e.g., within 13 blocks or less).

Correlational Analysis

A preliminary correlational analysis using scatterplots revealed that there appeared to be many correlations between subtest and full scale scores on the WISCIII^{UK} with equivalence training, symmetry testing, transitivity testing, and combined symmetry and transitivity testing scores. The preliminary analysis suggested that several correlations may be statistically significant. Thus, the number of blocks of training and testing required to meet criterion for both symmetry and transitivity were each statistically correlated with individual participant's scores on subtests of the WISC-III^{UK}. (For each individual's WISC-III^{UK} subtest scores see Appendix 7). Specifically, Spearman's Rho correlational analyses examined the relationships between the total number of training blocks (i.e., to Phase 2 training, not including mandatory re-training blocks) and IQ subtest scores. Further analyses calculated the correlation between; 1) the total number of symmetry testing blocks required to meet criterion for Phase 3 and IQ subtest scores; 2) the total number of blocks required for meet criterion for transitivity testing (Phase 4) and IQ subtest scores and; 3) the total number of both symmetry and transitivity testing blocks combined and IQ subtest scores. Table 3 indicates the correlation coefficients for each of these analyses for each participant in this experiment.

Table Number 3. Spearman's RHO Correlations for baseline conditional discrimination training, symmetry testing, transitivity testing and symmetry and transitivity testing combined with all IQ subtests.

IQ Measure/ Subtest	Transitivity Training (not incl. re-training) r values	P values	Symmetry Testing r values	P values	Transitivity Testing r values	P values	Symmetry and Transitivity Combined Testing r values	P values
Full Scale IQ	-.110*	.748	.189*	.577	.481**	.134	.168*	.622
Verbal IQ	-.419**	.200	-.480**	.135	-.017	.962	-.187*	.582
Performance IQ	-.011	.973	.026	.939	.727***	.011	.340**	.306
Picture Completion	-.068	.843	-.070	.838	.296*	.376	.191*	.574
Information	-.047	.891	-.266*	.430	.386**	.242	.147*	.667
Coding	.300**	.371	-.171*	.614	.276*	.412	.024	.944
Similarities	.085	.804	.073	.831	.202*	.552	.133*	.696
Picture Arrangement	.229*	.498	.135*	.692	.217*	.522	.103*	.764
Arithmetic	-.431**	.186	-.436**	.180	-.232*	.493	-.373**	.259
Block Design	-.012	.973	.162*	.634	.648***	.031	.426**	.192
Vocabulary	-.463**	.152	-.646***	.032	-.076	.823	-.359**	.279
Object Assembly	-.225*	.505	-.053	.876	.409**	.211	.098	.774
Comprehension	-.323**	.333	-.473**	.142	-.184*	.588	-.335**	.314
Symbol Search	.058	.866	-.698***	.017	-.081	.813	-.405**	.216
Digit Span	-.244*	.470	.031	.927	.027	.938	.078	.819

Note. The numbers in this table are raw "r" scores, or correlations. One asterisk indicates a mild correlation. Two asterisks indicate a moderate correlation and three asterisks indicate a strong correlation which is significant at the 0.05 level (2-tailed).

Cohen's (1988) system of interpreting values between 0 and 1 has been used here. In this interpretation r=.10 to .29 or r = -.10 to -.29 indicates a small or mild correlation. R= .30 to .49 or r= -.30 to -.49 indicates a medium or moderate correlation. R= .50 to 1.0 or -.50 to -.1.0 indicates a large correlation. (See Pallant, 2001 for detailed discussion of this system). Significant correlations are shown in red. In Table 3, it can be seen that there are small, medium and large (or mild, moderate and strong) correlations between most subtests of the WISC-III^{UK} and some element of the stimulus equivalence task. However, a point which has not yet been comprehensively addressed is the direction of these correlations. A negative correlation indicates that as the number of blocks to complete conditional discrimination training, symmetry testing, transitivity testing, or symmetry and transitivity combined testing decreases (indicating higher skill level at that task), IQ score increases. A positive correlation indicates that as the number of blocks to meet criterion on these tasks increases (indicating weaker performance on these tasks), IQ also increases. It is interesting that correlations are as expected for symmetry testing, particularly on verbal IQ score and Arithmetic subtests rather than tests for spatial relations or concrete manipulation of objects. It thus appears that symmetry is a good test of IQ subtest scores. However, for transitivity, correlations are weaker and often in the "wrong" direction. It is likely that poor range of scores across participants led to this outcome. Specifically, for symmetry, scores range from 1-15 blocks of testing before participants met criterion for that task, compared to transitivity where participants range from 1 to 35 blocks, where 35 may be considered an outlier as all other scores lie between 1 and 8. This single high score of 35, given such a low number of participants in the data set can be seen to strongly influence the outcome of the correlations. The inclusion of this outlier in the analyses may lead to correlations between IQ and transitivity being lost or reversed for some items. The range of scores for symmetry shows no extreme outliers and, as such, is a more normally distributed data set. Thus, the foregoing pattern appears to represent, for the most part, a floor effect in the data (and a ceiling effect in relational skill) with the exception of the outlier. The reader should note that while it would greatly improve upon the research to run more participants so that outliers don't have such a drastic effect on the data set, it was not feasible due to the intensive nature of the training and testing protocol and the large amount of time required to administer individual IQ tests. This issue will be revisited in the General Discussion.

According to the RFT analysis provided thus far, the correlations found would be expected to be negative rather than positive. This was not always the case in this analysis. The current findings indicate that, as expected, there is a significant (at the 0.05 level, 2 tailed) negative correlation between the Vocabulary subtest and symmetry testing. There was also a significant (at the 0.05 level, 2 tailed) negative correlation between the Symbol Search subtest and number of blocks to meet criterion on the symmetry test. This was unexpected as the Symbol Search subtest is categorised in the performance, rather than the verbal, section of the WISC-III^{UK}.

The strong and significant correlation found between the Vocabulary subtest scores and symmetry testing scores provides further support that there exists a functional relationship between verbal skills as measured by standardised IQ tests items and relational skills. This finding also supports the findings of O'Hora et al. (2005) and it is interesting that of all the correlations emerging here, this is one of the strongest. However, another significant correlation points to the surprising result that there are also strong and significant correlations between Derived Relational Responding (DRR) ability and a subtest in the performance (rather than verbal) domain on the WISC-III^{UK}. Thus, while many of the correlations seem to be in an expected direction (e.g., as IQ subtest score goes up, DRR measure goes down), there are also many that were not in an expected direction (e.g., as IQ subtest score went up, so too did the number of blocks required to pass the DRR measure). In fact, two of the strongly significant correlations (performance IQ with transitivity testing, Block Design with transitivity testing) seem to suggest that as the IQ measure goes up, so does the number of blocks required to pass a transitivity test. These curious findings point to the problem of conducting a large number of correlational analyses in a study employing relatively few participants. In contrast to the current study, the O'Hora et al. (2005) and the O'Toole et al. (in press) study employed a large number of participants. Therefore, the correlations observed in that study were less likely to be influenced by outlying data, ensuring that less surprising correlations emerged. In addition, the large participant pool used in those studies were made possible by a reduced IQ testing battery as well as the fact that participants in that study were adults who required little supervision compared to children. In the current study, in contrast, there was a balance of time and resources to be maintained by trading off numbers of participants against the demands of the study (i.e., which required many weeks of both the participants' and their parents' time). In effect, the lower number of subtests administered (chosen on the basis of pre-experimental expectations) in the O'Hora et al. (2005) study, and the lower number of subtests administered in the O'Toole et al. (in press) study reduced the likelihood of observing a range of unexpected correlations.

These issues notwithstanding, it may also be that SE tasks do not covary reliably with IQ tests and their subtest scores as a more complex task. In addition, it may be that at a relatively early stage of intellectual development the functional relationships between repertoires of relational skills and other intellectual skills have not yet been fully established. Put simply, we may see a wide variety of skills differences across children as well as differences in the trajectory of development of relational skills. That is, intellectual development is in a state of flux at this time of a child's life and subject to perhaps more variables than that of an adult whose intellectual development may be considered more "stabilised".

On balance, however, it should be pointed out that the strongest of the correlations were observed in the negative direction, as predicted, for the data set that was best distributed (i.e., symmetry testing). Thus, the strongest negative correlations, or for present purposes, predictors of high IQ, were observed for symmetry rather than transitivity. This may seem like a surprising outcome but in fact may also reflect the importance of simple symmetrical relations in language and vocabulary and consequently in the development of IQ. Thus, while a range of mild and moderate, positive and negative correlations were seen that were unexpected, the expected and most meaningful correlations from an RFT point of view tended to be the strongest. Nevertheless, the correlations in Experiment 1 are indeed difficult to interpret with the current data set. These issues will be revisited in later studies, and will also be addressed in the General Discussion in the current chapter.

Experiment 2

At the beginning of this chapter, it was stated that it is important to know if IQ and DRR actually correlate, but more importantly that a detailed analysis of the subtest scores on one or several IQ tests and their correlations with DRR was needed. Such an analysis has been conducted and does indeed shed some light on which aspects of IQ are most highly related to DRR skills. It has come to light that there are at least weak correlations between many IQ subtest items and DRR ability as measured by conditional discrimination training, symmetry testing, transitivity testing and symmetry and transitivity combined testing scores. It would seem that the

Vocabulary subtest and the Symbol Search subtest correlate most highly with DRR ability. Of course, as mentioned in the previous section, many of these correlations are difficult to interpret at this stage of the research, and suggestions for pursuing this issue empirically are needed and will be provided throughout the thesis (see Chapter 6).

While the foregoing data seem to suggest a functional relationship between IQ and relational skills (albeit one that requires significantly more empirical research), it has not been shown that IQ scores are dependent on relational skills. In other words, increasing relational skills may not necessarily increase measured IQ score. It may well be that both IQ and DRR are related via a third yet unknown variable. Indeed, even if these two variables are functionally related the correlation was always unlikely to be 100% for all subtests and DRR, as indeed was found in the current study. Thus, it seems likely that a myriad of variables are at work in determining both IQ and DRR, even though from a RFT perspective these two variables still bear an important functional relationship to each other. In effect, DRR is likely a necessary but not sufficient variable for determining IQ.

One obvious way in which to further explore the extent of this functional relationship is to conduct a brief intervention to examine whether or not an increase in DRR skill leads to an increase in IQ. While this may appear ambitious, it must be remembered that clinically relevant rises in IQ or subtest scores are not required in a preliminary laboratory investigation. Instead, what is needed is to establish that any numerical rises in subtest scores or full scale IQ scores are obtainable with a controlled DRR intervention. In effect, what is needed is an "in principle" effect of a DRR intervention. Such an exercise, if successful, would have two major research implications. The first implication would be that such an exercise might form the basis of an intervention to raise IQ or individual subtest scores on an IQ test. The second implication would be that such an exercise could have the potential to more fully establish the precise nature of the functional relationship between DRR and IQ. Given the foregoing, in Experiment 2, child participants from Experiment 1 were re-recruited to take part in a follow-up study. Eight of the twelve participants from Experiment 1 were available to take part.

Method

Participants

Eight of the participants from Experiment 1 agreed to participate in Experiment 2. Four were randomly allocated to the experimental group, or MET group, and four were randomly allocated to the control group, or non-MET group.

Apparatus

Materials used were identical to those used in Experiment 1.

General Experimental Sequence

Upon arrival at the experimental setting for the start of Experiment 2, each experimental participant was seated comfortably in front of a lap top computer. Each participant was given a verbal reminder of the generic introduction to the purpose of the study. All subsequent instructions were delivered via the *Psyscope* computer programme. The computer instructed each child how to perform each task, when to take breaks and when to keep going (see Appendix 5 and 6 for training and testing instructions. Pre-recorded audio instructions in the experimenter's voice, along with the popular song included in the previous experiment were included in an identical manner as in Experiment 1). The experimenter was always present when the participants were engaging in the tasks. However, as in Experiment 1, the experimenter only answered questions asked of her and did not give any instructions that had not already been addressed by the computer-based instructions. In fact, the experimenter merely reiterated the instructions already given.

The children in the control (non-MET) condition (P4, P6, P7, P9) did not receive the intervention. These participants were required to wait for a period of 3-4 months before being exposed to the follow-up conditional discrimination training and testing phases (employing novel stimuli, stimulus set 7). The number of conditional discrimination training, symmetry and transitivity test blocks required to reach criterion in the follow-up phase were compared to those required in Experiment 1 in order to calculate the relative effects of the intervention and its absence on participants' ability to derive symmetrical and transitive relations.

The multiple exemplar training (MET) intervention for the experimental participants proceeded in Experiment 2 only for the four experimental participants (P1, P2, P3 and P8). The purpose of the intervention was to increase the fluency of the DRR skills of the experimental participants by training participants to produce perfect (i.e., 100% accurate) DRR on a novel set of stimuli in the absence of feedback on their first exposure to a test phase. All experimental participants were exposed to training and testing and MET across a predetermined number of stimulus sets (e.g., five novel sets, in addition to Set 1 employed in Experiment 1 and Set 7 employed following the intervention) designed to produce this level of fluency. However, all experimental participants were exposed to all training and testing stages with these stimulus sets even if the fluency criterion had been reached before completing the final training and testing phase of the intervention employing Set number 6 prior to the follow-up whereby the final set (stimulus Set 7) was employed. The control group, however, were exposed only to the final stimulus set (i.e., stimulus Set 7) during a training and testing session that was repeated until a pass was produced (i.e., they were not required to pass on the first exposure). In effect, while there was a correctness criterion employed for the control participants, there was no fluency criterion (i.e., requirement to pass a test on the first exposure) enforced for this group.

Sessions in Experiment 2 consisted of eight phases in total. The experimental participants were exposed to all eight phases, while the control group were only exposed to the final four phases. Phase 1 included administration of a series of conditional discriminations using a novel stimulus set via a laptop computer to experimental participants only. As in Experiment 1, these conditional discriminations formed the conditional discrimination training procedure. Phase 2 included administration of a symmetry test. However, this symmetry test was not a test in the traditional sense of the word. In this symmetry test, corrective feedback was provided following every choice that each participant made. Thus, it was different from the symmetry tests in Experiment 1 where no feedback was provided. Phase 3 included the administration of more conditional discrimination training on the same stimulus set employed in Phases 1 and 2. This training served to re-familiarise participants with the conditional discriminations that they had learned during Phase 1. Phase 4 included the administration of a transitivity test. As in Phase 2, corrective feedback was provided following each choice made by each participant. Phase 5 included administration of a series of conditional discriminations utilising a further novel stimulus set. Both the experimental and control groups were exposed to this phase. These conditional discriminations formed the conditional discrimination training procedure for the second novel set. Phase 6 included administration of a symmetry test. Both the experimental and control groups were exposed to this phase. In Phase 6, no corrective feedback was provided following participants' choices.

Phase 7 included administration of more conditional discrimination training on the same novel stimulus set employed in Phase 6. This training served to re-familiarise participants with the conditional discriminations that they had learned during Phase 6. Finally, Phase 8 included the administration of a transitivity test. As in Phase 6, no corrective feedback was provided. For experimental participants, Phases 1-8 were cycled using each of the stimulus sets 2-6 in a sequential order. Participants were required to successfully complete Phases 1-8 using one set before being exposed to these phases with the subsequent stimulus set. All experimental participants were exposed the same number of stimulus sets even if they met criterion on an earlier stimulus set. All control participants were then exposed to Phases 5-8 with stimulus Set 7 (which differed from stimulus Set 1 employed during Experiment 1).

Phase 1: Conditional Discrimination Training

Each experimental participant was seated comfortably at a large table in a quiet room in their own home. The experimenter sat opposing the participant. No one else was present in the room, although in every instance, the participants' parent(s) were either in the next room or in another room in the house. As in Experiment 1, parents were also informed that they should feel free to come in and out of the room as they chose. As in the previous experiment, no parent did this. The participant was then asked if they were willing to participate in the research. If they responded, "yes" (e.g., that they were willing), then the experimenter proceeded to administer the intervention. Every participant agreed to continue. Each session took between one and two hours and each child was permitted to take as many breaks as they needed, although, as in Experiment 1, few ever took a break outside of an occasional trip to the bathroom.

During conditional discrimination training (as in all phases of Experiment 2), participants were seen individually. As in Experiment 1, a standard matching-to-sample procedure (e.g., Sidman 1971, 1986) was employed to train and test two three-member equivalence relations (i.e., A1-B1-C1 and A2-B2-C2), in a “linear” training protocol. This well-established methodology (see Barnes, 1994; see also Arntzen & Holth, 2000) involved the presentation of a sample stimulus in the centre of a computer screen, and the presentation of two further choice stimuli at the bottom of the screen. The participant’s task was to choose which of the comparison stimuli was equivalent to, or “goes with”, the sample. Arbitrary nonsense syllables were once again employed as stimuli. (All nonsense syllables employed across Experiments 1 and 2 can be seen in Appendix 4). Participant’s choices were guided by reinforcing feedback, delivered by the computer.

The two three-member equivalence relations were trained using the following tasks; A1-B1 (not B2), A1-C1 (not C2), A2-B2 (not B1) and A2-C2 (not C1). All stimuli, including A1 and A2, were presented in black font. There were two baseline discrimination training tasks, comprised of four matching-to-sample tasks (see Figure 4). Each of the four tasks was presented four times in a block of sixteen in a quasi-random order. This included no more than two consecutive exposures to any one task.

Prior to training, participants were presented with brief instructions requesting them to use the computer mouse to click on the comparison stimulus they believed to be correct (see Appendix 5 for training instructions. Pre-recorded audio training instructions in the experimenter’s voice were included in the same manner as in Experiment 1. The same popular song used in Experiment 1 was also played briefly prior to the audio and visual training instructions here). Participants matched the comparison stimuli (e.g., Cug, Paf) to the sample (e.g., Ler, Vek) by clicking on the comparison of their choice using the computer mouse and cursor. All choices were followed by feedback delivered by the computer. Feedback appeared in red 12-point font in the centre of the screen for 1500 milliseconds and then the screen was cleared (i.e., went blank). The feedback informed participants as to whether their choice was correct or incorrect, by the presentation of the word “correct” or “incorrect” appearing on the screen. The word “correct” was accompanied by a high-pitched beep, and the word “incorrect” was accompanied by a low-pitched beep. Participants were required to reach a pre-set criterion of 100% correct responding across a block of 16 trials (i.e., four exposures to each of the four tasks) in order to finish training. The programme continued to deliver training blocks until criterion was reached (or up to a maximum of 1000 cycles).

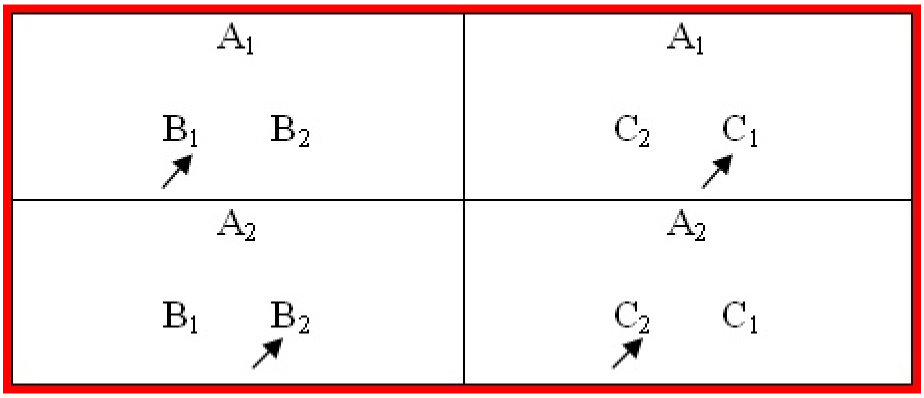


Figure 4. Baseline Conditional Discrimination Training Tasks. Training tasks employed during phase 1. Alphanumerics represent the nonsense syllables employed as stimuli. Arrows indicate correct response in the presence of feedback.

Identical to the proceedings in Experiment 1, follow-up appointments were always set up with each child’s parents for a date within one week of each session. Appointments were set for 1-2 hour sessions at the family’s convenience. As before, the second hour allowed time for set-up prior to commencement of each session and data entry, as well as packing up materials following each session. Appointments were not set at the end of a session if a child had already reached criterion on each and all of the training and testing cycles for each of the stimulus sets employed.

Phase 2: Symmetry Testing

After reaching criterion during training, experimental participants were offered a short break if needed and then were exposed immediately to a block of 16 symmetry testing tasks comprising for four exposures to each of four tasks (B1-A1, C1-A1, B2-A2, C2-A2, see Figure 5 below). Instructions for Phase 2 were identical to instructions for training. (Computer generated audio and visual instructions were employed as before). These four tasks were presented in a quasi-random order for 16 trials (i.e., four exposures to each of the four tasks) with no more than two successive exposures to any task. Each task was presented once in each block of four trials. Participants were provided with corrective feedback during this phase in an identical manner to Phase 1.

As in Experiment 1, the symmetry testing block was cycled as many times as necessary to a maximum of 24 cycles or until the participant responded with 100% accuracy on the same stimulus set. If the participant did not meet criterion after 24 cycles, they were provided with computerised feedback which told them that they were finished with this part of this experiment, asked them to notify the experimenter, and thanked them for their participation. No participant failed to meet criterion. Like before, it was pre-determined that if a participant failed to meet criterion after 24 cycles that they would have to be eliminated from the experiment. Participants who failed to meet criterion during a session simply continued with testing during the following session. The total of 24 test cycles, therefore, applied across sessions as well as within sessions (i.e., no participant was to be exposed to more than 24 cycles of symmetry testing, regardless of the number of sessions required to reach this limit).

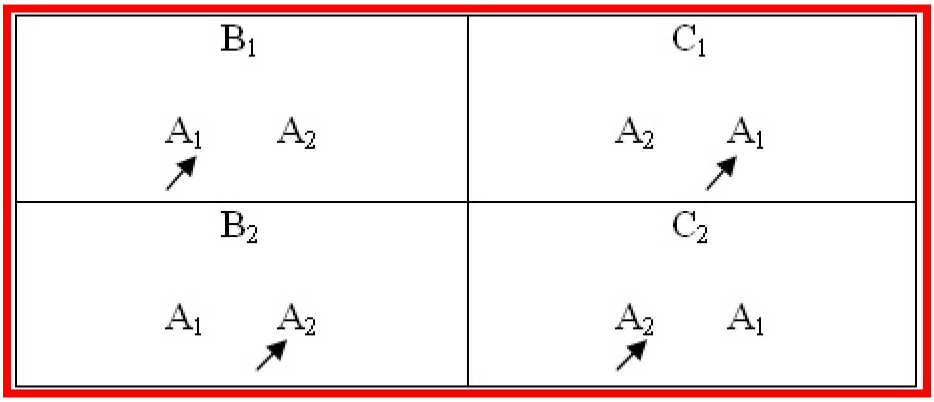


Figure 5. Symmetry Testing Tasks. Test tasks employed during Phase 2. Alphanumerics represent the nonsense syllables employed as stimuli. Arrows indicate correct choices in the absence of feedback.

Phase 3: Conditional Discrimination Re-Training

The purpose of this phase was to re-familiarise participants with the conditional discriminations following the symmetry test, but before a test for transitive relations. The same procedures and criteria were employed as for Phase 1.

Phase 4: Transitivity Testing Phase

In the transitivity testing phase, there were four tasks with each task presented four times in a block of 16. Each of the four tasks was presented once in a block of four trials, which was repeated four times (i.e., 16 trials in total). No one task was presented more than twice in succession. Once again, experimental participants were required to reach a pre-set criterion of 100% correct responding across a block to finish testing. Participants were provided with corrective feedback during this phase. As before, testing proceeded without a break, in blocks of 16 trials with these blocks cycled to a maximum of 24 cycles, or until the participant responded with 100% accuracy on the same stimulus set. As in Experiment 1, if the participant did not meet criterion after 24 cycles, they were provided with computerised feedback, which told them that they were finished with this part of the experiment, thanked them for their participation and also asked them to notify the experimenter. No participant failed to meet criterion on this section although time spent on this phase varied from participant to participant. Also, identical to the symmetry testing phase, it had been previously determined that if a participant failed to meet criterion after 24 cycles in one session, then they would have to be eliminated from the experiment. Participants who failed to meet criterion during a session simply continued with testing during the following session. The total of 24 test cycles, therefore, applied across sessions as well as within sessions (i.e., no participant was to be exposed to more than 24 cycles of transitivity testing, regardless of the number of sessions required to reach this limit).

The transitivity testing tasks were B1-C1 (not C2), B2-C2 (not C1), C1-B1 (not B2) and C2-B2 (not A1), whereby the stimuli in parentheses indicate incorrect choices (see Figure 6 below). During testing participants were expected to match B1 to C1, C1 to B1, B2 to C2, and C2 to AB, thereby demonstrating transitivity (see Barnes, 1994; Fields, Adams, Verhave, & Newman, 1990; Sidman, 1986).

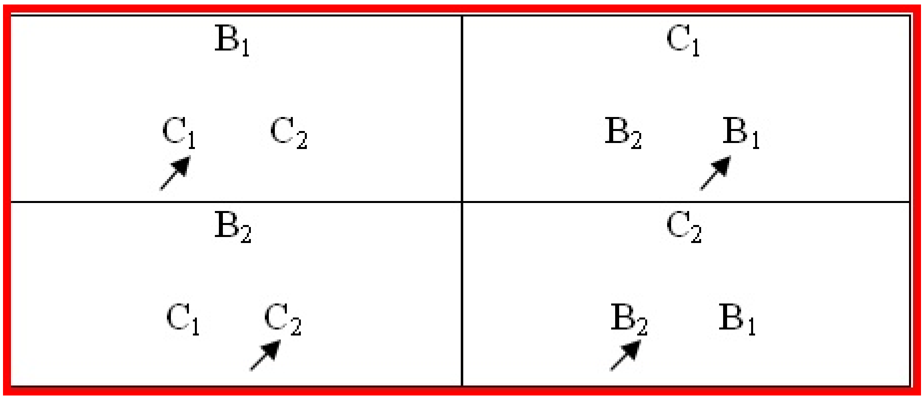


Figure 6: Transitivity Testing Tasks. Test tasks employed during Phase 4. Alphanumerics represent the nonsense syllables employed as stimuli during this phase. Arrows indicate correct choices in the absence of feedback.

Phase 5: Conditional Discrimination Training with Novel Stimulus Sets.

Phase 5 proceeded in the identical manner to Phase 1, but employed a unique novel stimulus set on every successive cycle through Phases 1-8 for experimental participants. This phase was administered to participants in both the experimental and the control groups. However, the experimental participants were cycled through this and subsequent phases five times (in addition to the baseline phases from Experiment 1), each time with a novel stimulus set.

Phase 6: Symmetry Testing (without feedback).

Phase 6 proceeded in an identical manner to Phase 2. However, participants were not presented with corrective feedback following responses. This phase was administered to participants in both the experimental and the control groups.

Phase 7: Conditional Discrimination Re-training.

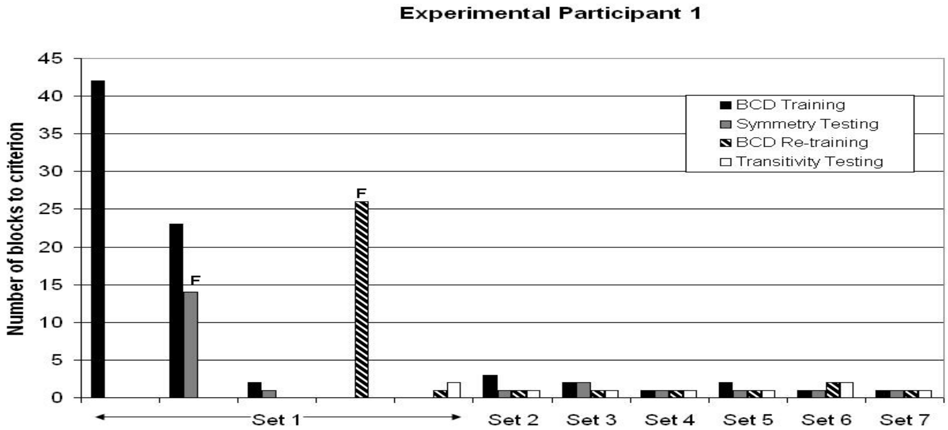
Phase 7 proceeded in an identical manner to Phase 3 and employed the same novel set as was employed in Phase 5 on each successive cycle. This phase was administered to participants in both the experimental and the control groups.

Phase 8: Transitivity Testing (without feedback).

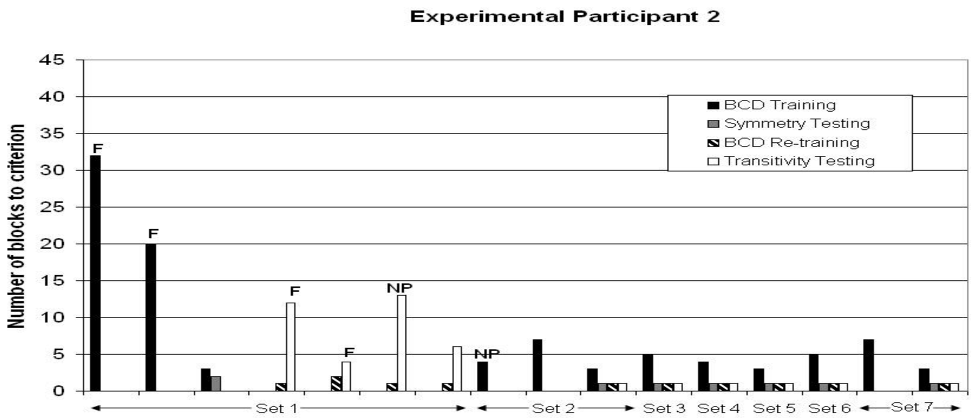
Phase 8 proceeded in an identical manner to Phase 4, with the exception that corrective feedback was not provided following participants' choices. This phase was administered to participants in both the experimental and the control groups.

Results and Discussion

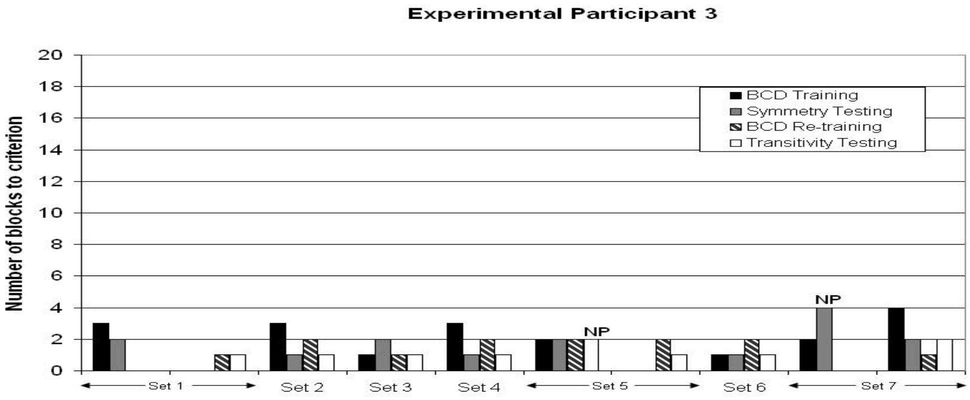
All IQ measures and baseline symmetry and transitivity scores were retained from Experiment 1 in order that they could be compared to the follow up IQ test scores, follow up symmetry and transitivity scores at the end of Experiment 2. The results for each individual participant, in terms of their rate of acquisition, from baseline through to follow up are listed in this section. The acquisition graphs are broken up into two figures. Figure 7 displays the experimental participants' results in terms of acquisition of criterion level responding for each stimulus set that participants were exposed to. Figure 8 lists the control participants' results in terms of acquisition of criterion level responding for each stimulus set that control participants were exposed to. Following the acquisition graphs, IQ changes will be discussed.



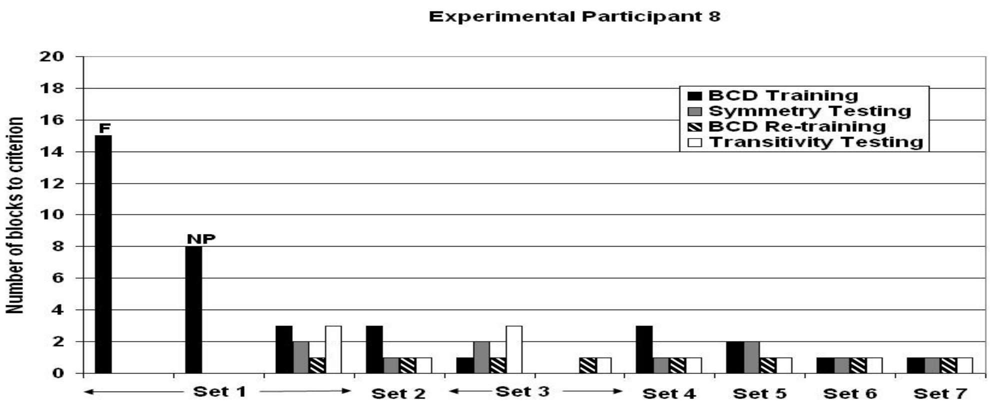
Panel 1.



Panel 2.



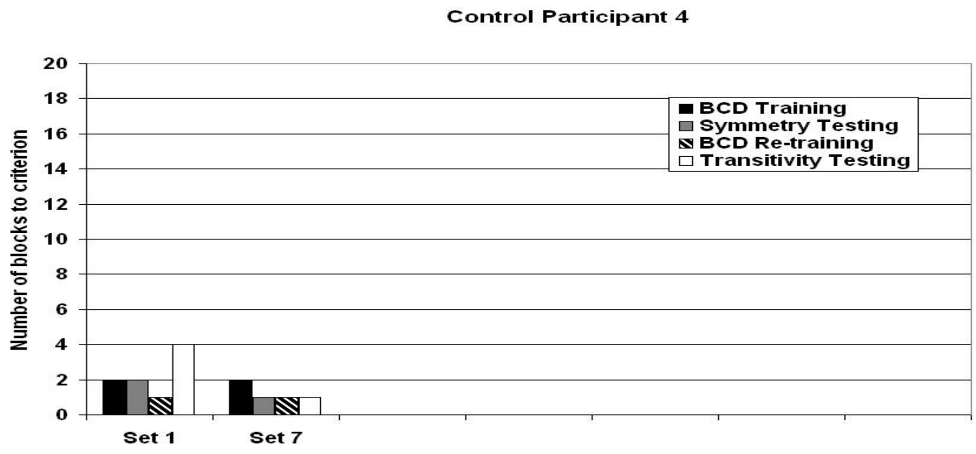
Panel 3.



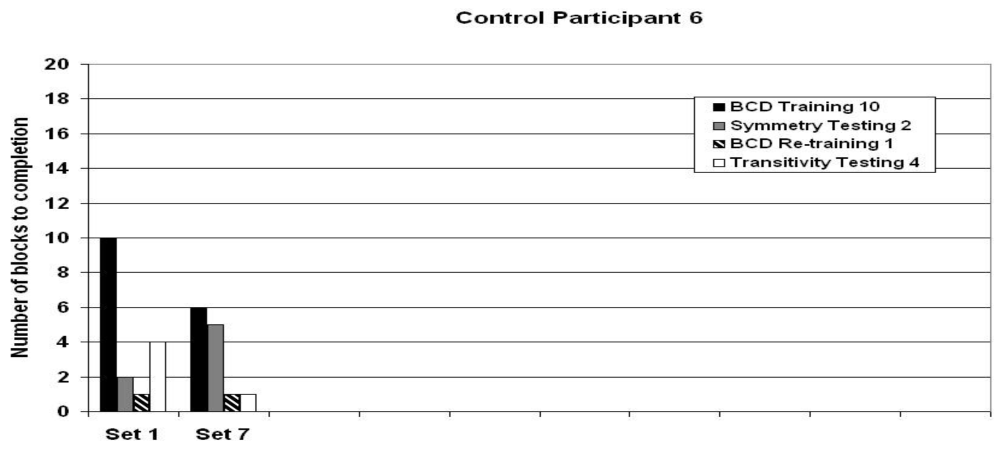
Panel 4.

Figure 7. Panels 1-4. Experimental (MET) participants' acquisition of baseline conditional discrimination training (BCD), symmetry testing, BCD re-training, and transitivity testing. The Y-axis shoes the number of blocks to criterion for each measure. The X-axis shows the individual novel stimulus sets. The legend key shows what each of the bars in the graph represents. "BCD" in black represents "baseline conditional discrimination training". Symmetry testing is shown in dark grey on the legend key. Baseline Conditional Discrimination (BCD) re-training is shown in black and white stripes. Transitivity testing is represented in solid white. An "F" inside these panels indicates that the participant failed to meet criterion. A "NP" indicates that the participant asked to take a break and therefore stopped the session and did not pass that set. Spaces between bars indicate a session break in real time.

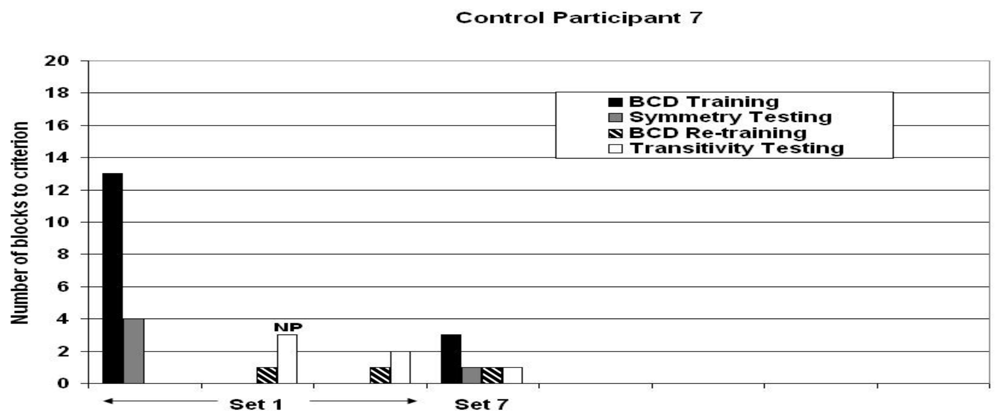
Figure 8 shows the acquisition graphs for control group (non-MET group).



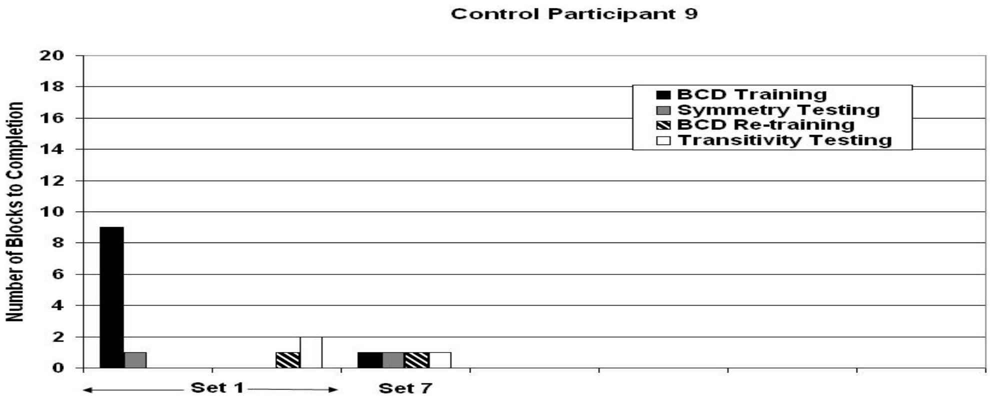
Panel 1.



Panel 2.



Panel 3.



Panel 4.

Figure 8. Panels 1-4. Control (non-MET) participants' acquisition of baseline conditional discrimination training (BCD), symmetry testing, BCD re-training, and transitivity testing. The Y-axis shoes the number of blocks to criterion for each measure. The X-axis shows the individual novel stimulus sets. The legend key shows what each of the bars in the graph represents. "BCD" in black represents "baseline conditional discrimination training". Symmetry testing is shown in dark grey on the legend key. Baseline Conditional Discrimination (BCD) re-training is shown in black and white stripes. Transitivity testing is represented in solid white. An "F" inside these panels indicates that the participant failed to meet criterion. A "NP" indicates that the participant asked to take a break and therefore stopped the session and did not pass that set. Spaces between bars indicate a session break in real time.

It would appear that, on the whole, participants' symmetrical and transitivity performance improved throughout the course of the experiment. However, both groups' data represents somewhat of a ceiling effect, in that participants produced fluent relational responding even at the baseline test (employing stimulus Set 1). More specifically, there was not much room for improvement in symmetrical or transitivity testing scores in either group as these skills were high for both groups even before an intervention was commenced with the experimental group. (See Table 4 for a summary of acquisition for both groups).

Table 4. Total number of test blocks to complete symmetry and transitivity tests at baseline and follow-up for Experimental (EXP) participants and Control (CT) participants.

Participant	Group	Symmetry	Transitivity	Follow-up Symmetry	Follow-up Transitivity
1	EXP	15	2	1	1
2	EXP	4	43	1	1
3	EXP	2	1	2	2
8	EXP	2	1	1	1
4	CT	2	1	1	1
6	CT	2	1	5	1
7	CT	4	5	1	1
9	CT	1	2	1	1

Table 4 shows this ceiling effect clearly as many participants required only one to two blocks of testing before meeting criterion at the baseline and thus, differences between the experimental and control groups' stimulus equivalence performance are difficult to interpret. At the level of acquisition alone, it looks like there is little difference in the relational skills acquired regardless of whether participants were members of the control (CT) group or the experimental (EXP) group. However, as noted in the previous section, the experimental participants did receive far more practice than did the control group as the control group were not exposed to any intervention. (This issue will be addressed in detail in the General Discussion). In fact, while the experimental group was engaged in the multiple exemplar relational intervention, the control participants were not engaged with at all. The control participants were revisited when the experimental participants had completed the intervention. Upon completion of the intervention, both the experimental group and the control group were revisited and administered follow-up conditional discrimination training on another novel stimulus set, a test for the emergence of symmetrical responding on this test, conditional discrimination re-training on this same novel set and finally a test for the emergence of transitive responding on this novel set. Modest differences in IQ subtest, full scale, performance and verbal scores did emerge between the experimental and control groups at the baseline and follow-up phases. These differences will now be discussed (see Table 5).

Table 5. All participants' baseline and follow-up full scale IQ scores (FSIQ), verbal IQ scores (VIQ) and performance IQ scores (PIQ).

Participant	Group	Baseline FSIQ	Follow-up FSIQ	Baseline VIQ	Follow-up VIQ	Baseline PIQ	Follow-up PIQ
1	EXP	98	105	103	102	92	107
2	EXP	119	118	120	120	115	112
3	EXP	109	111	113	109	103	112
8	EXP	96	107	101	100	91	115
4	CT	104	101	111	98	94	104
6	CT	111	111	111	111	107	107
7	CT	104	111	101	108	107	113
9	CT	107	106	110	113	103	96

It can be seen from Table 5 that the experimental group showed a mean rise of 4.75 full scale IQ points. The control group showed a mean rise of .75 full scale IQ points. The experimental group showed a mean fall of 1.5 verbal IQ points and the control group show a mean fall of .75 verbal IQ points. However, the experimental group showed a mean performance IQ rise of 11.25 points and the control group showed a mean performance IQ rise of 2.25 points.

The experimental and control participants' performances will be discussed in detail here. Experimental participant, P1 presented at the start of the intervention with a full scale IQ score of 98, which was comprised of a verbal IQ score of 103 and a performance IQ score of 92. Following the intervention, P1 presented with a full scale IQ score of 105, a verbal IQ score of 102 and a performance IQ score of 107. This displays a rise of seven full scale IQ score points, comprised of a drop of one verbal IQ score point and a rise of 15 performance IQ score points. Experimental participant P2's full scale IQ score showed a decrease of one point from baseline to follow-up (from 119 to 118). This full scale IQ score was comprised of a verbal IQ score which stayed the same (120 at baseline and follow-up testings) and a performance IQ score which dropped three points from baseline (115) to follow-up (112). Experimental participant P3 presented with a baseline full scale IQ of 109 and a follow-up full scale IQ of 111. P3 presented with a four-point drop from 113 to 109 for verbal IQ at baseline and follow-up. P3 presented with a nine-point rise in performance IQ from baseline (103) to follow-up (112). Experimental participant P8 presented with a full scale IQ rise of 11 points, from baseline (96) to follow-up (107). P8 presented with a one-point fall in verbal IQ from baseline (101) to follow-up (100).

Control participant P4 presented with a three-point drop in full scale IQ from baseline (104) to follow-up (101). P4 presented with a 13-point drop in verbal IQ score from baseline (111) to follow-up (98). P4 presented with a ten-point rise in performance IQ score from baseline (94) to follow-up (104). Control participant P6 presented with no change in full scale IQ score (111), verbal IQ (111) score or performance IQ (107) score from baseline to follow-up. Control participant P7 presented with a seven-point full scale IQ rise from baseline (104) to follow-up (111). P7's full scale IQ rise was comprised of a verbal IQ rise of seven points from baseline (101) to follow-up (108) and a performance IQ score rise of six points from baseline (107) to follow-up (113). Control participant P9 shows a full scale IQ fall of one point from baseline (107) to follow-up (106). P9 presented with a verbal IQ rise of three points from baseline (110) to follow-up (113) and a performance IQ fall of seven points from baseline (103) to follow-up (96).

With regard to the foregoing experiments, it is interesting to note that differences in relational fluency are not discernible easily from Figures 11 and 12. The reader should note that differences may still exist, but it is possible that these differences were not easily measured by graphs due to the ceiling effect described previously. The difference in exposure to the multiple exemplar training intervention for the experimental participants versus the absence of any intervention (control participants) is the essential difference between the two groups. It was predicted that the presence of an intervention would in turn lead to differences in IQ subtest measures. Table 5 shows that these predictions were correct in that the presence of an intervention seems to have had some measurable effect on overall IQ scores. (The individual subtest scores which contributed to these overall scores for each participant can be viewed in Appendix 9).

General Discussion

The findings from Experiment 1 suggest a possible functional relationship between IQ and relational skills (as measured by stimulus equivalence responding). However, it was noted that this relationship may not constitute proof of functional dependence. It was also suggested that it is possible that IQ and DRR may be related through a third and yet unknown variable. Indeed, even if these two variables are functionally related, it has been pointed out that the correlation was always unlikely to be 100% for all subtests of the WISC-III^{UK} and DRR. Thus, it seems likely that a myriad of variables are at work in determining both IQ and DRR, even though, from a RFT perspective these two variables still appear to bear an important functional relationship to each other. In effect, as predicted, it looks as though DRR is perhaps a necessary but not sufficient variable for determining IQ.

In Experiment 1, it was noted that a negative correlation indicated that as the number of blocks to complete conditional discrimination training, symmetry testing, transitivity testing, or symmetry and transitivity combined testing decreased (indicating higher skill level at that task), IQ score increased. A positive correlation indicated that as the number of blocks to meet criterion on these tasks increased (indicating weaker performance on these tasks), IQ also increased. It was interesting that correlations were as expected for symmetry testing, particularly on the verbal IQ scale and on Arithmetic subtests, rather than tests for spatial relations or concrete manipulation of objects. In addition, the current findings indicate that, as expected, there is a significant (at the 0.05 level, 2 tailed) negative correlation between the Vocabulary subtest and symmetry testing. It would appear therefore that symmetry is a good test of IQ subtest scores. The strong and significant correlation found between the Vocabulary subtest and symmetry testing provided further support that a functional relationship indeed exists between verbal skills as measured by standardised IQ tests items and relational skills. This finding also supported the findings of O’Hora et al. (2005) and of O’Toole et al. (in press). It should be reiterated at this point, that it is not surprising that of all the correlations which emerged in Experiment 1, that the one between Vocabulary and symmetry was one of the strongest. However, perhaps surprisingly, there was a significant (at the 0.05 level, 2 tailed) negative correlation between the Symbol Search subtest and number of blocks to meet criterion on the symmetry test. This was unexpected as the Symbol Search subtest is categorised in the performance, rather than the verbal, section of the WISC-III^{UK}.

For transitivity, correlations were often weak or in an unexpected direction. For instance, a strong and significant negative correlation was found between the

Symbol Search subtest (a performance subtest of the WISC-III^{UK}) and the number of blocks to produce transitive relations. Two of the strongest positive correlations found were between performance IQ and the number of blocks required to produce transitive relations and between the Block Design subtest and the number of blocks required to produce transitive relations. It is difficult, at this early stage of the research to explain these various outcomes. Indeed, any extended interpretive exercise may distract from the fact that ultimately these are empirical questions rather than conceptual ones. Nevertheless, some speculation on the possible sources of control over these effects is worthwhile.

It is likely that the poor range of scores across participants led to the emergence of many of these surprising correlations during the data analysis. Specifically, for symmetry, scores ranged from one to 15 blocks of testing before participants met criterion for that task, compared to transitivity in which participants' scores ranged from one to 35 blocks (where 35 was an extreme outlier and all other scores lie between one and eight). Of course, the single high score of 35 strongly influenced the outcome of the correlations, given the low number of participants in the data set. The range of scores for symmetry test performances showed no extreme outliers and thus, was a more normally distributed data set. Therefore, the foregoing pattern for transitivity skill levels appears to represent a floor effect in the data, with the exception of a single outlier. While it might be argued that a large participant pool would minimise the effects of outliers on the data set, it was not feasible to include more participants due to the intensive nature of the training and testing protocol as well as the large amount of time required to administer individual IQ tests.

It could be argued that the low total number of blocks required for children to meet criterion on the transitivity task were due to pre-existing differences in relational fluency. In other words, the participants' pre-experimentally high relational skills with regard to transitivity have made it difficult to observe differences across their performances. In effect, the acquisition of the skill for both groups was roughly similar, irrespective of whether individuals were members of the MET or non-MET group. Thus, comparing the two groups in terms of relative rates of acquisition of transitivity yielded little differences.

The unexpected correlations observed in Experiment 1 may not have emerged in a study employing a larger number of participants. Indeed, in the O'Hora et al. (2005) and the O'Toole et al. (in press) studies, the number of participants employed were large. However, in those cases, the participants were adults that required little supervision compared to children. Furthermore, the large participant pool was made possible by reducing the IQ testing battery. In the current study, in contrast, there was a balance of time and resources to be maintained by trading off numbers of participants against the demands of the study (i.e., which required many weeks of both the participants' and their parents' time). It is possible that O'Hora et al. (2005) would have found the same types of functional relationships (and the same amount of variance) if researchers involved in that research had administered all subtests of the WAIS. Similar findings may have emerged in the work of O'Toole at al. (in press) if a more comprehensive IQ test had been employed. However, it should also be noted that it is possible that the variance evident in the current research may not have been apparent with the large participant pool employed in the O'Hora et al. or the O'Toole et al. research.

It is interesting that the number of blocks required to produce symmetry responding shows a much stronger correlation with IQ measures than transitivity does because, at face value, symmetry appears to represent an easier task than transitivity (i.e., it is more basic). Instead, the data showed that typically participants needed larger numbers of blocks to meet criterion on symmetry rather than on transitivity tests. It is likely that this result points directly to the novelty of the symmetry testing context as the source of poor stimulus control over symmetrical responding. In addition, perhaps participants' performances improved as the context became more familiar and this set the occasion for improved abilities at derived relational responding in general, including transitivity skills. In other words, once a participant met criterion on a symmetry test, the participant may have been more likely to produce other familiar types of relational performances in the same way as during school level mathematics, one established skill can lay the foundation for another. The way in which this may have occurred is, of course, purely speculative at this point. However, it is reasonable to suggest that familiarity with the training and testing contingencies may have aided in the acquisition of the transitive relations, thereby normalising acquisition rates across participants.

The foregoing paradoxical outcome may be unexpected given a strictly mathematical definition of equivalence as suggested by Sidman (1971, 1986). In Sidman's definition, it is suggested that the abilities to produce reflexivity, symmetry and transitivity occur together. However, Pilgrim and Galizio (1995, 2000) and Smeets, Y. Barnes-Holmes, Apkinar and D. Barnes-Holmes (2003) have all shown the independent demonstration of symmetry and transitivity. Pilgrim and Galizio (1995) found that equivalence-class performances may be more easily disrupted in young children than in adults following reversals of the trained reflexivity, symmetry, and transitivity/equivalence probes. While it has been shown (Pilgrim & Galizio, 2000) that there is some variability across methodologies in the disruption of equivalence class performances, an abundance of data supporting the phenomena exists. For example, across two experiments, Smeets et al. (2003) demonstrated that regardless of training protocol, almost all members of a participant pool of 66 participants were able to reverse equivalence relations. In other words, these pieces of research provided evidence that reflexivity, symmetry and transitivity may not be functionally dependent and may not be best considered as occurring together as a single functional unit called stimulus equivalence. In effect, it may well be possible that there are lines of fracture between symmetry and transitivity that manifest themselves in the independent acquisition of each skill.

The foregoing issues notwithstanding, it has also been pointed out that it is possible that SE tasks simply do not co-vary reliably with IQ tests and their subtest scores as a more complex task. In addition, it may be that at a relatively early stage of intellectual development the functional relationships between repertoires of relational skills and other intellectual skills have not been fully established. Thus, there may exist many differences in the skills of young children as they continue in their intellectual development. Intellectual development in adults may not present the same levels of variability. On balance, however, it should be noted that the strongest of the correlations were observed in a negative direction as predicted for the data set that was best distributed (i.e., symmetry testing). Thus, while a range of mild and moderate positive and negative correlations were seen that were unexpected, the expected and meaningful correlations from an RFT point of view tended to be the strongest.

The fact that the strongest negative correlations, or for present purposes, best predictors of high IQ, were observed for symmetry rather than transitivity has been highlighted as a surprising outcome. However, it may in fact, also reflect the importance of simple symmetrical relations in language and vocabulary and consequently in the development of IQ. More specifically, vocabulary consists largely of large numbers of bi-directional relations between words and objects and between different words with the same meaning (e.g., synonyms). Thus, it may be said that vocabulary in the early stages only or mainly involves symmetry. Children learn names for a large range of objects. For those names to function symbolically, a mutually entailed relation between object and word must also be derived. So at least in early years the ability to derive mutually entailed symmetry relations, which is crucial to a good vocabulary, may well be a good predictor of verbal ability later on. According to Barnes-Holmes, Y., Barnes-Holmes, Y. Roche, Healy, Lyddy, Cullinan and Hayes (2001, p. 160-161), it is no surprise that vocabulary would emerge as a primary factor in verbal intelligence. "Relational frames originally emerge in this context, and acquiring elaborated networks of verbal content develops and applies every relational frame in common use. Thus, persons with a highly elaborated vocabulary will tend also to have highly elaborated relational repertoires. It is the relational skills that are key, not merely verbal content in a formal sense. A task such as learning to spell is far less relationally rich than learning word meanings, and thus it is no surprise that spelling performance will correlate less with overall levels of intellectual behavior than will vocabulary, even though both tasks involve verbal material". Given the foregoing conceptual analysis, it may not be surprising that IQ subtest items, such as vocabulary, turn out to correlate well with level of proficiency at symmetrical responding.

Experiment 2 further examined the nature of the functional relationship between IQ and DRR successfully. The findings from Experiment 2 provide evidence that DRR is indeed an operant phenomenon and that modest rises in IQ scores were generated using the DRR intervention. All four experimental participants reached the pre-determined fluency criterion given the DRR intervention, and this skill was shown to generalise to novel stimuli. However, relatively weak differences across experimental groups were observed due to rapid acquisition of relational skills by all participants. It is possible that for the control participants, exposure to the training and testing protocol served as a form of practice that may have been somewhat effective in facilitating the acquisition of relational skills in the absence of a MET intervention. While not immediately obvious, this in itself could provide support for the idea that relational skills can be acquired through operant conditioning. More specifically, by the strict definition of MET, practice in the absence of corrective feedback does not qualify as a relational intervention. However, by definition, if exposure to training and testing with multiple stimulus sets leads (as it did) to an improvement in relational fluency, this demonstrates that relational skills must, in principle, be subject to practice effects. While we have no technical definition of what constitutes practice, for learning to take place across training and testing sessions, some form of reinforcement must be occurring on a trial-by-trial and a test-by-test basis (i.e., operant conditioning). In the absence of any obvious source of reinforcement being delivered by the computer or experimenter, it is likely that some form of conditioned reinforcer for relational consistency must be operating during the test tasks. Interestingly, RFT, places a heavy emphasis on the role of conditioned reinforcers in precisely this situation. RFT would suggest that perhaps verbal consistency could have acted here to establish relational skills for the control participants in the absence of explicit reinforcement. This idea is worth considering briefly in more detail.

Relational Frame Theory predicts that identifying correspondences in a relational network (i.e., verbal coherence) will function as a relatively powerful reinforcer for relational activity itself, and that this will be an important feature of the verbal behaviour of most individuals. Indeed, several studies have shown that say-do correspondences can be reinforced (Baer, Deitrich, & Weninger, 1988; Catania, Shimoff, & Mathews, 1987; Lloyd, 1980; Paniagua, 1985; Ribero, 1989; Riegler & Baer, 1989 in Roche, Y. Barnes-Holmes, D. Barnes-Holmes, Stewart & O'Hora, 2002). Moreover, children are taught to be consistent in what they say and do from the time they can control their social environments through verbal interaction. There can even be punishments carried out against individuals for being inconsistent in their behaviour (see Guerin, 1994). Other researchers have examined the social processes involved in the maintenance of verbal consistency across verbal episodes (see Schauss, Chase, & Hawkins, 1997). Given the foregoing, it should not be surprising if verbal consistency, and behaviour-behaviour consistency more generally, were to become a conditioned reinforcer for verbal behaviour itself.

According to Roche et al. (2002), while behavioural consistency may be trained in even the simplest of organisms, it is thought that a special variety of consistency is available to the verbal organism. More specifically, given a history of DRR, a verbal organism may respond to consistencies across behavioural or verbal episodes that are topographically different. For example, a speaker may respond to two different statements made in different languages as having the same meaning. Furthermore, they may respond to the verbal coherence that obtains across the base and target domains specified in analogies (e.g., "hand is to glove as foot is to shoe") and metaphors (e.g., "cats are dictators") in a way that would seem impossible for a nonverbal organism (see Stewart, Barnes-Holmes, Roche, & Smeets, 2001). From the RFT perspective therefore, once an extensive network of relational frames is established, and a history of reinforcement is provided for producing coherent relational networks (i.e., not contradicting oneself), coherence will serve as a continuously available reinforcer for derived relational responding. That is, language will become a process that is self-sustaining because the very product of relational responding (i.e., coherent relational networks) is itself a conditioned reinforcer for further relational activity.

In the context of Experiments 1 and 2 of the current research, it is entirely possible that consistency in responses across trials was actually reinforced in the laboratory context. In other words, a social history in which consistency was established as a reinforcer may not have been necessary. That is, the cessation of training and testing phases when participants produced consistent response patterns may have served as a reinforcer for appropriate levels of relational consistency across the numerous training and testing cycles. In effect, consistency was established as a conditioned reinforcer within the research programme itself. In addition, inconsistency in relational responding was punished by the delivery of repeated training/testing phases. Thus, using the concept of consistency as a conditioned reinforcer, RFT can help to understand how exposure to the training and testing format alone may have some of the effects of explicit MET for the non-MET (i.e., control) group. This is in support of some recent findings by Vitale, Y. Barnes-Holmes, D. Barnes-Holmes and Campbell (2008). Vitale et al. (2008) found that exposure to unreinforced testing in an intervention format yielded some improvements in DRR repertoires for More-Than/Less-Than relational responding in normally functioning adults. Vitale et al. (2008) noted that reinforcement, or feedback, also seemed to be a crucial factor in the work of Y. Barnes-Holmes, D. Barnes-Holmes, and Roche (2001), Y. Barnes-Holmes, D. Barnes-Holmes, Roche, and Smeets (2001), Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, Strand, and Friman (2004), and Y. Barnes-Holmes, D. Barnes-Holmes, and Smeets (2004) even though practice at the task (i.e., exposure to the training and unreinforced testing format) did have some effect.

Of course it could be also be argued that the foregoing constitutes a fatal design flaw in Experiment 2 insofar as groups were not only administered different training and testing protocols (i.e., presence or absence of feedback during testing), but groups also differed in terms of the overall number of blocks of training and testing delivered. This difference could be the source of differences in relational skills at the follow-up test and the modest rises in IQ observed following the intervention. It is important to remember however, that Experiment 2 was exploratory in nature and was intended only to establish an in-principle effect of MET on DRR fluency and IQ scores. It was intended that increased training and testing block numbers across a large number of stimulus sets over that received by a control group would itself constitute part of the intervention. Of course, it becomes difficult when such a procedure is used to determine post hoc the relative importance of MET over any possible effects of increased exposure to training and testing per se in determining IQ rises. These issues will be addressed in further chapters which will more carefully control the number of training and testing blocks to which experimental and control participants are exposed.

Perhaps the most exciting finding of this study relates to the modest IQ rises that were observed for most participants in the experimental group. While the sample size employed was undeniably modest, it is important to remember that such onerous research places considerable limits on the amount of training and testing that can be administered in an individual experiment by an individual researcher. Nevertheless, even with such a small sample size, the rises in IQ appear to suggest a clear pattern of influence by the DRR intervention. Regardless of the precise process involved in these rises (e.g., practice alone versus the effects of explicit MET), these recorded rises in what is otherwise considered a psychological invariant should be of great interest to educational psychologists, psychometricians, and specialist educators working in a wide variety of domains. Any doubt over the nature of the relevant psychological process can be further investigated in future research both within the current thesis and beyond.

Interestingly, the IQ rises observed in Experiment 2 would be of particular interest to those working in applied settings insofar as these rises in several cases exceeded the rises typically expected across time under normal conditions, even given repeated exposure to the IQ test. More specifically, the stability of scores on the WISCIII^{UK} has been assessed in a study of 353 children who were tested twice (Wechsler, 1992). In the standardisation studies, the intervals between testings ranged from 12 to 63 days with a median re-test interval of 23 days. These data showed an increase of approximately 7-8 points in the FSIQ score over a short re-test period. Discrepancies in score, due to practice effects, are smaller for the verbal IQ (VIQ) than for the performance IQ (PIQ) score. However, practice effects on the WISC-R are smaller over longer test-re-test intervals (see WISCIII^{UK} manual; see also Juliano, Haddad, & Carroll, 1988, for further discussion). Thus, the rises found here seem to be larger than the standard rises associated with practice effects. In addition, it should be noted that practice effects on IQ scores are typically short-lived when measured in test-retest studies (i.e., a median time of 23 days in the WISCIII[UK] standardisation research). However, the length of time between baseline and follow-up administrations of the IQ test in the current study was between 90 and 120 days. In effect, the rises in IQ observed in the current study are unlikely to be accounted for by test-retest effects, even by those who have developed the Wechsler scales.

It is exciting to note that one IQ rise from baseline to follow-up was so large that it could potentially move a child from one qualitative IQ range to another (e.g., from sub-normal to normal). More specifically, P8 in the experimental condition presented with an eleven-point rise in full scale IQ score from baseline to follow-up. Although, this particular participant's scores were both within normal range, the implications of such a rise with a learning disabled child are obvious. Such an IQ score increase could move a child from presenting within the Borderline Learning Disabled range to presenting within the normal range. Alternatively, such a rise could move a child from the Mild General Learning Disability range to the Borderline range. These types of classifications may seem superfluous to the casual eye. However, they would have great impact on the types of services available to individuals and on the way these individuals might be treated by peers, teachers, other professionals and even their own parents.

It could be argued at this point that participants' IQ rises might be due to participant expectancies. However, the reader is reminded that MET group participants were given no more details on the purpose of the study than the non-MET group were given. No participant or parent/guardian of that participant was privy to information on which group participants' were a member. In addition, participants had no access to software outside of normal session hours. Thus, the participants were completely unable to practice the prescribed tasks. Parents and guardians were also naïve to the nature of the training and testing protocols, so there was also no opportunity for parents/guardians to assist participants in between sessions. Also, the intervention software was not left at participants' houses and thus, participants had no opportunity to practice without the experimenter's knowledge. In addition, IQ was roughly normally distributed across groups (bearing in mind the small sample size) before the beginning of the intervention. That is, at the beginning of the research study all except one participant (P2) were in the average IQ range (90-109). The remaining participant (P2) presented in the high average range (110-119).

Given the findings of Experiment 2, one aspiration for the current research is to design practical interventions that can be employed by teachers and caregivers to give intellectually challenged children a better opportunity to reach their intellectual potential. Therefore, the true value of this research will be found in the educational programmes that might be established to produce changes in the intellectual abilities of children. At present, such efforts focus almost exclusively on learning styles and knowledge content rather than on established psychological processes, such as derived relational responding. Thus, armed with RFT as a conceptual framework and a touchstone for the development of practical interventions, we are poised to make a real and measurable contribution to the fields of intellectual disability, educational psychology and behaviour analysis more generally. For example, such contributions might focus on increasing abilities to generalise mathematical principles rather than the ability to memorise addition or subtraction tables.

An intervention of the suggested kind would need to circumvent the problem of rapid acquisition of relational skills for control participants by involving more complex relations. For example, the inclusion of training and testing protocols for increasing relation skills involving the relations of Same, Opposite, More-Than, Less-

Than, Before/After, et cetera might be appropriate. Another approach to circumventing the problem of rapid acquisition of relation skills might be to use a multi-component task such as the protocols employed in O'Hora et al. (2005). Another requirement might include increasing the level of fluency required of participants. The foregoing could be facilitated by requiring participants to meet criterion on multiple tasks or by increasing the number of dependent measures to include task completion time. Such strategies would also serve to differentiate participants of different relational abilities. In other words, participants might be required to achieve 100% and then required to get 100% 80% of the time. All of these suggested measures might serve to circumvent the difficulties with interpretation of data when there are floor effects present such as the effects witnessed in this chapter.

The development of a rich and intensive intervention including the features suggested here might yield valuable information for the development of a relational intervention to raise IQ. Building on the foregoing suggestions, future work in this thesis will employ a more intensive intervention that draws on multiple approaches to increasing relational skills as efficiently as possible. The first step, however, is to focus on the intervention format before returning to the issue of IQ rises and their functional relationship to DRR.

Chapter 3

An Examination of Multiple Exemplar Training as a Means of Improving Same and Opposite Relational Responding in Children and Adults

Experiment 3

Introduction

In Chapter 2, it became evident that IQ rises as a result of the MET intervention may be difficult to demonstrate conclusively. Conclusions about the IQ rises observed in Chapter 2 can only be tentative for several reasons. Firstly, there was a small number of participants in the subject pool employed. As a result of the low number of subjects, any variations in measured IQ, in either direction, cannot form the basis of generalisations. A larger participant sample would be needed to make firmer conclusions regarding the observed patterns. This is undoubtedly a limitation of this research but is a natural feature of studies with small sample sizes often conducted in the context of such labour intensive investigations. This matter will be returned to in later chapters. Secondly, the ceiling effect in DRR skills observed in Experiment 2 of Chapter 2 led to the observation of similar rates of acquisition of both symmetrical and transitive relations across the experimental and control groups. Thus, the effect of multiple exemplar training (MET) on IQ was obscured. It was suggested in the previous chapter that the surprisingly similar rates of acquisition of DRR across groups may have been the result of the unexpectedly powerful effects of practice provided to the control group during repeated testing. This may partly explain the convergence in DRR skills across time for both groups.

A third reason that conclusions regarding IQ rises and the effects of MET are tentative relates to an important methodological shortcoming. That is, as pointed out in Chapter 2, the experimental and control groups were not only administered different training and testing protocols but these two groups also differed in terms of the overall number of blocks of training and testing delivered. This confound in the implementation of the independent variable could have influenced the observed results. As noted previously however, Experiment 2 of Chapter 2 was exploratory in nature and was intended only to establish an in-principle effect of MET on measured IQ scores. It was intended, therefore, that increased practice across a large number of stimulus sets over time for a control group would itself constitute the intervention. Nevertheless, these procedural issues must be addressed and will be in later sections of this thesis (see Chapter 4). What is important at this stage is to focus on the matter of developing interventions that work clearly to improve DRR. These are in-principle procedures that can be used to raise relational skills in children reliably and quickly.

Given the foregoing, the experiment in the current chapter, and those subsequent, are an attempt to establish more reliable procedures for establishing increases in DRR skill with improved experimental designs. The experiment in the current chapter will examine the acquisition of the relational responding in accordance with Same and Opposite relations. The tasks in this experiment have been designed to circumvent the problem of overly rapid acquisition (i.e., a ceiling effect in relational skills across experimental and control participants) by introducing more relationally complex tasks. In addition, the following experiment employed an alternative training procedure to MTS in an attempt to allow a very large amount of relational training to take place in as short a period of time as possible. In other words, if the purpose of the procedures being developed here is to establish the basis of an intervention for intellectual deficit, it is important that the procedure is as easy to establish with participants as possible (especially given the increased complexity of relational frame training tasks over equivalence tasks). A new procedure, which is an innovative combination of the Relational Evaluation Procedure (see Cullinan, Barnes-Holmes, & Smeets, 2001) and the

Yes-No (see Fields, Adams, Verhave, & Newman, 1990) procedure, was used here. This chapter will examine this procedure with children and adult participants. Five normally developing children and six normally developing adults took part in Experiment 3.

The experiments in Chapter 2 employed the well-established MTS procedure. In contrast, the current chapter will employ a less common training protocol. Thus, it is worth briefly exploring the purpose of these procedural changes. The procedure used in Experiment 3 employed a novel combination of other established procedures. While this precise procedure has not been used before in precisely this format, the idea of expanding upon old procedures and exploring new training and testing procedures has been suggested by other researchers (D. Barnes-Holmes, Y. Barnes-Holmes, Smeets, Cullinan, & Leader, 2004). D. Barnes-Holmes, Y. Barnes-Holmes, Smeets et al. (2004) suggested that the development of a wide range of new procedures may considerably enhance the empirical and theoretical analysis of stimulus equivalence (SE) and deriving relations more generally. They also noted that there are few published studies that have attempted to break free of Matching to Sample procedures (e.g., Cullinan, Barnes, & Smeets, 1998; Cullinan, Barnes-Holmes, & Smeets, 2000, 2001; Fields, Reeve, Varelas, Rosen, & Belanich, 1997) and that there is much to be learned about the phenomena of stimulus equivalence and derived stimulus relations by making such a break. For instance, some attempts have been made to extend the range of methods used to study SE by using variations of the standard MTS procedure (see Markham & Dougher, 1993; Schenk, 1995; Smeets, Schenk, & Barnes, 1994; Smeets & Striefel, 1994; Stromer, McIlvane, & Serna, 1993; see also Dube, Green, & Serna, 1993). One procedure, in particular, that has been developed in the context of RFT is the Relational Evaluation Procedure (REP). This procedure is similar to but extends upon earlier procedures, such as go/no go procedures (e.g., D'Amato & Colombo, 1985), go left/go right or yes/no (see D'Amato & Worsham, 1974) and same/different (see Edwards, Jagielo, & Zentall, 1982; see also Fields, Reeve, Varelas, Rosen, & Belanich, 1997, for a review of a variety of psychological experiments employing such procedures). In an attempt to further explore procedures not employing MTS, D. Barnes-Holmes, Y. Barnes-Holmes, Smeets et al. (2004) built on these procedures and created the REP. The earliest incarnation of this procedure was called the precursor to the REP (see Cullinan, Barnes, & Smeets, 1998). In the precursor to the REP, a participant is not asked to choose a comparison that "goes with" a sample. Rather, a participant is asked to choose a response option that evaluates the positive or negative relation that obtains between the comparison and sample stimuli on a given trial. Over time, this methodology has been further developed and has been labelled the Relational Evaluation Procedure or REP (Barnes-Holmes, Healy, & Hayes, 2000; Hayes & Barnes, 1997; O'Hora, Barnes-Holmes, Roche, & Smeets, 2004; Stewart, Barnes-Holmes, & Roche, 2004). Importantly, this procedure was explicitly designed to allow for the rapid training of large numbers of complex relational networks in a short space of time. This consideration is highly relevant to the current research.

What is required for the current research is a training format that allows for the rapid acquisition of derived relations using a task presentation format that itself requires no additional skills to be established in the laboratory. More specifically, a training and testing format is needed that does not require the establishment of control by sample and comparison stimuli. Also, the elimination of positional responding through the counterbalancing of comparison positions may increase the speed of the training and testing process. Ideally, therefore, the task should function like a reading task in the same way that stimuli presented outside the laboratory do in printed text. Admittedly, no systematic data exist to support the efficacy of any novel procedures over the MTS format, but preliminary data are promising (see Leader et al., 1996, 2000). A procedure like the REP would at least seem intuitively simpler for participants and may be of benefit for the purposes of training large and fluent relational networks. Thus, the following experiments will employ a novel training process, not unlike the REP, to establish the derived relational networks relevant in each case.

The novel procedure employed for the current research involved first exposing participants to a traditional pre-training procedure (see Steele & Hayes, 1991) to establish the arbitrary cues for Same and Opposite responding. This pre-training procedure did not involve the use of novel methodologies because the establishment of contextual cues to be used during the subsequent training and testing of relational networks was separate to and of little consequence to the acquisition of the relational networks themselves. Participants were then trained to respond to Yes and No verbal stimuli with 100% accuracy given a variety of stimulus pairs (e.g., A1 and B1, B1 and C1, C1 and D1) in the presence of the Same contextual cue. In effect, participants were trained to evaluate specific stimulus pairs as related or unrelated according to Same and Opposite relations. Importantly, the cue was positioned as it would be in a normal sentence printed in the English language. Thus, the stimulus-cue-stimulus string could be read like a sentence. Reinforcement was delivered for responding correctly to the relations presented in this way by appropriately consequating mouse clicks on the words Yes or No on the computer screen (i.e., the participant was required to confirm whether or not the relational statement was correct or incorrect).

Following training to criterion on a series of Same and Opposite stimulus relations, probes for derived relations of Same and Opposite responding among the stimuli were presented (e.g., C Same A, D Same A, C Opposite A, D Opposite A). In Experiment 3, testing for both Same and Opposite relational responding included an additional phase within the testing protocol which provided corrective feedback (MET testing) if the participant did not pass the standard Same or Opposite testing phase on the first exposure. This MET testing was administered repeatedly for each stimulus set to criterion. More specifically, testing tasks were presented as before. However, during the MET testing, participants were provided with corrective feedback on their responses. In this manner, a method of cycling training to criterion, testing, re-training on novel stimulus sets and MET testing formed the procedure for the training and testing protocol for Experiment 3.

Normally functioning children and adults with no history of training in relational responding in the laboratory context made up the participant pool for Experiment 3. This experiment involved an examination of the acquisition of the relational responding in accordance with Same and Opposite relations.

Method

Participants

Five normally developing children (ages 8-12 at the start of the experiment) were recruited from a local school (see Table 6 for exact ages). Three of these child participants were female and three were male. As in Experiments 1 and 2, children were chosen based on availability, and also on having been identified by parents and teachers as students who were not presenting with any known or suspected learning difficulties. These participants had no prior history of DRR and thus were considered naïve.

Table 6. Child Participants' ages at start of the experiment.

Participant ID	Age
P13	9 years, 4 months
P14	9 years, 4 months
P15	10 years, 2 months
P16	8 years, 0 months
P17	12 years, 0 months

Six normally functioning adults between the ages of 18 and 40 were also recruited from friends and family of the researcher to form the adult participants in Experiment 3. It was pre-established that all adult participants should have a minimum of a Leaving Certificate Examination level of education so that the participants were not different from one another before the experiment started. In fact, all adult participants had received university level education. Three of the participants had postgraduate level training and were engaged in gainful employment (P18 was a nurse. P21 was a full time academic researcher. P23 was a building contractor). One further participant was enrolled in a doctoral level research program in psychology (P22). The final two adult participants were enrolled in the final year of their undergraduate degrees in psychology (P19, P20). None of the participants had any knowledge of the current research programme and had not taken part in previous experiments related to DRR. Participants were chosen from a sample of convenience and did not present with any known or suspected learning difficulties.

Setting and Materials

Each child participated in the experiment in his/her own home with a parent present in the house, but not in the same room. As in experiments 1 and 2, a quiet private room was always used to minimise distractibility. Times for sessions were also chosen based on when there were least likely to be external distracters (such as the presence of other family members) in the house. A room was also chosen in each house based on the presence of a large table where each child could work comfortably.

Each adult participant for this experiment was engaged either in the computer laboratory at National University of Ireland, Maynooth (P19, P20, P21 and P22) or in their own homes (P18 and P23). As with child participants, a quiet private room was always used to minimise distraction. Times for sessions were also chosen based on when participants were available and when it was certain that there would be no external distracters in the computer laboratory (such as other classes or experiments) or in the participants' homes (e.g., other family members). Both in the laboratory and in participants' homes, it was ensured that the participants were seated comfortably in front of their designated lap top computers.

Pre-training (to establish the arbitrary cues for responding in accordance with Same and Opposite) and all relational training and testing for the establishment of Same and Opposite relations were administered on a Macintosh™ *ibook* lap top computer. The relational training for the establishment of Same and Opposite responding and testing for emergent Same and Opposite relations were controlled by software written by the author using *Psyscope*. *Psyscope* is a graphic interface and scripting language application for the creation of computer-controlled experiments in Psychology (Cohen, McWhinney, Flatt, & Provost, 1993; See also Roche & Dymond, 1999). There were a total of 27 pre-training figures (see Appendix 10). There were a total of 60 nonsense syllables employed for DRR in accordance with Same (see Appendix 11) and a total of 138 nonsense stimuli employed for DRR in accordance with Opposite (see Appendix 12)

Ethics

Experiments 1 and 2 of this thesis received ethical approval by the NUIM Ethics Committee. As the format of Experiment 3 was very similar to Experiments 1 and 2, ethical approval was not sought specifically for Experiment 3. However, at all times, ethical guidelines established by the Psychological Society of Ireland, and the British Psychological Society were observed. In addition, as in previous and subsequent experiments in this thesis, consent was obtained from each participant or from the parent/parents of each participant. Verbal consent was also obtained from each child who participated. Generic information about the general nature of the study was also provided to each adult participant and to parents of child participants prior to commencement of the research. Twice weekly sessions to train and test for the establishment of relational responding in accordance with Same and Opposite were limited to one hour sessions or until a participant requested to finish or showed any signs of distress, such as crying (for child participants), or hesitation in continuing. Each participant was permitted to take as many breaks as they needed, although few ever took a break apart from an occasional trip to the bathroom. Parents were always in the vicinity of the experiment and were welcome to be present in the room during the experiment for the child participants in this experiment. It should be noted that no parent accepted this offer.

General Experimental Sequence

The experiment took place over the course of several one to two hour sessions, depending on the availability of the participants and/or their caregivers. Research commenced following the provision of detailed descriptions of the research to the relevant authorities and to parents. All parents of child participants were also provided with a detailed consent form and weekly consultations regarding their child's participation and progress through the research programme. The reader should note that these consultations included only generic information, such as, whether or not each child was enjoying the tasks and moving at a rate that meant that they could continue to participate in the study. Thus, parents were never informed of the precise details of the Same and Opposite training and testing protocol. This was done in order to prevent confounding of the measure by parental interventions (verbal or otherwise) between sessions without the experimenter's awareness. Adult participants were merely informed that they were not yet finished with the experiment. Detailed information was only provided to the adult participants once they had completed the experiment. During sessions, each participant was seated comfortably at a large table in a quiet room or in the computer lab. The experimenter sat directly opposite the participant when in their homes and behind the participant when in the computer lab. No one else was present in the room, although for child participants, in every instance, the participants' parent(s) were either in the next room or in another room in the house. Parents were always informed that they should feel free to come in and out of the room as they chose. As mentioned previously, no parent did this. The participant was then asked if they were willing to participate in the research. If they responded, "yes" (e.g., that they were willing), then the experimenter proceeded with pre-training. Every participant agreed to continue. Sessions in this experiment consisted of five phases in total.

Phase 1 included pre-training to establish the arbitrary cues for relational responding in accordance with Same and Opposite. Phase 2 included administration of a series of relational training tasks via a lap top computer. These relational training tasks formed the training procedure for the establishment of responding in accordance with Same (i.e., Same training). Phase 3 included administration of a test for derived, or emergent, Same relations (i.e., Same testing). Phase 4 included administration of relational training tasks for the establishment of relational responding in accordance with Opposite relations (i.e., Opposite training). Finally, Phase 5 included the administration of a test for derived relational responding (DRR) in accordance with emergent Opposite relations (i.e., Opposite testing).

Phase 1: Same/Opposite Relational Pre-training.

During the pre-training phase, each participant was seated in front of the lap top computer. As in Chapter 2, the computer first displayed a screen which presented the statement "Thank you for participating." in orange coloured, size 48, Chalkboard style font. This message was accompanied by a randomly selected song from a list of popular dance music which would have been popular at the time of the experiment (see Appendix 13 for a list of the titles and artists of all songs employed). The message and music faded after 9000 milliseconds. Immediately after that, a new computer screen appeared which presented the participants with brief instructions requesting them to use the computer mouse to click on the comparison stimulus they believed to be correct (see Appendix 14 for training instructions). These instructions were presented visually on the computer screen in orange coloured, size 24 font in Chalkboard style. Identical instructions, pre-recorded by the experimenter, in the experimenter's voice, were presented at the same time in audio format. Following these audio and visual instructions, an arbitrary symbol for either Same (!!!!!!) or Opposite (%%%%%) appeared at the top of the computer screen. Then a non-arbitrary sample stimulus appeared in the middle of the screen. Following this, three non-arbitrary comparison stimuli appeared at the bottom of the screen. Of these three stimuli, one was the Same as the picture in the middle of the screen, one was the Opposite to the picture in the middle of the screen, and one was a picture of a similar stimulus that was neither the Same nor Opposite. Therefore, Opposite was defined on a continuum, so that Opposite responding was established as choosing the "most Opposite" of three stimuli. Choosing a comparison stimulus that was merely not the Same as the comparison in the presence of the Opposite cue was never reinforced. For example, in the presence of the Opposite cue, if presented with a short line as a sample, the comparisons would be another short line, a medium length line and a long line. In this example, the long line is the "most Opposite" and therefore the correct answer. The non-arbitrary stimulus sets employed here are listed in Appendix 10. See Figure 9 for an example of the tasks used in Same and Opposite pre-training.

As in Experiments 1 and 2 of the previous chapter, each session was set for one-two hours and each participant was permitted to take as many breaks as they needed, although few of them ever took a break outside of an occasional trip to the bathroom. During the pre-training phase, participants were exposed four times to each of four tasks in a randomised order, which meant that there were 16 trials in a block. Participants were re-exposed to this training until they attained a correct response on every task or until a block was cycled four times. Participants were provided with corrective feedback following their responses. When a participant attained a score of 16 correct out of 16 on a single block, they were considered to have met criterion and were exposed to a further novel stimulus set. (Stimulus sets were composed of 27 well-known shapes such as circles, lines, triangles, boxes, et cetera). Participants were again exposed four times to each of four tasks with the new stimulus set. Once again, there were 16 trials in a block. Participants were provided with corrective feedback during this portion of the training also. Once again, a block was cycled until a participant met criterion up to a total of four cycles. If a participant did not meet criterion (i.e., 100% correct responding) during this training stage, the participant was exposed to as many novel stimulus sets as necessary until they attained a score of 16 out of 16 correct in a single block. Once a participant met this criterion, they were exposed to a further novel stimulus set in the same manner as before. However, upon meeting criterion on this novel stimulus set, participants were exposed to the same sequence with another novel stimulus set, only this time, corrective feedback was not provided following responses (see Appendix 15 for instructions for this phase). If a participant met criterion during the phase with no feedback, they progressed to the relational training phase for the establishment of Same and Opposite relational responding. If participants were unable to meet criterion during a phase without feedback, they were exposed to the sequence as many times as necessary with as many stimulus sets as necessary to enable a participant to attain a perfect score (i.e., 16 out of 16 in a single block). Every participant attained this criterion. Following successful completion of the pre-training phase, a message appeared on the computer screen informing the participant that they were finished with this stage, and asked them to notify the experimenter. This message was again accompanied by popular dance music randomly selected from the same list of songs employed during the training instructions. Participants then progressed to relational training for the establishment of relational responding in accordance with Same.

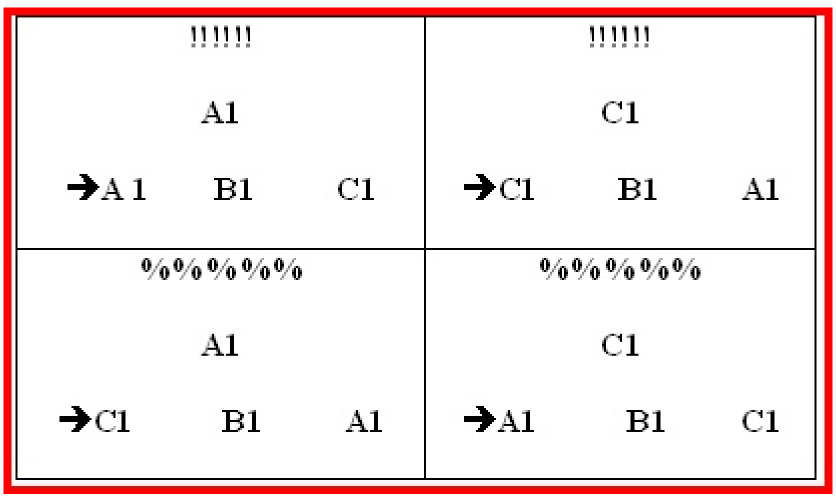


Figure 9. Tasks used for Same and Opposite relational pre-training in Phase 1. An arrow indicates the correct choice. The alphanumerics represent the non-arbitrary stimuli employed (see Appendix 13). Training and testing tasks were identical, except that feedback was provided during training, but not during testing.

Phase 2: Same Relational Training

During relational training for the establishment of responding in accordance with Same relations, participants were again seated comfortably at a large table with the experimenter seated opposing them directly across the table, or behind them if in the computer lab. Only one participant was seen at a time. Participants first received both audio and visual (written) instructions on how to proceed (see Appendix 16 for written instructions). The audio instructions were identical to the written instructions. Once again, these instructions had been pre-recorded by the author and were delivered by the computer programme. Popular dance music songs downloaded from i-tunes played softly in the background of the author's voice during the instructions. A combination of a Relational Evaluation Procedure (see Cullinan, Barnes-Holmes, & Smeets, 2001) and a Yes-No procedure (see Fields, Adams, Verhave, & Newman, 1990) was employed to establish three separate two-stimulus relations leading to the emergence of a four member relation of coordination during testing (i.e., A1 same as B1, B1 same as C1, C1 same as D1, in a linear training protocol). More specifically, on a given trial the two stimuli from a given stimulus pair (e.g., A1 and B1) were presented on screen along with the words "Yes" and "No", which were presented in the bottom left and right corner of the screen. Stimulus sets were composed of three letter nonsense syllables. Stimulus sets were selected randomly during Phase 2 and all subsequent phases. A total of 60 nonsense words were employed as stimuli (see Appendix 11). The stimuli were separated by one of the two contextual cues established in pre-training. (See Figure 10 for the tasks presented to each participant). The position of the "Yes" and "No" words was counterbalanced across trials. The participant's task was to choose "Yes" or "No" by clicking on the relevant word using the mouse and cursor. In the presence of the Same cue, therefore, if the stimulus pair were the same as each other the participant should click on the word "Yes". Alternatively, in the presence of the Opposite cue the participant was required to click on the word “No”. Choices were guided by corrective feedback following every response.

One further relation was also trained for reasons of experimental control. Specifically, participants were also trained to respond to the novel stimulus N1 as not the same as N2. This control task served the purpose of providing a history in which the word “Yes” was not chosen on every trial of this training phase in the presence of the Same cue, as was the case for the remaining relational training tasks. This served to control responding on a given trial by both the contextual cues and the relata.

The establishment of derived relational responding (DRR) in accordance with Same was trained using the following tasks; A1-B1 (yes), B1-C1 (yes), C1-D1 (yes) and N1-N2 (no), whereby the correct response is indicated here in parentheses. All stimuli were presented in black 18-point font. There were four relational evaluation training tasks comprised of four training tasks. Each of the four tasks was presented five times in a block of twenty in a quasi-random order, such that no one task was presented more than twice in succession. Feedback was presented immediately following responses in red for 1500 milliseconds and then the screen went blank. This feedback informed subjects as to whether their choice was correct or incorrect, by the word “correct” or “incorrect” appearing on the screen. The word “correct” was accompanied by a high-pitched beep and the word “incorrect” was accompanied by a low-pitched beep. Participants were required to reach a pre-set criterion of 100% correct responding across a block of 20 trials (i.e., five exposures to each of the four tasks) to finish relational training for the establishment of the emergence of Same relations. Participants were also provided with verbal (audio) feedback, delivered by the computer programme at the end of each trial if a participant had not met criterion. This verbal, or audio, feedback consisted of ten statements which had been pre-recorded by the author, in the author's voice, and which were selected randomly by the computer programme to be played at the end of each trial during which a participant had not met criterion. The feedback was intended to act as a reinforcer for continued participation. (See Appendix 17 for a list of all pre-recorded verbal feedback statements). The programme continued to deliver training blocks until criterion was reached (or up to a maximum of 20 cycles). Once criterion was reached, participants were informed by the computer in written (visual) form that they had passed the stage. A randomly selected popular dance music song from the list shown in Appendix 13 was once again played by the computer as reinforcement for continued participation. This music stopped when the participant clicked the mouse to continue. Participants then progressed to Phase 3, testing for DRR in accordance with Same.

A1 !!!!!! B1 →Yes No	B1 !!!!!! C1 →Yes No
C1 !!!!!! D1 →Yes No	N1 !!!!!! N2 Yes →No

Figure 10. Tasks used for Same relational training in Phase 2. An arrow indicates the correct choice in the presence of feedback. The alphanumerics represent the nonsense syllables employed as stimuli (see Appendix 11). !!!!!! represents the arbitrary cue for Same which had been trained during the pre-training phase.

Phase 3: Testing for Derived Same Relational Responding (Non-MET and MET)

After reaching criterion during relational training for Same relations, participants were offered a short break if they needed it, and then were exposed immediately to an almost identical procedure to Phase 2. Once again, participants first received both verbal (audio) and written (visual) instructions on how to proceed (see Appendix 18a & b for written instructions). As always, the verbal instructions were identical to the written instructions. Again, these instructions had been pre-recorded by the author, in the author's voice, and were delivered by the computer programme. Popular dance music songs again played softly in the background of the author's voice during the instructions. (See Figure 11 for the tasks presented to each participant during this testing phase. The reader should note that while the training phase employed only Same training, the testing phase employed deriving relations of both Same and Opposite responding). The only difference in Phase 3 was that feedback was not provided on initial exposures to the test. However, if a participant failed to reach the criterion on the first exposure to the test phase, they were re-exposed to Phase 2 with a new stimulus set. Participants were then re-exposed to Phase 3 employing the new stimuli and providing corrective feedback on every trial of the test. Identical to the training phase, verbal reinforcement from the selected list of pre-recorded statements was also included after each trial where criterion was not met. On blocks of testing where feedback was provided, the participant was also tested to criterion (i.e., MET testing). In other words, participants were provided with feedback on tests until such time as they attained perfect responding. Tests not involving feedback (non-MET tests or probe tests) were presented only once. Participants were exposed to this cycle of training with a novel stimulus set, testing on the novel stimulus set with no feedback for one exposure only. If a participant met criterion at this stage, the participant could move on to Phase 3. If a participant did not meet criterion on their first (and only) attempt at a test without feedback, the participant was exposed to Phase 2 again with a further novel set to criterion and then exposed to Phase 3 with feedback until the participant met criterion. A participant was required to meet criterion on their first exposure to a test without feedback (non-MET test) to move on to Phase 4. In other words, a participant was required to attain a score of 20 correct out of 20 in order to move on to training for the establishment of relations of opposition.

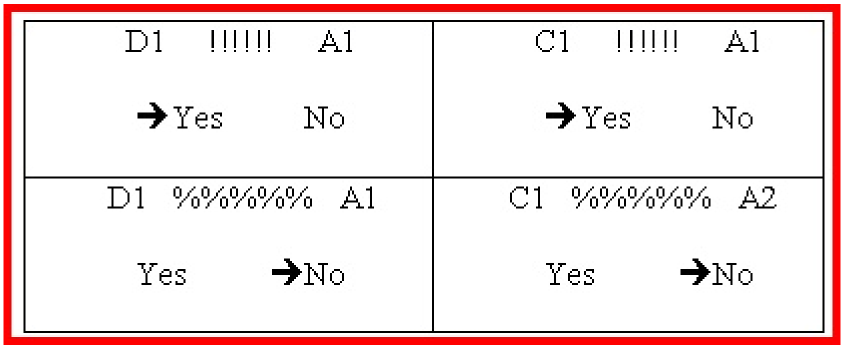


Figure 11. Tasks used for Same relational testing in Phase 3. An arrow indicates the correct choice in the presence (MET phases) or absence (non-MET phases) of feedback. The alphanumerics represent the nonsense syllables employed as stimuli (see Appendix 11). !!!!!! and %%%%% represent the arbitrary cues for Same and Opposite respectively which had been trained during the pre-training phase.

Phase4: Opposite Relational Training.

The training procedure used in Phase 4 was identical to that employed in Phase 2, with only one exception. In this phase, participants were trained to respond to the following relations; A1 opposite B1, B1 opposite C1, C1 opposite D1, N1 not opposite N2. (See Figure 12 for the tasks presented to each participant during Phase 4). The reader should also note that new stimulus sets were employed during this phase (see Appendix 12 for all nonsense syllables employed). All other instructions, procedures and criteria were employed as in Phase 2.

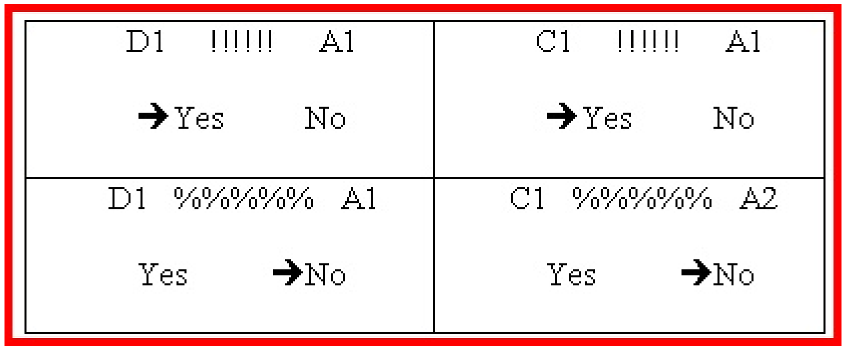


Figure 12. Tasks used for Opposite relational training in Phase 4. An arrow indicates the correct choice in the presence of feedback. The alphanumerics represent the nonsense syllables employed as stimuli (see Appendix 12). !!!!!! and %%%%% represent the arbitrary cues for Same and Opposite respectively which had been trained during the pre-training phase.

Phase 5: Testing for Derived Opposite Relations (Non-MET and MET)

Phase 5 was identical to Phase 3, except that this phase probed for derived relational responding in accordance with a frame of Opposite (see Appendix 18a+b) for visual, or written, instructions which were identical to audio instructions delivered by the computer programme as before). Once again, participants were presented with alternating test blocks, one without feedback and one with feedback. (See Figure 13 for the tasks presented to each participant during Phase 5. The reader should note that while the Opposite training phase employed only Opposite training, the testing phase employed deriving relations of both Same and Opposite responding). As soon as the participant achieved 100% correct responding on a block with no feedback, they were considered to have met criterion. If a participant did not meet criterion on a test for the demonstration of relational responding in accordance with Opposite, the participant returned to Phase 4 for relational training for the establishment of the emergence of responding in accordance with Opposite on a new stimulus set. Once the participant met criterion (i.e., 100% correct) on a novel training set, they were exposed to a test for the emergence of opposite relations, but this time with feedback, until such time as the participant produced 100% correct responding in a given block. As in previous training and MET tests, verbal reinforcement was once again delivered by the computer from the selected list of pre-recorded encouragement statements in Appendix 17. Once the participant met responding criterion, they were returned to Phase 4 and trained using a further novel stimulus set, and then presented with another probe test for the emergence of opposite relations. This probe test was identical to the first probe test for relational responding in accordance with Opposite in that the participant only received one exposure and did not receive any feedback on any response. This cycle continued until such time as each participant met criterion on a novel stimulus set on a probe phase. Participants were only considered finished with the experiment after they had met criterion on the probe test phase. (See Figure 14 below for a schematic of all training and testing tasks employed in all five phases of Experiment 3).

D1 %%%%%% A1 ➔Yes No	C1 %%%%%% A1 Yes ➔No
D1 !!!!!! A1 Yes ➔No	C1 !!!!!! A1 ➔Yes No

Figure 13. Tasks used for Opposite relational testing in Phase 5. An arrow indicates the correct choice in the presence (MET phases) or absence (non-MET phases) of feedback. The alphanumerics represent the nonsense syllables employed as stimuli (see Appendix 12). !!!!!! and %%%%% represent the arbitrary cues for Same and Opposite respectively which had been trained during the pre-training phase.

See Figure 14 below for a schematic of all training and testing tasks employed during Experiment 3.

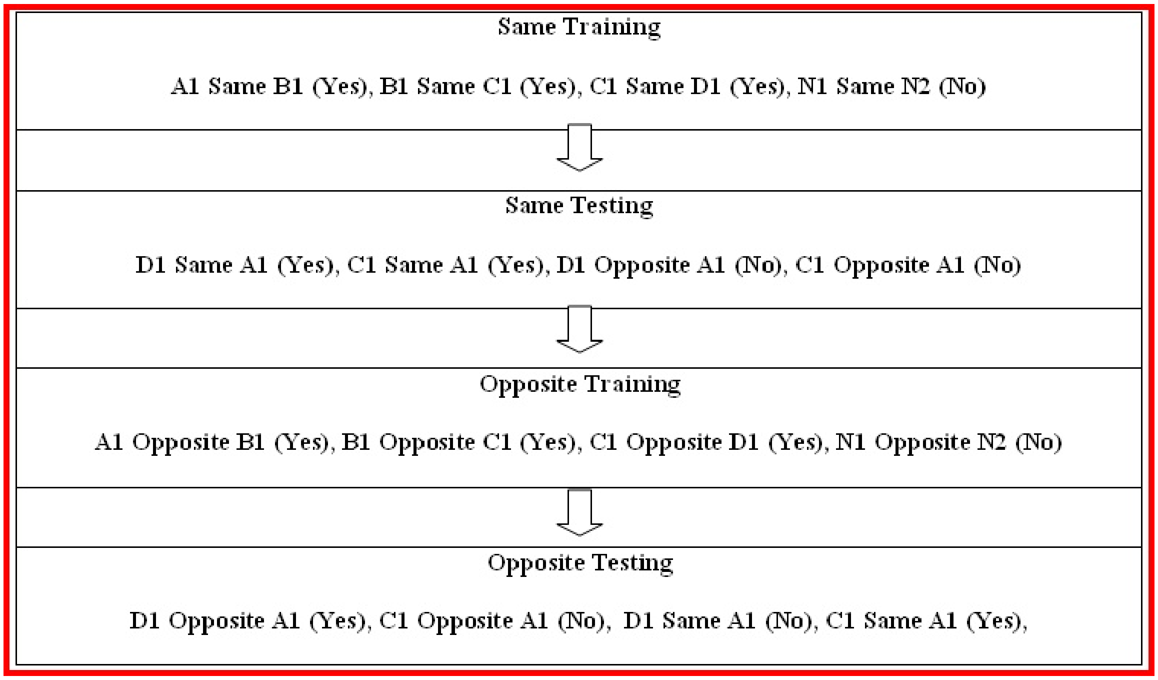


Figure 14. Schematic of the training and testing tasks employed during the five phases of Experiment 3 where Yes and No in parentheses indicate correct responses.

Results and Discussion

The data listed in this section consist of the total number of blocks required to pass the Phase 1 pre-training, Phase 2 Same relational training, the total number of blocks to pass the Phase 3 test for derivation of Same relations, the total number of blocks required to meet criterion on the Phase 4 relational training for Opposite relations and, finally the total number of blocks to pass the Phase 5 test for derivation of relational responding in accordance with Opposite. Table 7 details the number of blocks required by each participant to meet criterion for each phase of Experiment 3.

Table 7. The number of pre-training, training, testing, multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion on each of the five phases of Experiment 3.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (Same relational training)	Blocks of Same relational training	Blocks of Same relational testing	Blocks of Same Multiple Exemplar Testing	Novel stimulus sets (Opposite relational training)	Blocks of Opposite relational training	Blocks of Opposite relational testing	Blocks of Opposite Multiple Exemplar Testing
13	21	7	11	4	11	15	20	8	32
14	48	3	13	2	3	9	29	5	68
15	18	3	13	2	3	11	43	6	49
16	19	1	11	1	0	15	57	8	46
17	9	1	5	1	0	9	10	5	7
18	5	1	2	1	0	13	17	7	16
19	11	3	6	2	2	5	10	3	5
20	11	1	2	1	0	5	7	3	3
21	8	2	1	1	0	7	11	4	5
22	5	3	5	2	1	7	9	4	5
23	5	3	6	2	1	5	9	3	4

Table 7 shows the number of pre-training, training, testing, multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion on each of the five phases of Experiment 3. It can be seen from Table 7, that there was some variation among the child participants regarding number of blocks required to meet criterion on all measures. However, participants appear to have produced derived relational responding (DRR) in accordance with Same relations in fewer blocks as compared to Opposite. The results of child participants (P13-P17) will be discussed first, followed by the results of the adult participants (P18-P23).

As was expected, there was some variation among child participants in the number of blocks required to meet criterion for each task. However, more generally, it is clear that child participants required fewer exposures to training, multiple exemplar testing and testing across novel sets to meet criterion for relational responding in accordance with Same (i.e., Same relational responding) than child participants required to meet criterion for relational responding in accordance with Opposite (i.e., Opposite relational responding). Two child participants met the criterion for Same relational responding on their first exposure to a test. However, these two participants showed differences in the amount of relational training required to meet criterion. P16 required training, multiple exemplar testing, and testing across one novel stimulus set before reaching criterion for Same relational responding. P17 required exposure to training and testing across one novel stimulus set before reaching criterion for relational responding in accordance with Same. P15 required training (including retraining), multiple exemplar testing, and testing across three novel stimulus sets before reaching criterion for Same relational responding. P14 also required training (including retraining), multiple exemplar testing, and testing across three novel stimulus sets before reaching criterion for Same relational responding. P13 required training (including retraining), multiple exemplar testing, and testing across seven novel stimulus sets before reaching criterion for Same relational responding.

During relational responding in accordance with the frame of Opposition (i.e., Opposite relational responding), child participants required a higher number of exposures to training (including retraining), multiple exemplar testing, and testing across the novel stimulus sets before reaching criterion than they did to complete Same relational training. P13 required training (including retraining), multiple exemplar testing, and testing across fifteen novel stimulus sets before reaching criterion for relational responding in accordance with Opposite. P14 required training (including retraining), multiple exemplar testing, and testing across nine novel stimulus sets before reaching criterion for Opposite relational responding. P15 required exposure to training (including retraining), multiple exemplar testing, and testing across eleven novel stimulus sets before reaching criterion for Opposite relational responding. P16 required exposure to training (including retraining), multiple exemplar testing, and testing across fifteen novel stimulus sets before reaching criterion for Opposite relational responding. P17 required exposure to training (including retraining), multiple exemplar testing, and testing across nine novel stimulus sets before reaching criterion for Opposite relational responding. All child participants, except P17, required a high number of exposures to the sequence of training (including re-training), multiple exemplar testing, and testing across novel stimulus sets before reaching criterion for Opposite relational responding.

It can also be seen from Table 7 that there was less variation among the adult participants regarding number of blocks required to meet criterion on all measures than was seen among child participants in this experiment. However, as with child participants, adult participants produced DRR in accordance with Same relations in fewer blocks than required to produce DRR in accordance with Opposite relations.

As was expected, there was again some variation in the number of blocks required to meet criterion for each task. It is clear from Table 7 that adult participants required fewer exposures to training (including re-training), multiple exemplar testing and testing across novel sets to demonstrate DRR in accordance with Same than were required to demonstrate DRR in accordance with Opposite relations. Three adult participants (P18, P20 and P21) met criterion for relational responding in accordance with Same on their first novel stimulus set. These three participants also demonstrated identical requirements for exposure to training cycles. P18, P20 and P21 also required the same amount of exposure to testing (one cycle) and had no need of multiple exemplar testing. The other three adult participants (P19, P22 and P23) showed similar patterns of responding in accordance with Same relations in that they met criterion after exposure to training (including re-training), multiple exemplar testing and testing across three novel stimulus sets.

Similar to the results of child participants, all adult participants met criterion for both Same and Opposite relational responding. However, despite the similar pattern, adults on the whole, required many less blocks of Opposite relational training, multiple exemplar testing, and testing for derivation of relational responding in accordance with Opposite than did the child participants. One adult participant (P18) was an exception, in that she required greater exposure to relational training (including re-training), multiple exemplar testing, and testing across more stimulus sets than her adult counterparts in the current experiment. P18 aside, other adult participants in this group met criterion with exposure to relatively few training and testing cycles. P19, P20 and P23 required training (including re-training), multiple exemplar testing and testing across exposure to five novel stimulus sets before meeting criterion. P21 and P22 required training (including re-training), multiple exemplar testing, and testing across exposure to seven novel stimulus sets before demonstrating successful generalisation of relational responding in accordance with Opposite.

As with child participants, all adult participants met criterion for derived relational responding in accordance with both Same and Opposite relations. Once again, it is clear that adult participants required more total block numbers across training and testing before meeting criterion on deriving relations in accordance with Opposite than they required in meeting criterion on deriving relations in accordance with Same. However, adults, on the whole, required fewer total blocks of relational training for the derivation of relational responding in accordance with Opposite and fewer blocks of testing to demonstrate the successful derivation of relational responding in accordance with Opposite than did the child participants in Experiment 3. Thus, it would seem that either the age of the adult participants in Experiment 3 or their higher levels of exposure to relational responding in the natural environment (due to age and/or general experience) may have had an impact on the speed with which these participants were able to demonstrate the successful generalisation of relational responding in accordance with the relation of Opposite. In terms of relational responding in accordance with Same, it would seem that adults may also be able to more quickly demonstrate generalisation of relational responding. However, this discrepancy is not as apparent as it was with relational responding in accordance with Opposite.

All child and adult participants met criterion for derived relational responding in accordance with both Same and Opposite relations. Thus, it appears that the multiple exemplar intervention employed here was successful in establishing a repertoire of derived relational responding that was previously absent or weak for all of the participants. These data compliment those of Y. Barnes-Holmes, D. Barnes-Holmes, & Roche (2001) and of Y. Barnes-Holmes, D. Barnes-Holmes, Roche, & Smeets (2001) as detailed in Chapter 1 of this thesis. However, the important difference, here, is that these improvements in DRR were established with an entirely automated procedure that could be administered by anyone, even without knowledge of the DRR literature or methods. Secondly, these improvements were demonstrated with entirely arbitrary experimental stimuli that are commonly used in research of this kind with adults. In the Y. Barnes-Holmes, D. Barnes-Holmes, & Roche (2001) and Y. Barnes-Holmes, D. Barnes-Holmes, Roche et al. (2001) studies, the stimuli employed were everyday objects. To this extent, it could be argued that the DRR established in the current research was generalised along even more abstract continua than was demonstrated in the previous work of Y. Barnes-Holmes, D. Barnes-Holmes, & Roche (2001) and of Y. Barnes-Holmes, D. Barnes-Holmes, Roche et al. (2001). Finally, the use of abstract arbitrary nonsense syllables as stimuli provides a greater degree of experimental control over the history of exposure to those stimuli outside of the laboratory context, and therefore, renders the current findings more reliable than they may be if real-world objects, with unknown histories of association for each participant had been employed as stimuli.

It is clear that all participants required fewer exposures to training (including retraining), multiple exemplar testing and testing across novel stimulus sets before reaching criterion for Same relational responding than they required before reaching criterion for Opposite relational responding. Additionally, it was noted that the youngest child participant in this group (P16) was the only participant that requested regular breaks throughout the experiment. P16 was also observed to ask more frequent questions than the other participants and to engage in off-task behaviour, such as asking the experimenter about other topics, telling jokes and stories, et cetera. Thus, the question arose as to whether or not the task of deriving relations in accordance with Opposite might be difficult for all participants or merely for younger participants. This issue of potential confounds due to differing ages of child participants is addressed in all subsequent chapters by more closely matching child participants for age. In addition, in subsequent experiments, child and adult participants will be examined in separate experiments to further isolate age as a factor influencing DRR.

One might reasonably speculate that multiple exemplar training and testing is a powerful tool for establishing repertoires that are absent or weak in normally functioning participants. Such speculation could be based on the data outlined in this experiment, in which both children and adults demonstrated successful generalisation of relational responding in accordance with Opposite when provided with multiple exemplar testing blocks interspersed with normal testing blocks (i.e., non-multiple exemplar blocks).

It is interesting to note that the participants who required the most exposure to novel stimulus sets across training and testing in this experiment for both Same and Opposite relational responding were child participants. The idea that age may be a factor in skill at deriving relations or the speed with which one can derive relations has been pointed out before. Not surprisingly, it once again appears, at least at a cursory glance, that deriving relations may be easier for older participants. This is not to suggest that the DRR skills cannot, or do not, emerge for younger participants. It merely points to the evidence gathered so far in this exploratory work which may suggest that age plays some part in how quickly a participant can demonstrate DRR skills across relations.

Another interesting finding relates to differences, or the lack of differences, in the performances of adults and children for DRR in accordance with Same. All participants, regardless of age, demonstrated relational responding in accordance with Same relations. Adults required fewer blocks of training and testing across fewer novel stimulus sets to demonstrate DRR in accordance with Same. Differences in Opposite responding are difficult to interpret as such poor performances were observed for both adult and child participants.

It is interesting that adults performed so poorly on Opposite relational responding. We might have expected that for at least some participants, the Opposite relational responding repertoire would be in place. No published data exists to help us to make sense of this finding but it would appear that if Opposite relational repertoires are established outside the laboratory such that rudimentary Opposite relational responding occurs in daily life, it cannot be easily brought under the control of explicit contextual cues in the laboratory. Perhaps the Opposite relational responding of adults outside of the laboratory is at a much lower level than researchers may expect. This should be reflected in relatively low levels of ability at abstract reasoning among many adults.

According to RFT, many forms of relational responding in the world outside the laboratory are under forms of contextual control that may mix truly arbitrarily applicable relational responding with forms of responding that are under direct control by discriminative stimuli. For example, much relational responding involves “non” deriving and much more “non” deriving under direct control involves deriving. In other words, DRR likely quickly comes under the control of directly established discriminative cues. For instance, reading a stop sign for the first time in a foreign language while driving likely involves purely abstract DRR in accordance with a frame of coordination (i.e., direct translation from one language to another). Once the relation has been derived and behaviour (i.e., of stopping) has been controlled effectively, the sign can function as a direct discriminative stimulus without derived functions on future occasions. Thus, on the second encounter with the stop sign, it likely functions as a discriminative stimulus for stopping, particularly if direct consequences, such as the issuing of a fine, have occurred for failing to respond appropriately in its presence. Thus, the delineation between pure DRR in the real world and simple forms of discriminative control is blurred in reality, and it is likely that very little purely abstract DRR occurs for most adults on a daily basis. If this is so, it may not be surprising that such poor levels of generalisation of the Opposite relational responding repertoire were observed using arbitrary stimuli in the current experiment. Of course, this idea is merely speculative and itself requires further empirical investigation.

It should also be pointed out at this point that one specific potential design flaw may have at least partially obscured the effect of the testing practice on the emergence of perfectly accurate derived relational responding. This possible design flaw is related to the training and testing procedures employed. Specifically, in Experiment 3, all participants were necessarily trained to criterion on each exposure to a novel stimulus set. As can be seen in the schematic of the training and testing procedures (see Figure 14), and as outlined in Experiment 3, the "N" stimuli employed served the purpose of ensuring experimental control. Specifically, in addition to being trained on the relation; A1 Same B1 (yes), B1 Same C1 (yes) and C1 Same D1 (yes), participants were also trained to respond to the novel stimulus "N1" as not the same as "N2". This control task served the purpose of providing a history in which the word Yes was not chosen on every trial of this training phase in the presence of the Same cue, as was the case for the remaining relational training tasks. This served to control responding on a given trial by both the contextual cues and the relata. The "N" stimuli served a similar purpose in training for relational responding in accordance with Opposite. Thus, it is of course possible to speculate that participants might merely have learned that "yes" was the correct answer for every task except the tasks involving N stimuli. In other words, it is possible that during training and testing, only equivalence and non-equivalence (or "respond away") relations were established. While this was possible, it seems highly unlikely. More specifically, if a participant employed this strategy, they would not be able to meet criterion during the test. That is, during the test phases, several of the probes required participants to respond in novel ways in the presence of both samples and comparisons. For example, following training for Same relations in Experiment 3 when presented with C-A Opposite, the correct response is No. Thus, participants responded correctly by responding to the "No" stimulus in the presence of both familiar samples and a cue not employed during relational training by choosing "No". This response could not have been controlled by either the samples, or the cue alone, and thus illustrates the generalised relational nature of the response as predicted by Relational Frame Theory.

Another reason why it is speculated that the foregoing was not an experimental flaw was that it did not seem to present itself in Experiment 3 for the generalisation of relational responding in accordance with Same or Opposite. Nevertheless this is at least a tenable conclusion, but one which rightly requires empirical investigation. Thus, the following chapter attempts an improvement in the design of the multiple exemplar training and testing format before returning to other unanswered questions raised at the end of Chapter 2 and at the beginning of Chapter 3.

General Discussion

Chapter 3 set out to establish more reliable procedures for establishing increases in DRR skill in accordance with Same and Opposite relations by employing an improved experimental design and by introducing more relationally complex tasks than were employed in previous experiments. The aim of Experiment 3 was to circumvent the problem of rapid acquisition of DRR such that the ceiling effect encountered in Chapter 2 would not present itself. An alternative training procedure to MTS was employed in an attempt to allow a very large amount of relational training to take place in as short a period of time as possible. In that manner, a new procedure, which can be described as a combination of the Relational Evaluation Procedure (see Cullinan, Barnes-Holmes, & Smeets, 2001) and the Yes-No (see Fields, Adams, Verhave, & Newman, 1990) procedure, was used here. While this new procedure seemed to be an effective training strategy in that no participant in this chapter failed to meet criterion for training, it is difficult to demonstrate conclusively the effectiveness of this new procedure without making systematic experimental comparisons to other procedures. That endeavour is beyond the scope of the current thesis. Nevertheless, it does appear to have at least not been detrimental to the acquisition of trained and tested relations. Of course, the purpose of the current research was not to test the novel procedure systematically but simply to employ any and all techniques possible to help in the establishment of rapid and fluent DRR. Insofar as these types of procedures are more intuitive from the point of view of the participants (see Cullinan et al., 1998), it would appear to merit inclusion in future interventions for DRR.

It was found in Experiment 3 that relational responding in accordance with the frame of Opposition required a larger number of novel stimulus sets across training and testing than was required for these participants during relational responding in accordance with the frame of Same. Another small issue that was pointed out in Experiment 3 was that the youngest child participant (P16) was noted to engage in more off-task behaviour than other child participants. For this reason, the next experiment will also aim to match child participants more closely for age so that comparisons of participants within and across subsequent experiments involving children can be more clearly analysed. In addition, child participants will be chosen from a slightly older subject pool. In this manner, differences in the ages or possibly in the developmental, or attentional capabilities of participants due to their ages, may not be considered a confounding variable.

Several conclusions can be drawn from the foregoing findings. The first and most obvious conclusion is that, regardless of age, all participants in this chapter found DRR in accordance with Opposition more difficult to demonstrate than DRR in accordance with Same relations. This can be said due to the higher number of blocks of training, testing, re-training and MET testing required of participants throughout Experiment 3 to establish Opposite relational responding over Same relational responding. Thus, perhaps it is possible to argue that DRR in accordance with Opposite is a more complex relational skill than is DRR in accordance with Same. Consequently, it might be suggested that DRR in accordance with Same should be trained before DRR in accordance with Opposite in any intervention designed to establish a comprehensive repertoire of derived relational responding. In such an intervention, each form of relational responding effectively should act as a building block for the next. These findings, therefore, may provide the first clues as to whether or not it is possible to speculate on the order in which relations should be trained. (The reader is reminded that this question was also raised in Chapter 2). However, this suspected sequence of relational acquisition is still open to investigation and to date no sound empirical evidence for any fixed sequence exists.

The reader should note at this point that while a thorough investigation into the optimal sequence of relational training across a whole range of relational types is of interest in the current context, such an endeavour is not feasible given the ultimate applied goals of the current work. That is, the pragmatic goal of bringing a preliminary intervention package to a special needs population is now in sight. This goal requires that a keen eye be kept on developing an intervention that approximates an effective test treatment for intellectual deficit in as short a space of time as possible. To this extent, the current work is attempting to bridge the gap in applied research domains by using basic research strategies to devise preliminary, albeit not completely researched, intervention strategies to populations in need. Thus, further experiments will bear in mind the need to decide upon the optimal sequence of trained relations in an applied intervention, but will not treat the identification of the ideal sequence as a primary aim.

Another point which can be raised based on the current findings is that participants with a prior history of relational responding (adults) seemed to demonstrate DRR in accordance with Opposite across exposure to fewer novel stimulus sets and fewer exposures to training and testing (MET and non-MET) than did participants with no history of relational responding (children). Thus, perhaps the differences between child and adult participants in demonstrating DRR might be attributed to a history of DRR in the natural environment. This conclusion is of course speculative at this point, but may merit further analysis.

Despite these promising results, the scientist must always be cautious in interpreting data obtained from such a small subject pool. In addition, it remains unclear how resistant to extinction and how extensive the generalisation of this skill is to contexts outside the laboratory at this stage. More specifically, in the current chapter, Opposite relations were not studied in isolation. Thus, future experiments beyond those conducted in this thesis might need to compare Opposite relational responding acquisition across two groups, one of whom received training for the derivation of relational responding in accordance with Same (as in Experiment 3) and one of whom received only Opposite training. This issue will be addressed in the next chapter.

Chapter 4

Improving Same, Opposite and More-Than/Less-Than Relational Responding in Children and Adults Using an Amended Multiple Exemplar Procedure

Experiments 4-7

Introduction

In Experiment 3 of the previous chapter, it was found that relational responding in accordance with the frame of Opposition was established for participants across a larger number of novel stimulus sets across training and testing than was required for these participants during relational responding in accordance with the frame of Same. Several conclusions were drawn from these results. The first conclusion was that, regardless of age or exposure to DRR in a natural setting, all participants in Experiment 3 required a higher number of blocks of training, testing, re-training and MET testing to meet criterion for the Opposite task than they required for the Same task. Thus, perhaps it is possible to say that DRR in accordance with Opposite is a more complex relational skill than is DRR in accordance with Same. This does not allow us to say whether or not Same relational responding is established before Opposite relational responding, or that it necessarily forms a foundational building block. However, in the absence of literature on this issue, these may be reasonable suggestions at this point.

A second conclusion was that participants with a prior history of relational responding in the natural environment (adults) seemed to demonstrate DRR in accordance with Opposite relations across exposure to fewer novel stimulus sets and fewer exposures to training and testing cycles (MET and non-MET) than did participants with a lesser history of relational responding (children) in the natural environment.

As all participants in Experiment 3 demonstrated the derivation of Same and Opposite relational responding, it was suggested that the MET intervention was successful in generating these repertoires. However, it remained unclear precisely which feature of the MET training was efficacious. For instance, was it actually necessary to train and test participants to criterion on novel sets or should the mere effect of exposure to novel sets be the crucial factor in enhancing relational skills? This issue was also raised in Chapter 2 when it was discovered that exposure to training and unreinforced testing alone in the absence of corrective feedback on novel sets appeared to enhance relational skills. The reader is reminded that this finding from Chapter 2 was in support of recent findings by Vitale, Y. Barnes-Holmes, D. Barnes-Holmes, and Campbell (2008) who found that exposure to unreinforced testing in an intervention format yielded some improvements in More-Than/Less-Than repertoires of normally functioning adults. However, Vitale et al. (2008) also noted that feedback was the most effective intervention format for facilitating responding to the target relations (i.e., More-Than/Less-Than). In following up on work by Y. Barnes-Holmes, D. Barnes-Holmes, Smeets et al. (2004), Vitale et al. (2008) investigated the utility of an intervention which combined feedback and non-arbitrary trial types, and found that such a format yielded only marginally better results in terms of facilitating repertoires of More-Than/Less-Than relational responding.

It was also noted in Chapter 3 that it is possible that a design flaw was uncovered which may need to be remediated before this developed intervention can be administered to other populations to address the questions raised in this thesis. This potential design flaw in Chapter 3 related to the foils which were designed to prevent the participant from simply responding "Yes" in the presence of every Opposite cue. This involved presenting novel "N" stimuli in the following control task; N1 Opposite N2 (No), where the correct response is indicated in the parentheses. During training, if a participant responded to A1 as Opposite to B1, B1 as Opposite to C1, C1 as Opposite to D1 and N1 as Opposite to N2, this participant was considered to have passed training and was presented with the testing tasks. It was then assumed that the participant responded correctly because the participant had learned the Opposite relations among the stimuli presented. However, it is also possible that the participant had only learned to respond "Yes" to every stimulus pair presented except for the N1, N2 stimulus pair. In other words, the presence of the "N" stimuli may have functioned as a contextual cue for the reversal of the contextual functions of the Same and Opposite stimuli. Parenthetically, this hypothetical high level of confounding contextual control is at least as complex as the form of relational control which the experimenter was attempting to establish in the first instance. While it was unlikely, therefore, to have served as a "shortcut" to correct responding, this remains a possibility. In addition, if such control was exerted by the "N" stimuli, a participant would meet criterion for training every time, but would be unable to meet the testing criterion. Nevertheless, in the interests of tight experimental control, the experiments reported in the current chapter addressed this issue by adding two additional control tasks to the Same and Opposite training protocols to further ensure contextual control by the cues and by the relata.

One final issue arose in analysing the data from Chapter 3. This issue was that variance in performances across child participants may be difficult to relate to age with such a small sample size. Therefore, to eliminate these variances it may be ideal if participants are as equal in age as possible, at least for laboratory based studies. This issue was also addressed in the current chapter

Given the foregoing, the amended training procedure for Same relational responding in the current chapter trained the following relations; A1 Same B1 (yes), B1 Same C1 (yes), C1 Same D1 (yes), A1 Opposite B1 (no), N1 Same N2 (no), N1 Opposite

N2 (yes), whereby the correct response is indicated in parentheses. The testing procedure for Same relational responding was amended in one way. Specifically, testing was not administered to criterion, even on multiple exemplar blocks. The testing procedure for Same relational responding tested for the following relations; D1 Same A1 (yes), D1 Opposite A1 (no), C1 Same A1 (yes), C1 Opposite A1 (no) whereby the correct response is again indicated in the parentheses. The amended training procedure for Opposite relational responding trained the following relations; A1 Opposite B1 (yes), B1 Opposite C1 (yes), C1 Opposite D1 (yes), A1 Same B1 (no), N1 Opposite N2 (no), N1 Same N2 (yes).

The testing procedure for Opposite relational responding was also amended insofar as multiple exemplar testing was not administered to criterion. This was an important design change because testing to criterion on individual stimulus sets, as in previous experiments, could be considered to have had two adverse effects. Firstly, this procedure does not allow for the easy comparison of training and testing blocks required for participants to meet criterion because these variations must be expressed in additional terms by the number of stimulus sets involved in administering those blocks. In contrast, by administering each test block only once before moving to training using a novel stimulus set, the number of testing blocks, and the number of novel stimulus sets employed to establish DRR become equal. Secondly, the use of as many stimulus sets as possible in a given intervention is commensurate with the use of a RFT-based MET intervention. That is, it is precisely the generalisation of DRR across formal relations that facilitates its generalisation across arbitrary stimulus relations. Thus, by reducing the number of test blocks with individual sets, additional stimulus sets will likely be employed for each participant. The testing procedure for Opposite relational responding now tested for the following relations; D1 Opposite A1 (yes), D1 Same A1 (no), C1 Opposite A1 (no), C1 Same A1 (yes). In this manner, participants were required to learn all possible relationships among the stimulus sets before moving onto the testing phase for both Same and Opposite relational responding.

One final change was added to the multiple exemplar intervention for both Same and Opposite relations. Specifically, a remedial training and testing level was added for participants who were unable to pass a test with no feedback (i.e., a probe test) after a designated amount of exposure to training, probe testing, further training and multiple exemplar testing. This designated amount of exposure was set at seven cycles of relational training with alternating cycles of single test blocks which were either probe tests or multiple exemplar tests. The reason for adding a remedial level was that some normally functioning participants taking part in previous experiments struggled to meet criterion for Opposite relational responding across the maximum number of stimulus sets produced by the experimenter. As the long-term intended purpose of the foregoing procedures was to administer an intervention to a learning disabled population to increase relational skills, it was very important that all participants receiving the multiple exemplar intervention were able to meet criterion within a reasonable time frame (e.g., a school term).

Berens and Hayes (2007) have suggested that one possible way of facilitating the derivation of arbitrary relational responding when it is absent or weak would be to make use of non-arbitrary stimuli as well as arbitrary stimuli in training and testing protocols. (See also Y. Barnes-Holmes, D. Barnes-Holmes, Roche, & Smeets, 2004; Steele, &

Hayes, 1991; Vitale et al., 2008). With this in mind, the remedial level in the current chapter employed an identical procedure to the standard training and testing protocol described in Chapter 3, and with the amendments described above, but with one additional exception. Specifically, instead of training and testing with nonsense syllables, remedial levels involved training and testing with well-known objects, such as the ones employed in pre-training phases (See Appendix 10) of this experiment and those in previous chapters. In this manner, participants were exposed to practice at relational responding, but with stimulus sets that participated in non-arbitrary relations. For instance, participants were asked to relate abstract shapes that were the "same as" or "opposite to" each other along a physical dimension, such as size. Specifically, participants were presented with a shape such as a circle, followed by the arbitrary cue for either Same or Opposite, followed by another circle which was either the exact same as the first circle presented, much larger than the circle presented, or somewhere in between the size of the large and small circle. With the current example, if given the Same cue, reinforcement was provided for choosing the circle that was the exact same as the first circle (i.e., the small circle). If given the Opposite cue, reinforcement was provided for choosing the circle that was the most opposite to the first circle presented (i.e., the large circle). As in standard relational training and testing, participants indicated their choice by clicking on either "Yes" or "No" on the bottom left or right of the computer screen. In this regard, the remedial training was not unlike additional relational pre-training, but the format was identical to the standard relational training and testing protocol.

Five normally functioning adults with no history of training in relational responding in the laboratory context made up the participant pool for Experiment 4, which set out to examine the new multiple exemplar procedure. Ten normally functioning children with no history of training in relational responding in the laboratory context made up the participant pool for Experiment 5. These children were more closely matched for age than were child participants in previous experiments to increase the level of control in the small sample size. Half of the child participants in Experiment 5 formed the “MET” group, and thus received the multiple exemplar intervention, and the other half formed the “non-MET” group and thus received no multiple exemplar testing or remedial levels. Experiments 4 and 5 involved an examination of the acquisition of relational responding in accordance with Same and Opposite relations. These two experiments (and subsequent experiments in this chapter) employed the same alternative training procedure to MTS as was described in detail in Chapter 3, with the amendments described above. All participants were normally functioning.

Experiment 6 turned to the examination of the optimal sequence of training relational responding. More specifically, the question was raised; if a participant has a history of Same and Opposite responding, how might that impact on the acquisition of DRR for More-Than/Less-Than responding? To address this issue, all participants from Experiments 4 and 5 were re-recruited to form an “experienced” group of participants, who had previous exposure to training and testing for Same and Opposite relational responding. New participants (both adults and children) were recruited to form a “naïve” group who had no prior exposure to any type of relational responding history within the laboratory context. These groups were then compared to see whether or not differences were apparent in their skill at deriving More-Than/Less-Than relational responding. Experiment 7 raised a similar question regarding sequence of relational training. Specifically, Experiment 7 re-recruited the naïve child participants from Experiment 6 to ascertain whether or not a history of More-Than/Less-Than relational responding would have an impact on the acquisition of Opposite relational responding. By addressing these sequence issues, Experiments 6 and 7 were also checking whether or not the relations examined were functionally dependent on each other.

Method

Participants

Five normally developing adults were recruited from the Psychology Department at the National University of Ireland, Maynooth (N.U.I.M.). Thus, the participants were drawn from a sample of convenience. The adults taking part in Experiment 4 (P35, P36, P37, P38 and P39) were all over the age of 18, had a minimum of a Leaving Certificate Examination level of education and were in employment. All participants were engaged in doctoral level research in psychology but had no prior history of exposure to multiple exemplar training for the development of relational responding in the laboratory context. As such, all participants were considered “naïve”.

Setting and Materials

Each participant took part in the experiment in the computer laboratory at N.U.I.M. As in previous experiments, times for sessions were chosen based on when the laboratory was available. As in Chapter 3, pre-training (to establish the arbitrary cues for responding in accordance with Same and Opposite) and all relational training and testing for the establishment of Same and Opposite relations were administered on a Macintosh™ *ibook* lap top computer. The 27 different non-arbitrary stimuli used in pre-training and remedial levels included shapes instead of nonsense syllables. These non-arbitrary stimuli are the same as those employed in the pre-training phases for Same and Opposite relational responding in Chapter 3, and can be seen in Appendix 10. The relational training for the establishment of Same and Opposite responding and testing for emergent Same and Opposite relations were controlled by software written by the author using *Psyscope.* There were a total of 120 nonsense syllables (i.e., A, B, C, D, N1 and N2) employed in the training and testing programme for generating Same relational responding (see Appendix 19 for a list of all nonsense syllables employed in this and subsequent Same relational training and testing protocols). There were a total of 240 nonsense syllables (i.e., A, B, C, D, N1 and N2) employed in the training and testing programme for generating Opposite relational responding (see Appendix 20 for a list of all nonsense syllables employed in this and subsequent Opposite relational training and testing protocols). The cues used in training and testing for both Same and Opposite relational responding were established during pre-training phases and included "!!!!!!" to indicate a relation of Sameness and "%%%%%" to indicate a relation of Opposition.

Across Experiment 3 of Chapter 3, verbal (audio) reinforcement was delivered by the computer programme at the end of each training and MET testing (for MET groups) trial during which the participant did not meet criterion. This reinforcement included audio statements randomly selected from a pre-recorded list of ten varied encouraging statements, in the experimenter's voice, which intended to act as reinforcement for continued participation. (See Appendix 17 for a list of the encouragement statements employed.). This method of reinforcement was retained for this and subsequent chapters in an identical manner for training. However, as MET testing was no longer administered to criterion, the audio reinforcement was not retained for any testing phases. Another intended reinforcer for participation included the popular dance music which was played quietly at the beginning of verbal (audio) and written (visual) instructions, and also upon the completion of all phases. This method of reinforcement was also retained for the current and subsequent experiments in this thesis.

Ethics

Experiments 1 and 2 of this thesis received ethical approval by the NUIM Ethics Committee. As the format of Experiment 4 was very similar to Experiments 1 and 2, ethical approval was not sought specifically for Experiment 4. However, at all times, ethical guidelines established by the Psychological Society of Ireland and the British Psychological Society were observed. In addition, as in previous and subsequent experiments in this thesis, consent was obtained from all participants in verbal and written form.

General Experimental Sequence

The experiment took place over the course of several 1-2 hour sessions. Most participants took part twice weekly. Research commenced at the convenience of the participants and upon the signing of consent forms. Participants were also comprehensively debriefed on the exact nature of the procedure and purpose of the experiment when all participants completed the research. Participants were not ever informed as to the precise details of the Same and Opposite training and testing protocol while they were still engaged in the experiment. This was done in order to prevent confounding of the measure by participants discussing possible strategies for meeting criterion in between training and testing sessions outside of the experimenter's awareness. Sessions in this experiment consisted of seven phases in total. Two of the phases (i.e., remedial training and testing) were available to participants only if they were unable to meet criterion for Same and Opposite responding. Not all participants required these extra phases.

Phase 1 included pre-training to establish the arbitrary cues for relational responding in accordance with Same and Opposite. Phase 2 included administration of a series of relational training tasks via a lap top computer. These relational training tasks formed the training procedure for the establishment of responding in accordance with Same (i.e., Same training). Phase 3 included administration of a test for derived or emergent Same relations (i.e., Same testing). Phase 4 was a Same remedial phase which was administered to participants only if they were unable to pass Phase 3 after four exposures to the standard relational testing without feedback, each employing a unique stimulus set and each following training to criterion using that set. Phase 5 included administration of relational training tasks for the establishment of relational responding in accordance with Opposite relations (i.e., Opposite training). Phase 6 included the administration of a test for derived relational responding (DRR) in accordance with emergent Opposite relations (i.e., Opposite testing). Finally, Phase 7 was an Opposite remedial phase, which was administered to participants only if they were unable to pass Phase 6 after four exposures to the standard relational testing without feedback, each employing a unique stimulus set, and each following training to criterion using that set.

Phase 1: Same/Opposite Relational Pre-training

Pre-training in this experiment was identical to pre-training in the Experiment 3 of the previous chapter. Figure 15 displays an example of the tasks presented to each participant during this phase.

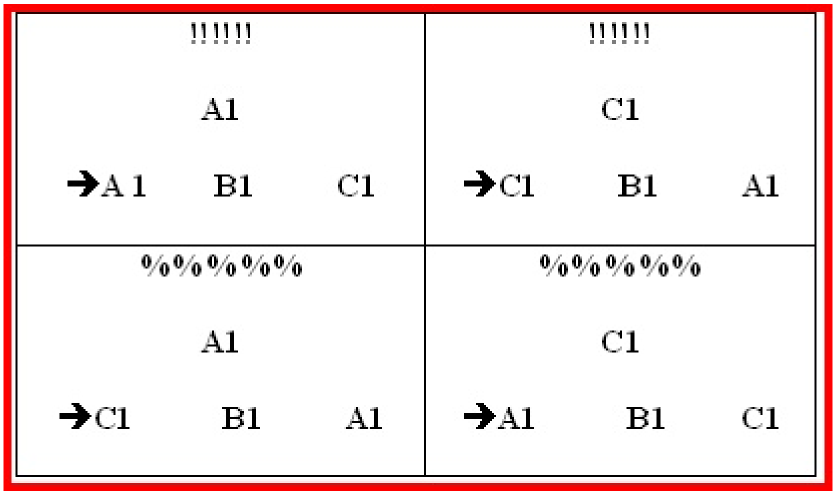


Figure 15. Tasks used for Same and Opposite relational pre-training in Phase 1. The alphanumerics represent the non-arbitrary stimulus sets employed (see Appendix 10). An arrow indicates the correct choice in the presence of feedback.

Phase 2: Same Relational Training

Same relational training in this experiment was identical to Same relational training in Experiment 3, Chapter 3 except for the addition of two control tasks which served as foils to ensure appropriate control by the contextual cues. Specifically, in addition to the tasks employed in Chapter 3, participants were also trained to respond “no” if presented with N1 same N2, and were trained to respond “yes” if presented with N1 opposite N2. These control tasks provided a history in which “yes” was not chosen on every trial of the training phase in the presence of the Same cue, as was the case for the remaining relational training tasks. This served to control responding on a given trial by both the contextual cues and the relata. In effect, six relations in total were established during this phase; A1 same B1 (yes), B1 same C1 (yes), C1 same D1 (yes), A1 opposite

B1 (no), N1 same N2 (no) and N1 opposite N2 (yes), whereby the correct response is indicated in parentheses. All stimuli were presented in black 18-point font. Each of the six tasks were presented five times each in a block of thirty in a quasi-random order, such that no one task was presented more than twice in succession. Feedback was presented in red immediately following responses in red for 1500 milliseconds and then the screen went blank. This feedback informed subjects as to whether their choice was correct or incorrect, by the word “correct” or “incorrect” appearing on the screen. The word “correct” was accompanied by a high-pitched beep and the word “incorrect” was accompanied by a low-pitched beep. Participants were required to reach a pre-set criterion of 100% correct responding across a block of thirty trials (i.e., five exposures to each of the six tasks) to finish relational training for the establishment of the emergence of Same relations. The programme continued to deliver training blocks until criterion was reached (or up to a maximum of 20 cycles). Figure 16 displays an example of the tasks used during this phase.

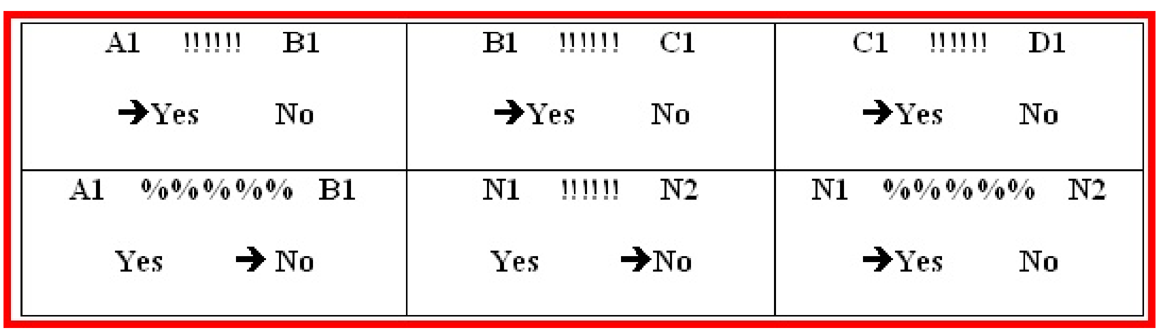


Figure 16. Tasks presented during Same relational training in Phase 2. The alphanumerics represent the nonsense syllables employed. The “!!!!!” and “%%%%%” are cues representing Same and Opposite which had been established during the Pre-training phase. An arrow indicates the correct choice in the presence of feedback.

Phase 3: Testing for Derived Same Relational Responding (Non-MET and MET)

After reaching criterion during relational training for Same relations, participants were offered a short break if required. Participants were then exposed immediately to an almost identical procedure to Phase 2. (Instructions for testing can be seen in Appendix 18a + b). See Figure 17 for the tasks presented to each participant during this testing phase. The only difference between Phase 2 and Phase 3 was that in Phase 3 feedback was not provided on initial exposures to a test with a given stimulus set and tests were not administered to criterion. Also, given that tests were no longer administered to criterion, the verbal encouragement statements were no longer employed for any testing phase during which participants failed to meet criterion.

In effect, Phase 3 consisted of alternate feedback and no-feedback tests until a participant reached 100% correct responding on a no-feedback test. Tests with feedback were indistinguishable from training except that they were not administered to criterion and did not include verbal encouragement from the pre-recorded list of statements when criterion was not met. If a participant failed to reach the criterion on the first exposure to a no-feedback test phase, they were re-exposed to Phase 2 training with feedback, using a new stimulus set. Participants were then exposed to a Phase 3 test with trial-by-trial feedback using these new stimuli. In effect, participants were trained to derive the appropriate Same relations on the test. They were then returned to training and then to another test without feedback using an entirely novel stimulus set. This process continued until the participant produced 100% correct responding on a test with no feedback.

On blocks of testing where feedback was provided (MET testing), the participant was also not tested to criterion. Thus, in the same manner as non-MET testing, MET tests or probe tests were presented only once. Participants were exposed to this cycle of training with a novel stimulus set (Phase 2), and testing on the novel stimulus set with no feedback (Phase 3) for one exposure only. If a participant met criterion at this stage, the participant was considered to have met criterion (i.e., demonstrated the derivation of Same relational responding), and was permitted to move on to Phase 5 (training for the establishment of DRR in accordance with the relation of Opposite). If a participant did not meet criterion on their first (and only) attempt at a test without feedback, the participant was exposed to Phase 2 training again with a further novel set to criterion, and then exposed to Phase 3 testing with feedback for one exposure only. A participant was required to meet criterion on their first exposure to a test without feedback (non-MET test) to move on to Phase 5. In other words, a participant was required to attain a score of 30 correct out of 30 in order to move on to training for the establishment of relations of Opposition. Given that MET tests were no longer administered to criterion, the verbal encouragement statements were no longer employed following any testing phase during which participants failed to meet criterion. Figure 17 below displays an example of the tasks used during this phase.

D1 !!!!!! A1 ➔Yes No	C1 !!!!!! A1 ➔Yes No
D1 %%%%% A1 Yes ➔No	C1 %%%%% A2 Yes ➔No

Figure 17. Tasks presented during Same relational testing in Phase 3. The alphanumerics represent the nonsense syllables employed. The “!!!!!” and “%%%%%” are the cues representing Same and Opposite which had been established during the Pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Phase 4: Same Remedial Training and Testing

If a participant was not able to pass Phase 3 after seven exposures to standard training, three exposures to multiple exemplar testing and four exposures to regular testing, each employing a unique stimulus set, and each following training to criterion using that set, the participant was exposed to Phase 4: remedial training and testing for Same relations. This phase was identical to standard training and multiple exemplar training in previous phases with one exception. Instead of utilising nonsense syllables, Phase 4 used non-arbitrary stimuli which were composed of pictures of objects used in pre-training phases (e.g., lines, circles, boxes, et cetera. See Appendix 10). In effect, six relations in total were established during this phase; A1 same B1 (yes), B1 same C1 (yes), C1 same D1 (yes), A1 opposite B1 (no), N1 same N2 (no) and N1 opposite N2 (yes), whereby the correct response is indicated in parentheses. Each of the six tasks was presented five times each in a block of thirty in a quasi-random order, such that no one task was presented more than twice in succession. Feedback was presented immediately following responses in red for 1500 milliseconds, and then the screen went blank. This feedback informed subjects as to whether their choice was correct or incorrect, by the word "correct" or "incorrect" appearing on the screen. As before, the word "correct" was accompanied by a high-pitched beep and the word "incorrect" was accompanied by a low-pitched beep. Participants were required to reach a pre-set criterion of 100% correct responding across a block of thirty trials (i.e., five exposures to each of the six tasks) to finish remedial relational training for the establishment of the emergence of Same relations. The software used throughout continued to deliver training blocks until criterion was reached (or up to a maximum of 20 cycles). It was decided in advance that if a participant did not meet criterion within the 20 cycles, the participant would be excused from further participation in the study. If a participant met criterion for the remedial relational training, then this participant moved on to remedial relational testing for Same relations. Figure 18 displays an example of the tasks used in this phase.

A1 !!!!!! B1 →Yes No	B1 !!!!!! C1 →Yes No	C1 !!!!!! D1 →Yes No
A1 %%%%%% B1 Yes → No	N1 !!!!!! N2 Yes →No	N1 %%%%%% N2 →Yes No

Figure 18. Tasks presented during remedial Same relational training in Phase 4. The alphanumerics represent the non-arbitrary stimuli employed. The "!!!!!" and "%%%%%" are cues representing Same and Opposite which had been established during the Pre-training phase. An arrow indicates the correct choice in the presence of feedback.

After reaching criterion during remedial relational training for Same relations, participants were offered a short break if required, and then were exposed immediately to an almost identical procedure to Phase 4 remedial relational training. (Instructions for testing can be seen in Appendix 18a+b. See Figure 19 for tasks presented to each participant during this testing phase. See Appendix 10 for all non-arbitrary stimuli used during Phase 4). The only difference between Phase 4 training and Phase 4 testing was that in Phase 4 testing, feedback was not provided on initial exposures to a test with a given stimulus set and tests were not administered to criterion.

Phase 4 remedial testing was identical to Phase 3 testing with the exception that non-arbitrary stimuli were used instead of arbitrary nonsense syllables. In effect, Phase 4 remedial testing consisted of alternate feedback and no-feedback tests, separated by training phases until a participant reached 100% correct responding on a no-feedback test. As before, tests with feedback were indistinguishable from training except that they were not administered to criterion. If a participant failed to reach the criterion on the first exposure to a no-feedback test phase, then they were re-exposed to Phase 4 remedial training with feedback using a new stimulus set. Participants were then exposed to a Phase 4 remedial test with trial-by-trial feedback using these new stimuli. In effect, participants were trained to derive the appropriate Same and Opposite relations on the test. They were then returned to remedial training wherein they received another remedial test without feedback using an entirely novel non-arbitrary stimulus set. This process continued until the participant attained 100% correct on a test with no feedback

On blocks of remedial testing where feedback was provided (MET remedial testing), the participant was not tested to criterion. Thus, in the same manner as non-MET remedial testing, MET remedial tests or probe remedial tests were presented only once. Participants were exposed to this cycle of training with a novel stimulus set and testing on the novel stimulus set, with no feedback, for one exposure only. If a participant met criterion at this stage, then the participant could move on to Phase 5. If a participant did not meet criterion on their first (and only) exposure to a test without feedback, the participant was exposed to Phase 4 remedial training again with a further novel set to criterion and then exposed to Phase 4 remedial testing with feedback for one cycle only. A participant was required to meet criterion on their first exposure to a remedial test without feedback (non-MET test) to progress to Phase 5. In other words, a participant was required to attain a score of 30 correct out of 30 in order to move on to training for the establishment of relations of Opposition. Figure 19 shows an example of tasks used during this phase.

D1 !!!!!! A1 →Yes No	C1 !!!!!! A1 →Yes No
D1 %%%%%% A1 Yes →No	C1 %%%%%% A2 Yes →No

Figure 19. Tasks presented during Same remedial relational testing in Phase 4. The alphanumerics represent the non-arbitrary stimuli employed. The “!!!!!” and “%%%%%” are the cues representing Same and Opposite which had been established during the Pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Upon reaching criterion on Phase 4, participants were re-exposed to Phases 2 and 3 until they met criterion (i.e., 100% correct responding on exposure to a relational test without feedback, or probe test, on the first attempt with a novel set of stimuli) or until they needed further exposure to another remedial phase. Phases 1 through 4 could be administered an infinite number of times if necessary. Participants only moved on to

Phase 5 when they had demonstrated the generalisation of Same relations with a novel stimulus set on a probe phase.

Phase 5: Opposite Relational Training.

The training procedure used in Phase 5 was similar to that employed in Phase 2. In this phase, participants were trained to respond to the following relations; A1 opposite B1 (yes), B1 opposite C1 (yes), C1 opposite D1 (yes), A1 same B1 (yes), N1 same N2 (yes), N1 opposite N2 (no), whereby the correct responses are indicated in parentheses. See Figure 20 below for the tasks presented to each participant during Phase 5. The reader should also note that new stimulus sets were employed during this phase. All other procedures and criteria were employed as in Phase 2.

A1 %%%%%% B1 ➔Yes No	B1 %%%%%% C1 ➔Yes No	C1 %%%%%% D1 ➔Yes No
A1 !!!!!! B1 Yes ➔ No	N1 %%%%%% N2 Yes ➔No	N1 !!!!!! N2 ➔Yes No

Figure 20. Tasks presented during Opposite relational training in Phase 5. The alphanumerics represent the arbitrary nonsense syllables employed. The "!!!!!" and "%%%%%" are the arbitrary cues for Same and Opposite respectively which had been established during pre-training. An arrow indicates the correct choice in the presence of feedback.

The reader is reminded that foils were utilised to prevent the participant from simply responding "Yes" in the presence of every Opposite cue. Presenting the participant with the sample, N1 opposite N2 (no), ensured that the participant could not merely respond "Yes" every time the Opposite cue was presented. In the same manner, A1 same B1 (no) and N1 same N2 (yes) were presented so that the participant could not merely respond "Yes" every time they were presented with the Same cue. These foils ensured tight contextual control by the cues and by the relata.

Phase 6: Testing for Derived Opposite Relations (Non-MET and MET)

Phase 6 was identical to Phase 3, with the exception that this phase probed for derived relational responding in accordance with a frame of Opposition. Once again, participants were presented with alternating test blocks, one without feedback and one with feedback. (See Figure 21 for the tasks presented to each participant during Phase 6). As soon as the participant achieved 100% correct responding on a block with no feedback, they were considered to have met criterion. If a participant did not meet criterion on a test for the demonstration of relational responding in accordance with Opposite, the participant returned to Phase 5 for relational training for the establishment of the emergence of responding in accordance with Opposite on a new stimulus set. Once the participant met criterion (i.e., 100% correct) on a novel training set, they were exposed to a Phase 6 test for the emergence of Opposite relations, but this time with feedback, for one exposure to a testing block. Whether or not the participant met this responding criterion, they were returned to Phase 5 and trained on a further novel stimulus set and then presented with another probe test for the emergence of Opposite relations. This probe test was identical to other probe tests for relational responding in accordance with Opposite in that the participant only received one exposure and did not receive any feedback on any response. This cycle continued until such time as each participant met criterion on a novel stimulus set on a probe phase. Participants were only considered finished with the experiment after they had met criterion on the probe test phase.

D1 %%%%%% A1 ➔Yes No	C1 %%%%%% A1 Yes ➔No
D1 !!!!!! A1 Yes ➔No	C1 !!!!!! A1 ➔Yes No

Figure 21. Tasks presented during Opposite relational testing in Phase 6. The alphanumerics represent the arbitrary nonsense syllables employed. The "!!!!!" and "%%%%%" are the arbitrary cues for Same and Opposite respectively which had been established during pre-training. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Phase 7: Opposite Remedial Training and Testing

If a participant was not able to pass Phase 6 after seven exposures to standard training (Phase 5), three exposures to Phase 6 multiple exemplar testing and four exposures to Phase 6 regular testing (each employing a unique stimulus set and each following training to criterion using that set), then a participant was exposed to Phase 7, remedial training and testing for Opposite relations. This phase was identical to standard training and multiple exemplar training in previous phases with one exception. Instead of utilising nonsense syllables, Phase 7 used non-arbitrary stimuli the same as those used in pre-training phases (e.g., lines, circles, boxes, et cetera. See Appendix 10). In effect, six relations in total were established during this phase; A1 opposite B1 (yes), B1 opposite C1 (yes), C1 opposite D1 (yes), A1 same B1 (no), N1 opposite N2 (no) and N1 same N2 (yes), whereby the correct response is indicated in parentheses. Each of the six tasks were presented five times in every block of thirty in a quasi-random order, such that no one task was presented more than twice in succession. Feedback was presented immediately following responses in red for 1500 milliseconds, and then the screen went blank. This feedback informed participants as to whether their choice was correct or incorrect, by the word “correct” or “incorrect” appearing on the screen. As before, the word “correct” was accompanied by a high-pitched beep and the word “incorrect” was accompanied by a low-pitched beep. Participants were once again required to reach a pre-set criterion of 100% correct responding across a block of thirty trials (i.e., five exposures to each of the six tasks) in order to finish remedial relational training for the establishment of the emergence of Opposite relations. The programme continued to deliver training blocks until criterion was reached (or up to a maximum of 20 cycles). Participants, once again, received verbal encouragement from the pre-recorded list of statements following those training blocks which participants failed. It was decided in advance that if a participant did not meet criterion within the 20 cycles, the participant would be excused from further participation in the study. If a participant met criterion for the remedial relational training, then this participant moved on to remedial relational testing for Opposite relations.

A1 %%%%% B1 ➔Yes No	B1 %%%%% C1 ➔Yes No	C1 %%%%% D1 ➔Yes No
A1 !!!!!! B1 Yes ➔ No	N1 %%%%% N2 Yes ➔No	N1 !!!!! N2 ➔Yes No

Figure 22. Tasks presented during remedial Opposite relational training in Phase 7. The alphanumerics represent the non-arbitrary stimuli employed. The “!!!!!” and “%%%%%” are cues representing Same and Opposite which had been established during the Pre-training phase. An arrow indicates the correct choice in the presence of feedback.

After reaching criterion during remedial relational training for Opposite relations, participants were offered a short break if they needed it, and were then were exposed immediately to an almost identical procedure to Phase 7 remedial relational training. (Instructions for testing can be seen in Appendix 18 a+b). See Figure 23 for tasks presented to each participant during this testing phase. See Appendix 10 for all non-arbitrary stimuli used during Phase 7. The only difference between Phase 7 training and Phase 7 testing was that in Phase 7 testing, feedback was not provided on initial exposures to a test with a given stimulus set and tests were not administered to criterion.

Phase 7 remedial testing was identical to Phase 6 testing with the exception that non-arbitrary stimuli were used instead of arbitrary nonsense syllables. In effect, Phase 7 remedial testing consisted of alternate feedback and no-feedback tests until a participant reached the 100% correct responding on a no-feedback test. As in remedial testing for relations of Sameness, tests with feedback were indistinguishable from training except that they were not administered to criterion. If a participant failed to reach the criterion on the first exposure to a no-feedback test phase, then they were re-exposed to Phase 7 remedial training with feedback using a new stimulus set. Participants were then exposed to a Phase 7 remedial test with trial-by-trial feedback using these new stimuli. In effect, participants were trained to derive the appropriate Same and Opposite relations on the test. They were then returned to remedial training wherein they were presented with another remedial test without feedback using an entirely novel stimulus set. This process continued until the participant attained 100% correct on a test with no feedback.

On blocks of remedial testing where feedback was provided (MET remedial testing) the participant was not tested to criterion. Thus, in the same manner as non-MET remedial testing, MET remedial tests were presented only once. Participants were exposed to this cycle of training with a novel stimulus set and testing on the novel stimulus set with no feedback, for one exposure only. If a participant met criterion at this stage, the participant could return to Phase 5 (training for Opposite relational responding). If a participant did not meet criterion on their first (and only) attempt at a test without feedback, the participant was exposed to Phase 7 remedial training again with a further novel set to criterion (with the computer generated verbal encouragement statements in the experimenter's voice following failed trials) and then exposed to Phase 7 remedial testing with feedback for one exposure only. A participant was required to meet criterion on their first exposure to a remedial test without feedback (non-MET test) to return to Phase 5 training. In other words, a participant was required to attain a score of 30 correct out of 30 in order to return to training for the establishment of relations of Opposition.

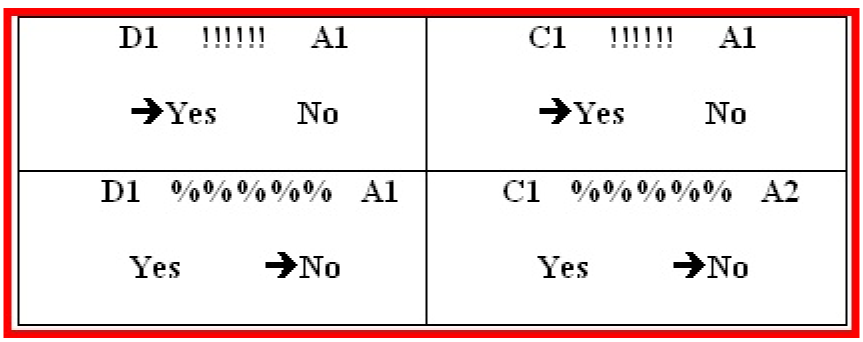


Figure 22. Tasks presented during Opposite remedial relational testing in Phase 7. The alphanumerics represent the non-arbitrary stimuli employed. The "!!!!!" and "%%%%%" are the cues representing Same and Opposite which had been established during the Pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Upon reaching criterion on Phase 7, participants were re-exposed to Phases 5 and 6 until they met criterion (i.e., 100% correct responding on exposure to a relational test without feedback, or probe test, on the first attempt with a novel set of stimuli) or until they needed further exposure to another remedial phase. Phases 5 through 7 could be administered an infinite number of times, if necessary. Participants ended their participation when they had demonstrated the generalisation of Opposite relations with a novel stimulus set on a probe phase. Figure 24 below shows a schematic of all training and testing during the four phases of Experiment 4 which employed arbitrary nonsense syllables.

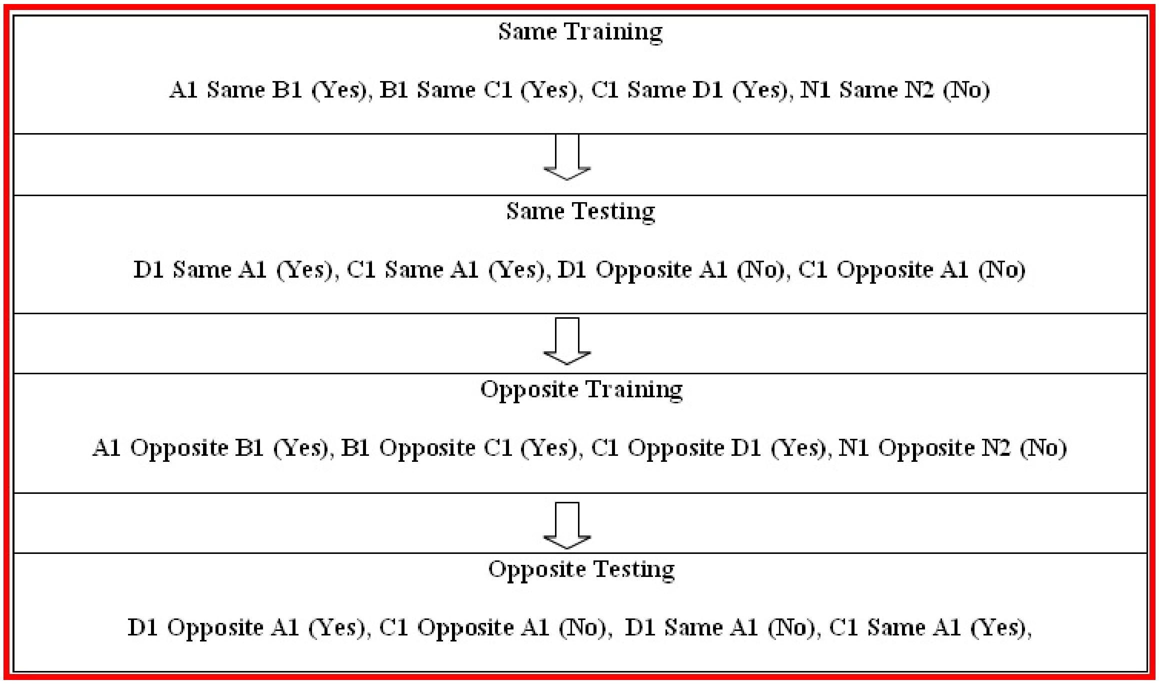


Figure 24. Schematic of the training and testing tasks employed during four of the seven phases of Experiment 4 where "Yes" and "No" in parentheses indicate correct responses.

Results and Discussion

The data listed in this section consist of the total number of blocks required to pass the Phase 1 Pre-training, the Phase 2 Same relational training, the total number of blocks to pass the Phase 3 test for derivation of Same relations, the total number of blocks required to meet criterion on the Phase 4 Same remedial training and testing, the total number of blocks required to meet criterion on the Phase 5 relational training for

Opposite relations, the total number of blocks to pass the Phase 6 test for derivation of relational responding in accordance with Opposite and finally, the total number of blocks required to meet criterion on the Phase 7 Opposite remedial training and testing. Table 8 details the number of blocks required by each participant to meet criterion for each phase of Experiment 4.

Table 8. The number of pre-training, training, testing, multiple exemplar testing blocks and the number of novel stimulus sets required by each adult MET participant to meet criterion for Same and Opposite relations on each of the phases of Experiment 4.

Participant No.	Blocks of relational pre-training required	Novel stimulus sets (Same relational training)	Blocks of Same relational training	Blocks of Same relational testing	Blocks of Same Multiple Exemplar Testing	Blocks of Same remedial training	Blocks of Same remedial testing	Novel stimulus sets (Opposite relational training)	Blocks of Opposite relational training	Blocks of Opposite relational testing	Blocks of Opposite Multiple Exemplar Testing	Blocks of Opposite remedial training	Blocks of Opposite remedial testing
35	8	1	2	1	0	0	0	5	8	3	2	0	0
36	5	3	6	2	1	0	0	1	2	1	0	0	0
37	6	1	3	1	0	0	0	13	27	7	5	1	1
38	7	3	8	2	1	0	0	7	11	4	3	0	0
39	5	1	2	1	0	0	0	3	5	2	1	0	0

Table 8 shows the number of pre-training, training, testing, multiple exemplar testing blocks and the number of novel stimulus sets required by each adult MET participant to meet criterion for Same and Opposite relations on each of the phases Experiment 4. It can be seen from this table, that there was little variation among the participants regarding number of blocks required to meet criterion for Same responding, and slight variation in the number of blocks required to meet criterion for Opposite responding. However, as in Experiment 3, Chapter 3, all participants produced derived relational responding (DRR) in accordance with Same relations in fewer blocks compared with the relation of Opposite.

As stated above, there was little variation among participants in the number of training and testing cycles required to meet criterion for Same responding. A higher amount of variation was seen among these participants in the number of training and testing cycles each required to meet criterion for Opposite relational responding. However, more generally, it is clear that participants required fewer exposures to training, multiple exemplar testing and testing across novel stimulus sets to meet criterion for relational responding in accordance with Same (i.e., Same relational responding) than participants required to meet criterion for relational responding in accordance with Opposite (i.e., Opposite relational responding). Three participants (P35, P37 and P39) met the criterion for Same relational responding on their first exposure to a test block. These three participants also required similar amounts of relational training to meet criterion (between two and three training blocks in total). P36 and P38 required six and eight blocks respectively of relational training across three novel stimulus sets. Both of these participants required three blocks of testing before reaching criterion for Same relational responding. In both cases, two of these blocks were relational tests with no feedback (i.e., probe tests), and one was a multiple exemplar test (i.e., with feedback). Together, P35, P36, P37, P38 and P39 required a mean of 4.2 blocks of relational training, a mean of 1.4 blocks of relational testing and a mean of 0.4 blocks of multiple exemplar testing before meeting criterion for Same relational responding.

During relational responding in accordance with the frame of Opposition (i.e., Opposite relational responding), participants required a higher number of exposures to training (including retraining), multiple exemplar testing and testing across the novel stimulus sets before reaching criterion than they required to complete Same relational training. (See Table 8 for a detailed account of how many blocks of training, re-training, testing and multiple exemplar testing were required by each participant during Experiment 4). P37 required training (including re-training), multiple exemplar testing and testing across 13 novel stimulus sets before reaching criterion for relational responding in accordance with Opposite. P38 required training (including re-training), multiple exemplar testing and testing across seven novel stimulus sets before reaching criterion for Opposite relational responding. P35 required exposure to training (including re-training), multiple exemplar testing and testing across five novel stimulus sets before reaching criterion for Opposite relational responding. P39 required exposure to training (including re-training), multiple exemplar testing and testing across three novel stimulus sets before reaching criterion for Opposite relational responding. P36 required exposure to training (including re-training), multiple exemplar testing and testing across one novel stimulus set before reaching criterion for Opposite relational responding. All participants, except P37, required a relatively low number of exposures to the sequence of training and re-training (between two and eleven total cycles of training). P37 required 27 cycles of training. All participants, except P37, also required low numbers of total exposures to multiple exemplar testing and testing (between one and seven) across novel stimulus sets before reaching criterion for Opposite relational responding. P37 required twelve exposures to testing (including multiple exemplar testing). P37 was also the only participant in the current experiment to require remedial training and testing. Taken together, P35, P36, P37, P38 and P39 required a mean of 10.6 cycles of relational training, a mean of 3.4 cycles of relational testing and a mean of 2.2 cycles of multiple exemplar testing before meeting criterion for Opposite relational responding.

All participants met criterion for derived relational responding in accordance with both Same and Opposite relations. Thus, it appears that the multiple exemplar intervention employed here was successful in establishing a repertoire of derived relational responding that was previously absent or weak for all of the participants or in improving the repertoire that was already in existence. Once again, these data compliment those of Y. Barnes-Holmes, D. Barnes-Holmes, and Roche (2001) and of Y. Barnes-Holmes, D. Barnes-Holmes, Roche, and Smeets (2001) as detailed in Chapter 1 of this thesis. Overall, it would seem that acquisition rates were as expected in Experiment 4. It also appears that the introduction of the extra control tasks prevented control by the contextual cues or sample stimuli alone. Thus, the additional control task introduced in Experiment 4 seem to have accelerated the rate of acquisition of DRR, or at the very least, did not impair acquisition rates for DRR for Same and Opposite relations. The addition of the remedial level, while not needed for most participants, possibly assisted one participant in producing DRR for Opposite relations. For this reason, the remedial level and the additional control tasks will be retained for future experiments in this thesis.

As in previous experiments, it is clear that participants required fewer exposures to training (including re-training), multiple exemplar testing and testing across novel stimulus sets before reaching criterion for Same relational responding than they required before reaching criterion for Opposite relational responding even when using the amended procedure. However, it can be argued that adults, on the whole, did not require a great deal of intervention in order to establish DRR even in accordance with the relation of Opposite.

It is important to understand that this experiment did not require the employment of control participants in order to fully test the utility of MET as opposed to no intervention or some other specific intervention for relational skills. This issue will be addressed again in subsequent chapters. The purpose of this experiment was to establish the in-principle utility of the particular MET training and testing protocol developed thus far before it was employed in an applied setting as a real intervention for children with learning difficulties. Thus, what is at issue, is not whether or not MET is more or less effective than any other form of intervention, but whether or not the current procedures would appear to be helpful in developing a MET intervention that can be taken to an applied setting to address precisely these types of questions in this and future research. However, before the suitability of the current procedure could be confirmed it was important to examine its effectiveness in establishing relational skills with a range of children. Experiment 5 employed identical procedures to those reported in Experiment 4, but employed children as participants.

Experiment 5

Experiment 4 aimed to amend the procedures employed in Chapter 3 by adding two extra tasks to increase experimental control over the acquisition rates of DRR skill in accordance with Same and Opposite. The intervention employed in Experiment 4 also utilised a new remedial level, which was available to participants if they were unable to meet criterion on a task after seven exposures to training and testing cycles with novel stimulus sets used during each cycle. It can be argued that the amended procedure employed in Experiment 4 succeeded in at least ensuring experimental control over DRR, if not increasing it. Thus, the amended procedure was employed in subsequent experiments. Experiment 5 employed the new procedure with a group of five normally functioning children in order to establish, in principle, that the procedure may be used to generate generalised DRR skills with such a population. To further evaluate the efficacy of this procedure, a control group composed of an additional five normally functioning children was also employed. This group was not exposed to any multiple exemplar testing or remedial levels for either Same or Opposite relational responding. However, these participants were exposed to standard multiple exemplar training (as described in Phase 2 of Experiment 4), and to the non-MET testing during which no corrective feedback was provided. The employment of the control group also allowed the experimenter to examine whether or not DRR in accordance with Same and Opposite relations emerged with exposure to relational training and unreinforced testing alone.

Method

Participants

All participants were recruited from 6th class at a local school. All participants were twelve years of age and had not taken part in any similar research investigating the emergence of DRR or the use of MET. Thus, all participants were considered naïve. As in previous experiments, children were chosen based on availability, and also on having been identified by parents and teachers as students not presenting with any known or suspected learning difficulties.

Setting and Materials

Each child participated in the experiment in their primary school in a private room set aside by the school for the children's participation in the research. Times for sessions were chosen based on when the room was available during the regular school day. All other setting and material information was identical to Experiment 4.

Ethics

Experiments 1 and 2 of this thesis received ethical approval by the NUIM Ethics Committee. As the format of Experiment 5 was very similar to Experiments 1 and 2, ethical approval was not sought specifically for Experiment 5. However, at all times, ethical guidelines established by the Psychological Society of Ireland and the British Psychological Society were observed. In addition, as in previous and subsequent experiments in this thesis, consent was obtained from all participants in verbal form and from their parents in written form.

General Experimental Sequence

The experimental sequence was identical to Experiment 5 for Participants P40, P41, P42, P43 and P44. The remaining participants, P45, P46, P47, P48 and P49 were exposed to the same sequence, but without exposure to any multiple exemplar testing or remedial phases. The experiment took place over the course of several 1-2 hour sessions. Most participants took part twice weekly. Research commenced at the convenience of the participants, their parents and teachers and upon the signing of consent forms. P40, P41, P42, P43 and P44 were exposed to all seven phases described in Experiment 4. As before, the two remedial training and testing phases were available to participants only if they were unable to meet criterion for Same and Opposite responding within the pre-established number of training and testing cycles. Not all participants required these extra phases.

Phase 1 included pre-training to establish the arbitrary cues for relational responding in accordance with Same and Opposite. Phase 2 included administration of a series of relational training tasks via a lap top computer. These relational training tasks formed the training procedure for the establishment of relational responding in accordance with Same (i.e., Same training). Phase 3 included administration of a test for derived or emergent Same relations (i.e., Same testing). Phase 4 was a Same remedial phase which was administered to participants only if they were unable to pass Phase 3 testing after four cycles of Phases 2 and 3. Phase 5 included administration of relational training tasks for the establishment of relational responding in accordance with Opposite relations (i.e., Opposite training). Phase 6 included the administration of a test for derived relational responding (DRR) in accordance with emergent Opposite relations (i.e., Opposite testing). Finally, Phase 7 was an Opposite remedial phase, which was administered to participants only if they were unable to pass Phase 6 testing after four cycles of Phases 5 and 6 training and testing. P45, P46, P47, P48 and P49 (the control participants) were not ever exposed to Phase 4 or 7 and exposure to Phases 3 and 6 for the control participants did not include any corrective feedback.

Results and Discussion

The data outlined in this section consist of; the total number of blocks required to pass the Phase 1 Pre-training, the Phase 2 Same relational training, the total number of blocks to pass the Phase 3 test for derivation of Same relations, the Phase 4 remedial training and testing for Same relations where necessary, the total number of blocks required to meet criterion on the Phase 5 relational training for Opposite relations, the total number of blocks to pass the Phase 6 test for derivation of relational responding in accordance with Opposite and finally, the Phase 7 remedial training and testing for Opposite relations where necessary. Table 9 details the number of blocks required by each participant in the experimental group to meet criterion for each phase of Experiment 5. Table 10 details the number of blocks required by each participant in the control group (i.e., no MET/remedial levels) to meet criterion for each phase of Experiment 5.

Table 9. The number of pre-training, training, testing, multiple exemplar testing blocks and the number of novel stimulus sets required by each child MET participant to meet criterion for Same and Opposite relations on each of the phases of Experiment 5.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (Same relational training)	Blocks of Same relational training	Blocks of Same relational testing	Blocks of Same Multiple Exemplar Testing	Blocks of Same remedial training	Blocks of Same remedial testing	Novel stimulus sets (Opposite relational training)	Blocks of Opposite relational training	Blocks of Opposite relational testing	Blocks of Opposite Multiple Exemplar Testing	Blocks of Opposite remedial training	Blocks of Opposite remedial testing
40	3	1	3	1	0	0	0	10	21	6	4	1	1
41	4	7	16	4	3	0	0	11	16	6	4	1	1
42	3	7	33	4	3	0	0	7	10	4	3	0	0
43	8	3	17	2	1	0	0	17	34	9	7	1	1
44	17	1	3	1	0	0	0	3	9	2	1	0	0

Table 9 shows the number of pre-training, training, testing, multiple exemplar testing blocks and the number of novel stimulus sets required by each child MET participant to meet criterion for Same and Opposite relations on each of the phases Experiment 5. It can be seen from this table, that there was little variation among the participants regarding number of blocks required to meet criterion on all measures. However, as in Chapter 3, it is clear that participants once again required fewer exposures to training, multiple exemplar testing and testing across novel sets to meet criterion for relational responding in accordance with Same (i.e., Same relational responding) than participants required to meet criterion for relational responding in accordance with Opposite (i.e., Opposite relational responding). In total, after establishing the contextual cues, P40, P41, P42, P43 and P44 required a mean of 14.4 exposures to relational training cycles and a mean of 2.4 exposures to relational testing cycles to meet the pre-established criterion for Same relational responding. P40 and P44 did not require any multiple exemplar testing. P41, P42 and P43 required exposure to three, three and one exposures to multiple exemplar testing cycles respectively to meet criterion for Same relational responding. For Opposite relational responding, P40, P41, P42, P43 and P44 required a mean of 18 exposures to relational training cycles and a mean of 5.4 exposures to relational testing cycles to meet the pre-established criterion for Opposite relational responding. All participants required multiple exemplar testing to meet criterion for Opposite relational responding (i.e., a mean of 3.8 testing cycles). Thus, all participants required a greater number of exposures to relational and multiple exemplar testing cycles to meet criterion for Opposite relational responding than they required for Same relational responding. It was also noted that child participants in Experiment 5 required exposure to training and testing across a larger number of novel stimulus sets than did the adult participants in Experiment 4 for both Same and Opposite relational responding. However, the differences between Same and Opposite DRR acquisition was not as pronounced for children in Experiment 5 as it was for adults in Experiment 4.

The participants in the control group were only exposed to relational training and probe testing. In other words, participants were trained to criterion for Same and Opposite relational responding and then tested without feedback. In effect, they did not receive any exposure to multiple exemplar testing or remedial training and testing. The control groups' results are detailed in Table 10. The data listed in Table10 consist of the total number of training blocks required to pass the Phase 1 Pre-training, Phase 2 Same relational training, the total number of testing blocks to pass the Phase 3 test for derivation of Same relations, the total number of training blocks required to meet criterion on the Phase 5 relational training for Opposite relations and finally, the total number of testing blocks to pass the Phase 6 test for derivation of relational responding in accordance with Opposite.

Table 10. The number of pre-training, training and testing blocks and the number of novel stimulus sets required by each child Non-MET participant to meet criterion for Same and Opposite relations on each of the phases in Experiment 5.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (Same relational training)	Blocks of Same relational training	Blocks of Same relational testing	Novel stimulus sets (Opposite relational training)	Blocks of Opposite relational training	Blocks of Opposite relational testing
45	8	1	8	1	7	16	7
46	7	2	7	2	23	37	23NP
47	4	2	4	2	23	61	23NP
48	9	2	6	2	12	23	12
49	12	2	12	2	23	50	23

Table 10 shows the number of pre-training, training and testing blocks and the number of novel stimulus sets required by each child Non-MET participant to meet criterion for Same and Opposite relations on each of the phases in Experiment 5. It can be seen from this table that there was more variation among the participants regarding number of blocks required to meet criterion on all measures than was seen for the experimental participants listed in Table 9. However, as in previous experiments, it is clear that participants once again required fewer exposures to relational training and testing across novel sets to meet criterion for relational responding in accordance with Same (i.e., Same relational responding) than participants required to meet criterion for relational responding in accordance with Opposite (i.e., Opposite relational responding). In total, after establishing the contextual cues, P45, P46, P47, P48 and P49 required a mean of 7.4 exposures to relational training blocks and a mean of 2.2 exposures to relational testing blocks to meet the pre-established criterion for Same relational responding. For Opposite relational responding P45, P48 and P49 required a mean of 14 exposures to relational training blocks and a mean of 29.6 exposures to relational testing blocks to meet the pre-established criterion for Opposite relational responding. P46 and

P47 were exposed to the pre-established maximum number of exposures to relational training and testing cycles without meeting criterion, and thus were excused from the experiment.

In summary, participants in both the experimental and the control group in Experiment 5 met criterion for derived relational responding in accordance with Same relations. Specifically, all participants, regardless of whether or not they received multiple exemplar testing, attained criterion for DRR in accordance with Same relations within a low number of total training and testing cycles. More specifically, all members of the control group demonstrated DRR for Same relations within one to two blocks of testing, whereas members of the multiple exemplar testing group required a slightly higher number of cycles (between one and four blocks of testing) to attain criterion. Thus, it would seem that there were no significant differences in acquisition rates for Same responding between the two groups. It would appear therefore, that multiple exemplar testing may not be necessary for the acquisition of Same relational responding. Most likely, multiple exemplar testing was not necessary for the demonstration of Same relational responding because the relational skill under examination was a relatively simple one that may in fact already have been well established in the repertoires of the participants. The reader is reminded that this point was also made with regard to derived equivalence relations in Experiment 2, Chapter 2, whereby it was suggested that little differences were seen between members of the control and members of the experimental group’s performance on the SE task because the task was a relatively basic task which was already well established in the repertoires of the participants. It was also pointed out in Experiment 2, Chapter 2, that training and unreinforced testing alone can be considered a form of MET. More specifically, it is possible that for the control participants in Experiment 2, Chapter 2, exposure to the training and testing protocol served as a form of practice that may have been somewhat effective in facilitating the acquisition of relational skills in the absence of a formal MET intervention. While practice in the absence of corrective feedback does not technically qualify as a relational intervention, it should be noted that if exposure to training and testing with multiple stimulus sets leads (as it did) to an improvement in relational fluency, then this demonstrates that relational skills must, in principle, be subject to practice effects. While we have no technical definition of what constitutes practice, for learning to take place across training and testing sessions some form of reinforcement must be occurring on a trial-by-trial and a test-by-test basis (i.e., operant conditioning). This point appears now to be applicable to the intervention for Same relational responding in the same manner as it was applicable to the SE relational responding tasks in Chapter 2.

The role of MET in the acquisition of Opposite relational responding is clearer than it is for the acquisition of Same relational responding. That is, all participants who received the multiple exemplar intervention met the pre-established criterion for Opposite relational responding. Two participants in the non-MET group did not demonstrate the generalisation of DRR in accordance with the relation of Opposite. This suggests that practice (i.e., training and unreinforced testing) alone was not so helpful for generating repertoires of Opposite DRR. Thus, it might be suggested that this provides evidence for the role of MET in the acquisition of relational responding, at least for those not already at a fluent rate in the repertoire of participants. This also supports the RFT position that DRR is a form of generalised operant behaviour that must be explicitly established.

Participants in Experiment 5 who did not receive multiple exemplar testing for Same relational responding met criterion within the same number of blocks, and in some instances in a lower number of blocks, than did the participants who received multiple exemplar testing. This does not preclude the possibility that the intervention was of some benefit in improving a DRR repertoire more generally. The rapid emergence of criterion level Same relational responding may simply result from that fact that this relational skill was at a high level of proficiency before the experimentation began. However, we do not yet know if multiple exemplar training and testing was nevertheless of benefit in some other way, such as in increasing fluency of skills that have already reached criterion level to even higher levels of fluency. This was not examined in the current research but should be considered in future research of this kind. (The issue of fluency criteria for DRR in applied interventions will be examined in Chapter 6).

In addition, MET interventions targeting fundamental relational skills may improve the likelihood that additional more complex relational skills may emerge as a result. Indeed this very issue is addressed in subsequent experiments. In effect, it is possible that the MET group employed here may have been meeting the pre-set criterion for Same relational responding in a shorter amount of time than the control participants, even if this was not apparent from the number of blocks of testing required. It would be worthwhile therefore to measure the time to criterion for all participants in future research and to explicitly use MET interventions to target this variable.

Berens and Hayes (2007) also refer to the issue of the length of time spent in training, as opposed to mere trial numbers as a possible index of relational skill. In that body of work, participants spent between two and seven months taking part. However, the varying lengths of time spent taking part in that intervention for each participant was partly related to the multiple baseline design of the study and partly related to uncontrollable factors (e.g., how many trials each participant required before meeting criterion on baseline or probe phases, how many sessions each participant was available for over the course of the study, whether or not participants were immediately exposed to probe phases following baseline phases or whether they needed to wait until the next session due to time constraints, et cetera). These time differences may partly explain differences in performance. However, because of the foregoing differences across participants, such a conclusion is difficult to arrive at in confidence.

In summary, Experiment 5 demonstrated that the current multiple exemplar intervention does seem to have been of benefit to the acquisition of DRR in accordance with Opposite. This can be seen by the lower number of blocks of relational testing required to meet criterion for participants in the MET group as compared to the non-MET group. Thus, it appears that the multiple exemplar intervention employed here was indeed more successful in establishing a repertoire of derived relational responding for Opposite relations than was relational training and unreinforced testing alone. With a working MET intervention now rapidly emerging for testing in an applied setting, a series of remaining questions need to be considered. For instance, it is still not clear what the optimal sequence of relational training should be across a battery of relational interventions comprised of Same, Opposite and perhaps others, such as More-Than and Less-Than. What needs to be addressed, therefore, is the matter of functional dependence between various repertoires of DRR. By determining which repertoires stand alone, and which require other relations to be first established to criterion, we will be in a better position to formulate an approximate sequence of training and testing for an applied intervention. Experiment 6 began by addressing these issues.

Experiment 6

Experiment 5 examined the amended procedure described in Experiment 4 with a group of ten normally functioning twelve-year-old children. The findings suggested that the developed procedure might function adequately as a preliminary intervention procedure in an applied setting to generate repertoires of DRR for Same and Opposite. However, what are also needed to produce a comprehensive intervention are procedures for training more complex relational repertoires, such as More-Than/Less-Than relations. Experiment 6, therefore, had two purposes. The first purpose was to administer a More-Than/Less-Than training procedure in addition to the Same and Opposite procedure developed across Experiments 3 and 4 and tested in Experiments 4 and 5. The second purpose was to examine the effect of the Same and Opposite multiple exemplar training (MET) procedure on the acquisition of More-Than/Less-Than relations. Thus, Experiment 6 began to directly address the question of the optimal sequence of relational training required to establish a repertoire of Same and Opposite and More-Than/Less-Than relations in the course of examining the possible functional dependence of More-Than/Less-Than relational responding on Same and Opposite relational responding. The outcome of this investigation is, at this time, difficult to predict. More specifically, it may emerge that a More-Than/Less-Than repertoire is acquired at similar rates regardless of previous Same and Opposite MET interventions to criterion. Alternatively, it may emerge that Same and Opposite MET interventions significantly increase acquisition rates of More-Than/Less-Than relational responding. If this occurs, then it is possible that acquiring More-Than/Less-Than relational responding is functionally dependent on a participant's previous acquisition of Same and Opposite relational responding.

The current experiment employed both adults and children. The adults were divided into three groups. The first group, (P35, P36, P37, P38 and P39), had a history of Same and Opposite relational responding (established in Experiment 4). The second group had no history of Same and Opposite relational responding. The second group was further divided into a group that received a multiple exemplar intervention for More-Than/Less-Than relational responding (MET group, P50, P51, P52, P53 and P54), and a group that was exposed to relational training and unreinforced testing alone (non-MET group, P55, P56, P57, P58 and P59) for More-Than/Less-Than responding. Similar groups were established for the child participants. Specifically, the MET group of children from Experiment 5 (P40, P41, P42, P43 and P44) were re-recruited to form the group with a history of Same and Opposite relational responding. A new group of ten children was also employed. Half of this group was exposed to a multiple exemplar intervention for More-Than/Less-Than relational responding (MET group, P60, P61, P62, P63, P64 and P65). The other half was exposed to relational training and unreinforced testing alone (non-MET group, P66, P67, P68, P69 and P70). The non-MET group from Experiment 5 (P45, P46, P47, P48 and P49) were also re-recruited as a comparison group for the naïve non-MET group.

Method

Participants

All adult participants were recruited from the Postgraduate Research Laboratory at National University of Ireland, Maynooth. These participants were not experienced at participating in research into derived relational responding, but several had basic knowledge of the concept of DRR from their undergraduate training. All child participants were recruited from the 6th class students at a local school. All adult participants were over the age of eighteen and were involved in full-time research in psychology. All child participants were twelve years of age. As in previous experiments of this thesis, children were chosen based on availability, and also on having been identified by parents and teachers as students who were not presenting with any known or suspected learning difficulties.

Setting and Materials

As before, each adult participated in the computer laboratory at National University of Ireland, Maynooth. Each child participated in the experiment in their primary school in a private room set aside by the school for the children's participation in the research. Times for sessions were chosen based on when these rooms were available during the regular school day. Thirty new non-arbitrary stimuli (see Appendix 21) were used for the More-Than/Less-Than pre-training phase and 120 new arbitrary nonsense syllables were used as stimuli for the More-Than/Less-Than training and testing protocols (see Appendix 22). All other setting and material information was identical to Experiments 4 and 5.

Ethics

All ethical information was identical to the ethical information in Experiments 4 and 5.

General Experimental Sequence

The experimental sequence for the More-Than/Less-Than relational training and testing protocol was similar to the sequence described for Same and Opposite relational responding described in Experiments 4 and 5. However, More-Than/Less-Than relations were trained and tested together in one single protocol. More specifically, sessions in this experiment consisted of four phases in total. Phase 4 (the remedial phase) was available to participants only if they were unable to meet criterion for More-Than/Less-Than relational responding. Not all participants required the remedial phase. In addition, only members of the MET group (P35, P36, P37, P38, P39, P40, P41, P42, P43, P44, P50, P51, P52, P53, P54, P60, P61, P62, P63, P64 and P65) were eligible to receive the multiple exemplar testing in Phase 3 and remedial training and testing (Phase 4). Members of non-MET groups (P55, P56, P57, P58, P59, P66, P67, P68, P69, P70, P45, P46, P47, P48 and P49) were only exposed to pre-training, relational training and relational testing.

Phase 1 included pre-training to establish the arbitrary cues for relational responding in accordance with More-Than and Less-Than relations. Phase 2 included administration of a series of relational training tasks via a lap top computer. These relational training tasks formed the training procedure for the establishment of responding in accordance with More-Than/Less-Than (i.e., More/Less training) relations. Phase 3 included administration of a test for derived or emergent More-Than/Less-Than relations (i.e., More/Less testing). Phase 4 was a More-Than/Less-Than remedial phase which was administered to MET group participants if they were unable to pass Phase 3 testing after four exposures. Phase 4 included remedial training and testing for More-Than/Less-Than relations using pictures (e.g., balls, houses, dogs) instead of arbitrary nonsense syllables (See Appendix 21 for all non-arbitrary stimuli used in this phase).

Phase 1: More-Than/Less-Than Relational Pre-training

Participants first received both verbal and written instructions on how to proceed (see Appendix 23). Following these verbal and written instructions, an arbitrary symbol for either More-Than ($$$$$) or Less-Than (*****) appeared at the top of the computer screen. Then, a non-arbitrary sample stimulus appeared in the middle of the screen. Following this, three non-arbitrary comparison stimuli appeared at the bottom of the screen. Of these three non-arbitrary comparison stimuli, one was More-Than the sample stimulus in the middle of the screen, one was Less-Than the sample stimulus in the middle of the screen and one was identical to the stimulus in the middle of the screen. Choosing a comparison stimulus that was identical to the sample stimulus was never reinforced in the presence of either the More-Than or the Less-Than cue. For example, in the presence of the More-Than cue, if presented with a two balls as a sample, the comparisons would be two balls, one ball and three balls. In this example, the comparison stimuli “three balls” is more than the sample stimuli “two balls” and is therefore the correct answer. See Figure 25 for an example of the tasks used in More-Than/Less-Than pre-training.

During pre-training, participants were exposed four times to each of four tasks in a randomised order, such that there were 16 trials in a training block. Participants were re-exposed to this training until they produced a correct response on every task or until a block was cycled four times. Participants were provided with corrective feedback following their responses. Participants were also once again provided with randomly selected verbal encouragement statements from the pre-recorded list (Appendix 17) following blocks during which they did not meet criterion. When a participant attained a score of 16 correct out of 16 on a single training block, they were considered to have met criterion and were exposed to a further novel stimulus set. Participants were again exposed four times to each of four tasks with the new stimulus set. Once again, there were 16 trials in a block. Participants were provided with corrective feedback and verbal encouragement statements following failed trials during training as always. Once again, a block was cycled until a participant met criterion or up to a total of four cycles. If a participant did not meet criterion (i.e., 100% correct responding) during this training stage, the participant was exposed to as many novel stimulus sets as necessary until they attained a score of 16 out of 16 correct in a single training block. Once a participant met this training criterion, they were exposed to a further novel stimulus set in the same manner as before. However, upon meeting criterion on this novel stimulus set for training, participants were exposed to the same sequence (i.e., with another novel stimulus set), except that corrective feedback and verbal encouragement statements were not provided following responses and failed trials. If a participant met criterion during this pre-training test with no feedback, the participant progressed to the relational training phase for the establishment of More-Than/Less-Than relational responding. If a participant was unable to meet criterion during the test without feedback, they were exposed to the sequence as many times, with as many stimulus sets as necessary, to enable a participant to attain a perfect score (i.e., 16 out of 16 in a single testing block). Every participant attained this criterion. Following successful completion of the pre-training phase, participants progressed to relational training for the establishment of relational responding in accordance with More-Than/Less-Than (i.e., Phase 2). Each session was set for 1- 2 hours and each participant was permitted to take as many breaks as they needed.

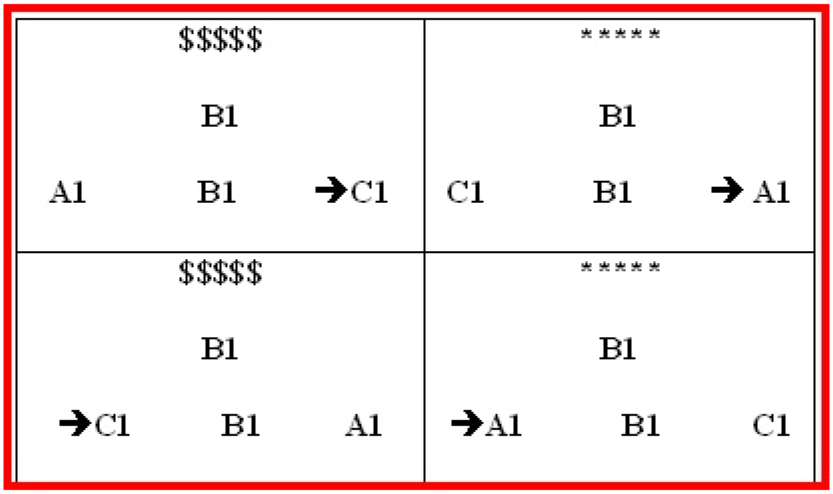


Figure 25. Tasks presented during More-Than/Less-Than relational pre-training in Phase 1. The alphanumerics represent non-arbitrary stimulus sets (see Appendix 21). The $$$$$ and ***** represent the cues for More-Than and Less-Than, respectively, which were established during this phase. An arrow indicates the correct choice in the presence (training phases) or absence (testing phases) of feedback.

Phase 2: More-Than/Less-Than Relational Training

During relational training for the establishment of responding in accordance with More-Than/Less-Than relations, (as in the previous chapter), a combination of a Relational Evaluation Procedure (see Cullinan, Barnes-Holmes, & Smeets, 2001) and a Yes-No procedure (see Fields, Adams, Verhave, & Newman, 1990) was employed to establish four separate two-stimulus relations leading to the emergence of a four member relation of coordination during testing. The trained relations were; A1 More-Than B1 (yes), B1 More-Than C1 (yes), C1 More-Than D1 (yes), A1 Less-Than B1 (no), in a linear training protocol, whereby the “yes” or “no” reflect the correct response in a given stimulus pair. (See Figure 26 below. See Appendix 24 for relational training instructions).

On a given trial the two stimuli from a given stimulus pair (e.g., A1 and B1) were presented on-screen along with the words “yes” and “no”, which were presented in the bottom left and right corner of the screen. Stimulus sets were composed of three letter nonsense syllables. Stimulus sets were selected randomly during Phase 2 and all subsequent phases. The stimuli were separated by one of the two contextual cues established in pre-training. The left-right position of the “yes” and “no” stimuli were counterbalanced across trials. The participant’s task was to choose “yes” or “no” by clicking on the relevant word using the mouse and cursor. In the presence of the More-Than cue, therefore, if the arbitrary nonsense syllable on the left of the screen was more than the arbitrary nonsense syllable on the right of the computer screen, the participant should click on the word “yes”. Alternatively, in the presence of the Less-Than cue the participant was required to click on the word “no” given the same set of circumstances. Choices were guided by corrective feedback following every response.

Two further relations were also trained for reasons of experimental control. Specifically, participants were also trained to respond “no” if presented with “N1 More-Than N2”, and were trained to respond “yes” if presented with “N1 Less-Than N2”. These control tasks served the purpose of providing a further history in which the word “yes” was not chosen on every trial of this training phase in the presence of the More-than cue. This served to control responding on a given trial by both the contextual cues and by the relata. In the absence of these control tasks, participants could merely choose "yes" in the presence of the More-Than cue and "no" in the presence of the Less-Than cue and could meet criterion without having actually learned the intended relations among the stimulus pairs presented.

During establishment of derived relational responding (DRR) in accordance with More-Than/Less-Than relations, all stimuli were presented in black 18-point font. There were six relational evaluations comprised of six relational tasks. Each of the six tasks were presented five times each in a block of thirty in a quasi-random order, such that no one task was presented more than twice in succession. Feedback was presented in red immediately following responses for 1500 milliseconds, and then the screen went blank. This feedback informed subjects as to whether their choice was correct or incorrect, by the presentation of the word "correct" or "incorrect" on the screen. The word "correct" was accompanied by a high-pitched beep, and the word "incorrect" was accompanied by a low-pitched beep. Participants were required to reach a pre-set criterion of 100% correct responding across a block of thirty trials (i.e., five exposures to each of the six tasks) to finish relational training for the establishment of the emergence of More-Than/Less-Than relations. As in previous experiments, participants were provided with verbal reinforcement from the pre-recorded audio list of encouragement statements following failed blocks of training. The programme continued to deliver training blocks until criterion was reached (or up to a maximum of 20 cycles).

A1 $$$$$ B1 →Yes No	B1 $$$$$ C1 →Yes No	C1 $$$$$ D1 →Yes No
A1 ***** B1 Yes → No	N1 $$$$$ N2 Yes →No	N1 ***** N2 →Yes No

Figure 26. Tasks presented during More-Than/Less-Than relational training in Phase 2. The alphanumerics represent the nonsense syllables employed as stimuli. The $$$$$ and ***** represent the cues for More-Than and Less-Than, respectively which had been established during the Phase 1 pre-training. An arrow indicates the correct choice in the presence of feedback.

Phase 3: Testing for Derived More-Than/Less-Than Relational Responding (Non-MET and MET)

After reaching criterion during relational training for More-Than/Less-Than relations, participants were offered a short break if they needed it, and then were exposed immediately to an almost identical procedure to Phase 2. (Instructions for testing can be seen in Appendix 25a+b. See Figure 27 for the tasks presented to each participant during this testing phase). However, in this phase no feedback was provided.

If a participant in the MET group failed to reach the criterion on the first exposure to the test phase, then they were re-exposed to Phase 2 training with a new stimulus set. Participants were then re-exposed to Phase 3 testing employing the new stimuli and with corrective feedback. On testing cycles where feedback was provided (MET testing), the participant was not exposed repeatedly to testing to criterion, but instead was required to reach criterion on the first exposure to a test. Thus, in the same manner as all non-MET testing, and in the same manner employed in Experiments 4 and 5, MET tests were presented only once. If a participant met criterion at this stage, the participant could progress to Phase 3 testing. If a participant did not meet criterion on their first (and only) exposure to a test without feedback, the participant was exposed to Phase 2 training again, with a further novel set to criterion, and then exposed to a Phase 3 test, with feedback, for one exposure only. A participant was required to meet criterion on their first exposure to a test without feedback (non-MET test) to be considered finished with the experiment. In other words, a participant was required to attain a score of 30 correct out of 30 in order to complete the phase. The reader should note that participants in the non-MET group were never provided with feedback during testing. Instead, this group was simply exposed to successive training and testing cycles, with novel stimuli, until they passed a test without feedback on the first block.

D1 $$$$$ A1 →Yes No	C1 $$$$$ A1 →Yes No
D1 ***** A1 Yes →No	C1 ***** A2 Yes →No

Figure 27. Tasks presented during More-Than/Less-Than relational testing in Phase 3. The alphanumerics represent the nonsense syllables employed as stimuli. The $$$$$ and ***** represent the cues for More-Than and Less-Than respectively which were established during the Phase 1 Pre-training. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Phase 4: *More-Than/Less-Than Remedial Training and Testing*

If a participant in the MET group was not able to pass Phase 3 after seven exposures to standard training (Phase 2), three exposures to multiple exemplar testing and four exposures to regular testing, then this participant was exposed to Phase 4, remedial training and testing for More-Than/Less-Than relations. This phase was identical to standard training and multiple exemplar training in previous phases with one exception. Instead of utilising nonsense syllables, Phase 4 used pictures of objects identical to those used in pre-training phases (e.g., dogs, cats, houses, et cetera. See Appendix 21). Progression through Phase 4 proceeded in an identical manner to remedial training and testing for Same and Opposite in Experiments 4 and 5. As in Experiments 4 and 5, participants in non-MET groups were never exposed to this phase. See Figure 28 for the tasks used during the More-Than/Less-Than remedial training phase.

A1 $$$$$ B1 →Yes No	B1 $$$$$ C1 →Yes No	C1 $$$$$ D1 →Yes No
A1 ***** B1 Yes → No	N1 ***** N2 Yes →No	N1 ***** N2 →Yes No

Figure 28. Tasks presented during remedial More-Than/Less-Than relational training in Phase 4. The alphanumerics represent the non-arbitrary stimuli employed. The “$$$$$” and “*****” are cues representing More-Than and Less-Than which had been established during the pre-training phase. An arrow indicates the correct choice in the presence of feedback.

See Figure 29 for the tasks used during More-Than/Less-Than remedial testing.

D1 ***** A1 →Yes No	C1 ***** A1 →Yes No
D1 $$$$$ A1 Yes →No	C1 $$$$$ A2 Yes →No

Figure 29. Tasks presented during Same remedial relational testing in Phase 4. The alphanumerics represent the non-arbitrary stimuli employed. The "*****" and "$$$$$" are the cues representing Less-Than and More-Than which had been established during the pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Figure 30 below shows a schematic of all training and testing during the two phases of Experiment 6 which employed arbitrary nonsense syllables.

More-Than/Less-Than Training
A1 More-Than B1 (Yes), B1 More-Than C1 (Yes), C1 More-than D1 (Yes), A1 Less-Than B1 (No), N1 More-Than N2 (No), N1 Less-Than N2 (Yes)
⇩
More-Than/Less-Than Testing
D1 More-Than A1 (No), C1 More-Than A1 (No), D1 Less-Than A1 (Yes), C1 Less-Than A1 (Yes)

Figure 30. Schematic of the training and testing tasks employed during two of the four phases of Experiment 6 (where "Yes" and "No" in parentheses indicate correct responses). The remedial level (Phase 4) is identical to More-Than/Less-Than training and testing phases, with the exception that instead of using nonsense words (as represented by alphanumerics in the schematic), non-arbitrary stimuli (images of cats, dogs, houses, et cetera.) are used. These can be seen in more detail in Appendix 21.

Results and Discussion

The data listed in this section consist of the total number of blocks required to pass the Phase 1 pre-training, the Phase 2 More-Than/Less-Than relational training, the total number of blocks to pass the Phase 3 test for derivation of More-Than/Less-Than relations and the total number of blocks to pass the Phase 4 remedial training and testing where applicable. Table 11 details the number of blocks required by each adult participant in the experimental (MET) group to meet criterion for each phase of Experiment 6. This MET group had prior exposure to DRR training and testing in accordance with Same and Opposite relations in Experiment 4. Table 12 details the number of blocks required by each adult participant in a further experimental group who did not have previous exposure to DRR with Same and Opposite relations. Table 13 details the number of blocks required by each adult participant in the control group (i.e., no MET or remedial levels) to meet criterion for each phase of Experiment 6.

Table 11. The number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each experienced adult MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing	Blocks of More/Less Multiple Exemplar testing	Blocks of More/Less remedial training	Blocks of More/Less remedial testing
35	6	1	2	1	0	0	0
36	5	1	1	1	0	0	0
37	5	5	10	3	2	0	0
38*	5	11	17	6	4	2	1*
39	4	1	1	1	0	0	0

*Note. This participant withdrew from the experiment and did not meet criterion.

The number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each experienced adult MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases Experiment 6 are shown in Table 11. It can be seen from this table that, on the whole, normally functioning adults did not require much, if any, intervention in order to demonstrate criterion level More-Than/Less-Than relational responding. Of the participants in Experiment 6, one participant (P38) withdrew from the experiment and her data will not be discussed. Three others (P35, P36, P39) did not require any multiple exemplar testing and none of the four required the remedial level. The one participant who did require multiple exemplar testing (P37), required two blocks before meeting criterion. P37 also required a higher number of novel stimulus sets (i.e., five sets) than other participants and a higher number of training blocks (i.e., ten blocks). In total, after establishing the contextual cues for More-Than/Less-Than relational responding, P35, P36, P37 and P39 required a mean of 3.5 exposures to relational training blocks and a mean of 1.5 exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding.

The performances of adults with no previous history of DRR in accordance with Same and Opposite relations were then examined. Table 12 details their performances. Overall, this group also showed a very high rate of More-Than/Less-Than relational responding, even before any multiple exemplar intervention was administered.

Table 12. The number of pre-training, training, multiple exemplar and relational testing blocks and the number of novel stimulus sets required by each naïve adult MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing	Blocks of More/Less Multiple Exemplar Testing	Blocks of remedial More/Less training	Number of blocks of remedial More/Less testing
50	5	1	2	1	0	0	0
51	4	5	22	3	2	0	0
52	7	3	5	2	1	0	0
53	7	1	2	1	0	0	0
54	8	1	2	1	0	0	0

Table 12 shows the number of pre-training, training, multiple exemplar and relational testing blocks and the number of novel stimulus sets required by each naïve adult MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases Experiment 6. It can be seen from this table that adults with no previous history of DRR in accordance with Same and Opposite responding performed at similar rates of acquisition to adults with a prior history of DRR in accordance with Same and Opposite relations. More specifically, it can once again be seen that, on the whole, normally functioning adults did not perform poorly on the More-Than/Less-Than relational responding tasks.

Of the participants in this group, all required between one and five novel stimulus sets to meet criterion during this task. Most participants required between two and five blocks of relational training in total. P51 was the exception to this in that she required 22 blocks of relational training in total. Two of the remaining four participants (P50 and P54) only required one block of relational testing before meeting criterion. P51 required three blocks of relational testing and P52 required two blocks of relational testing before meeting criterion. P50 and P54 did not require any multiple exemplar testing to meet criterion. P51 and P52 required two and one blocks of multiple exemplar testing respectively. None of these adult participants required the remedial levels of training and testing. In total, P50, P51, P52, P53 and P54 required a mean of 6.6 exposures to relational training blocks and a mean of 1.6 exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding. P51's requirement of 22 blocks of relational training may be considered a statistical outlier. Of course, one needs to recognise that in the experimental analysis of behaviour, outliers are perhaps of even more scientific interest than response patterns close to the norm. It is only by gaining control over such outliers (e.g., reducing the amount of training required) that we can be satisfied that we have established control over a behaviour of interest. This will be the pursuit of future researchers in this field. Nevertheless, at present it is important firstly to gain a picture of a typical training and testing requirement. For this reason, it should be borne in mind that the mean number of relational training and testing cycles are 2.75 and 1.25 blocks, respectively, when the performance of P51 is omitted. In effect, the performance of the typical adult shows a remarkably high level of proficiency at More-Than/Less-Than relational responding.

The performance of a group of adults who had no prior exposure to relational responding in the laboratory context, and who also had no access to the multiple exemplar intervention was then examined. This group demonstrated weaker levels of acquisition than their adult counterparts who received the MET intervention (both naïve and experienced participants). Table 13 details their performances.

Table 13. The number of pre-training, training and testing blocks and the number of novel stimulus sets required by each naïve adult non-MET (control) participants to meet criterion for More-Than/Less-Than relations on each of the phases in Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing
55	6	3	7	3
56	5	2	4	2
57	5	8	15	8
58	6	1	3	1
59	6	1	3	1

Table 13 shows the number of pre-training, training and testing blocks and the number of novel stimulus sets required by each naïve adult non-MET (control) participants to meet criterion for More-Than/Less-Than relations on each of the phases in Experiment 6. In this table, it can be seen that participants who had no prior history of DRR in accordance with Same and Opposite relational responding, and who were not exposed to the More-Than/Less-Than multiple exemplar intervention, demonstrated weaker levels of acquisition at More-Than/Less-Than relational responding than the adult participants who had previous exposure to the intervention and than the adult participants who had no previous exposure to DRR training and testing, but who were exposed to the multiple exemplar intervention for More-Than/Less-Than relations.

Of the participants in this group, after establishing the arbitrary cues for More-Than/Less-Than relational responding, two participants, P58 and P59, required only one novel stimulus set to demonstrate DRR in accordance with More-Than/Less-Than. These participants each required three blocks of relational training and one block of relational testing to demonstrate DRR in accordance with More-Than/Less-Than relations. P56 demonstrated DRR in accordance with More-Than/Less-Than relations across two novel stimulus sets, four blocks of relational training and two blocks of relational testing. P57 displayed the weakest acquisition rates by demonstrating DRR in accordance with More-Than/Less-Than relations across eight novel stimulus sets, fifteen blocks of relational training and eight blocks of relational testing. In total, after establishing the contextual cues for More-Than/Less-Than relational responding, P55, P56, P57, P58 and P59 required a mean of 6.4 exposures to relational training blocks and a mean of three exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding.

From this data set, it appears that having exposure to a history of Same and Opposite relational responding may not be necessary in order for adult participants to demonstrate the generalisation of More-Than/Less-Than relations. However, it is possible that having access to an intervention can increase the speed of acquisition for adults. (A measurement of speed per se was not employed in the current study, but this issue will be returned to in Chapter 6). It is as yet unclear whether the same findings hold true for child participants. The performances of children with various learning histories are now examined.

Table 14 details the acquisition rates of the first group of child participants who received this multiple exemplar intervention. These child participants had previous exposure to Same and Opposite relational responding in an intervention format (i.e., MET) in Experiment 5. Overall these participants required more cycles of training (including re-training), relational testing and multiple exemplar testing to meet criterion level responding than did adults with the same history.

Table 14. The number of pre-training, training (including re-training), relational and multiple exemplar testing blocks and the number of novel stimulus sets required by each experienced child MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing	Blocks of More/Less Multiple Exemplar Testing	Blocks of More/Less remedial training	Blocks of More/Less remedial testing
40	6	3	9	2	1	0	0
41	5	5	10	3	2	0	0
42	7	13	17	7	5	1	1
43	5	7	11	4	3	0	0
44	6	3	6	2	1	0	0

Table 14 shows the number of pre-training, training (including re-training), relational and multiple exemplar testing blocks and the number of novel stimulus sets required by each experienced child MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases Experiment 6. It can be seen from this table that child participants with a previous history of DRR with Same and Opposite relations required more cycles of training (including re-training), relational testing and multiple exemplar testing to meet criterion level responding than did adults with the same history. One child participant in this group also required one exposure to a remedial level of training and testing.

Overall, child participants in this group required between six blocks of relational training (P44) and seventeen blocks of relational training (P42), involving exposure to between three (P40 and P44) and thirteen novel stimulus sets (P42). Participants required between two (P40 and P44) and seven (P42) blocks of relational testing. Participants required between one (P40 and P44) and five (P42) blocks of multiple exemplar testing before demonstrating relational responding in accordance with More-Than/Less-Than relations. Only one participant (P42) required exposure to the remedial level. In total, after establishing the contextual cues for More-Than/Less-Than relational responding, child participants P40, P41, P42, P43 and P44 required a mean of 10.6 exposures to relational training blocks and a mean of 3.6 exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding. These participants also required a mean of 2.4 exposures to More-Than/Less-Than multiple exemplar testing blocks. Only one of these participants required exposure to cycles of remedial More-Than/Less-Than training and testing.

Children with no previous history of DRR in accordance with Same and Opposite relations were also administered a MET intervention for More-Than/Less-Than relations. These child participants were found to display comparable rates of acquisition of More-Than/Less-Than relational responding to their same age peers who had prior exposure to DRR with a Same and Opposite relational intervention. Table 15 details their performance across Experiment 6.

Table 15. The number of pre-training, training (including re-training), relational testing, multiple exemplar testing, and remedial blocks and the number of novel stimulus sets required by each naïve child MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing	Blocks of More/Less Multiple Exemplar Testing	Blocks of More/Less remedial training	Blocks of More/Less remedial testing
60	9	1	2	1	0	0	0
61	13	3	12	2	1	0	0
62	5	3	11	2	1	0	0
63	5	1	6	1	0	0	0
64	12	3	12	2	1	0	0
65	13	9	55	5	3	2	1

Table 15 shows the number of pre-training, training (including re-training), relational testing, multiple exemplar testing, and remedial blocks and the number of novel stimulus sets required by each naïve child MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases Experiment 6. It can be seen from this table that participants with no previous history of DRR in accordance with Same and Opposite relational responding attained similar rates of acquisition to their counterparts who had a history of DRR with Same and Opposite responding.

One participant (P65) required higher numbers of exposures to relational training (55 blocks), to relational testing (five blocks) and to multiple exemplar testing (three blocks) across nine novel stimulus sets. P65 was also the only participant to require the remedial training (two blocks) and remedial testing (one block). P65 aside, participants in this group required between two (P60) and twelve (P61 and P62) blocks of relational training, between one (P60, P63) and two (P61, P62 and P64) blocks of relational testing and between zero (P60, P63) and one (P59, P63, P64) block of multiple exemplar testing. In total, after establishing the contextual cues for More-Than/Less-Than relational responding, P60, P61, P62, P63, P64 and P65 required a mean of 16.3 exposures to relational training blocks, a mean of 2.1 exposures to relational testing blocks and a mean of one exposure to More-Than/Less-Than multiple exemplar tests to meet the pre-established criterion for More-Than/Less-Than relational responding. As noted, only one participant (P65) required any exposure to remedial training and testing cycles. Thus, with the exception of P65, participants acquired DRR in accordance with More-Than/Less-Than relations at a comparable or faster rate than did their counterparts who had a history of DRR in accordance with Same and Opposite relations.

The performances of a group of children who did not have access to MET (i.e., the control group or non-MET group), but who did have a previous history of non-MET DRR in accordance with Same and Opposite relations, were then examined in terms of relational responding in accordance with More-Than/Less-Than relations (see Table 16). The reader is reminded that this group formed the control group in Experiment 5. These participants required a greater number of exposures to training and testing cycles than did their peers in the MET groups, both naïve and experienced.

Table 16. The number of pre-training, training (including re-training) and testing blocks and the number of novel stimulus sets required by each experienced child non-MET (control) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing
45	8	13	18	13
46	15	16	18	16
47	6	15	27	15
48	7	6	10	6
49	6	16	27	16

Table 16 shows the number of pre-training, training (including re-training) and testing blocks and the number of novel stimulus sets required by each experienced child Non-MET (control) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6. It can be seen from this table that all participants in the non-MET, experienced group (P45, P46, P47, P48 and P49) required a greater number of exposures to relational training and testing blocks than did their counterparts in the MET group (P60, P61, P62, P63, P64 and P65), with the exception of P65.

Taken together, it can be seen in Tables 14 and 15 that, with the exception of P65, all participants in the MET group (P60, P61, P62, P63, P64 and P65) required fewer exposures to relational training and testing cycles than did their counterparts in the non-MET, experienced group (P45, P46, P47, P48 and P49).

The next group whose performance was examined was composed of children who had no previous exposure to DRR for Same and Opposite relational responding (either MET or non-MET). This group (P66, P67, P68, P69 and P70) was a control group and therefore had no access to multiple exemplar testing or to any remedial levels. Thus, this group was considered a naïve non-MET group. It was found that this naïve non-MET group demonstrated weaker levels of acquisition for More/Less relational responding than all adult and most child participants, regardless of prior history of relational responding or access to an intervention. These results are listed in Table 17.

Table 17. The number of pre-training, training and testing blocks and the number of novel stimulus sets required by each naïve child non-MET participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing
66	9	16	51	16
67	9	6	24	6
68	6	5	14	5
69	5	4	8	4
70	11	5	9	5

Table 17 shows the number of pre-training, training and testing blocks and the number of novel stimulus sets required by each naïve child Non-MET participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 6. It can be seen from this table that participants who had no prior history of DRR in accordance with Same and Opposite relational responding and who were not exposed to the multiple exemplar intervention demonstrated weaker levels of acquisition for More-Than/Less-Than relational responding than all adult and most child participants, regardless of prior history of relational responding or access to an intervention. This group required more total training blocks than their non-MET counterparts (from Experiment 5) who had a history of relational responding in accordance with Same and Opposite relations. However, some participants in this group required fewer exposures to testing blocks than some of their experienced counterparts from Experiment 5. Nevertheless, the differences between the child participant groups were small and given the small sample size, not appropriate for statistical analyses.

Of the participants in this group, P69 required four novel stimulus sets to demonstrate DRR in accordance with More-Than/Less-Than relations. This participant required eight blocks of relational training and four blocks of relational testing to demonstrate DRR in accordance with More-Than/Less-Than relations. P70 demonstrated DRR in accordance with More-Than/Less-Than relations across five novel stimulus sets, nine blocks of relational training and five blocks of relational testing. P68 demonstrated DRR in accordance with More-Than/Less-Than relations across five novel stimulus sets, fourteen blocks of relational training and five blocks of relational testing. P67 demonstrated DRR in accordance with More-Than/Less-Than relations across six novel stimulus sets, 24 blocks of relational training and six blocks of relational testing. P66 displayed the weakest acquisition rates by demonstrating DRR in accordance with More-Than/Less-Than relations across 16 novel stimulus sets, 51 blocks of relational training and 16 blocks of relational testing. In total, P66, P67, P68, P69 and P70 required a mean of 7.2 exposures to relational training blocks and a mean of six exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding.

The results so far have shown that adults who have had access to a MET intervention for More-Than/Less-Than relational responding (P35, P36, P37, P39, P50, P51, P52, P53, P54) outperformed their adult peers who had no access to a MET intervention of this type. This appeared to be the case regardless of whether or not participants had a history of laboratory based exposure to interventions for DRR skills in accordance with Same and Opposite relations. However, even adults who did not have exposure to a MET intervention for More-Than/Less-Than relational responding (P55, P56, P57, P58, P59), and who also had no history of prior exposure to laboratory based relational responding with other relations, demonstrated criterion level relational responding at a relatively high rate of proficiency. Thus, it appears that having a history of criterion level relational responding with earlier relations may not be necessary for adults in terms of their proficiency at demonstrating criterion level relational responding for More-Than/Less-Than relations. However, it does appear that having access to a MET intervention for More-Than/Less-Than relational responding can increase the speed of acquisition for demonstrating this relational skill.

Most child participants required higher numbers of blocks of training and testing to demonstrate relational responding in accordance with More-Than/Less-Than than did most adult participants. This applied regardless of whether or not participants had previous exposure to a MET intervention for earlier relational skills or not. This also applied even when the adults had no access to MET for More-Than/Less-Than relational responding (P55, P56, P57, P58 and P59). However, it should be noted that the differences between child and adult participants was not as noticeable when adults did not have access to the MET intervention (P55, P56, P57, P58 and P59).

Child participants who had access to the MET intervention for More-Than/Less-Than relational responding (P40, P41, P42, P43, P44, P60, P61, P62, P63, P64 and P65) displayed comparable acquisition rates, regardless of whether or not they had a history of exposure to a MET intervention format for relational responding in accordance with Same and Opposite relations. This indicates that having a history of DRR with Same and Opposite relational responding may not impact upon the acquisition of criterion level More-Than/Less-Than relational responding when participants have access to a MET intervention.

Finally, child participants who had no access to the MET intervention (P45, P46, P47, P48, P49, P66, P67, P68, P69 and P70) for More-Than/Less-Than relational responding showed weaker acquisition than child participants who had access to the MET intervention for More-Than/Less-Than relational responding (P40, P41, P42, P43, P44, P60, P61, P62, P63, P64 and P65). The child participants who had a history of exposure to the training and testing format in the absence of MET acquired criterion level relational responding in accordance with More-Than/Less-Than relational responding more quickly than did their counterparts who had no access to MET for More-Than/Less-Than relations and who also had no history of exposure to DRR with other relations (P66, P67, P68, P69 and P70). These naïve participants showed the weakest performance across all group types (i.e., adult, child, naïve, experienced, MET, non-MET). Thus, it appears that having a history of exposure to training and testing for DRR in accordance with other relations may have benefited at least one group of child participants (P45, P46, P47, P48 and P49).

Given the foregoing, some interesting points have emerged from the data sets in the current experiment. First, the reader is reminded that previous experiments have shown that Same seems to be a fundamental relation insofar as it appears to be already pre-experimentally developed to some extent for most normally functioning participants. As such, it is expected that Same relational responding should ideally be delivered at the outset of any relational intervention with an applied population. Surprisingly, the current data suggest that More-Than/Less-Than relational responding can emerge without intervention for adults, and to a lesser extent, for children also. This is the case even in the absence of a history of multiple exemplar training with other relations in the laboratory setting. Thus, More-Than/Less-Than also appears to be relatively fundamental as a relation as assessed by the speed with which criterion levels of arbitrary relational responding in accordance with this relation emerged in the current research. In addition, demonstrating More-Than/Less-Than relational responding does not appear to be functionally dependent on having an already established repertoire of Same and Opposite relational responding. Given that in the previous experiment (Experiment 5), two participants (P46 and P47) failed to demonstrate DRR in accordance with Opposite, (and that no participant has yet failed to demonstrate DRR for More-Than/Less-Than in the current experiment), we might speculate that demonstrating DRR in accordance with More-Than/Less-Than requires a less extensive training history, and possibly a smaller number of exemplars than does a demonstration of criterion level DRR in accordance with Opposite relations. This is surprising because Same and Opposite are generally considered to be established earlier in life than More-Than/Less-Than relations, and are therefore considered to be perhaps more fundamental to the acquisition of other forms of DRR. While this may indeed be the case for Same relations, the current findings suggest that this may not be the case for Opposite relations. More-Than/Less-Than relations appear to be established more quickly in the laboratory than researchers would expect, and from the data presented thus far in the current thesis, it looks like More-Than/Less-Than relations are established even more quickly than Opposite relations. Perhaps then, More-Than/Less-Than relations are more fundamental in the relational repertoires of developing children than had previously been expected. Consequently, More-Than/Less-Than relations should be trained before Opposite relations in an applied intervention for general intellectual deficit. Given the foregoing, an effective intervention might train participants to first acquire DRR in accordance with Same relations, then to acquire DRR in accordance with More-Than/Less-Than relations, and then finally to acquire DRR in accordance with Opposite. Whether or not this intervention format will prove effective is an empirical question, which Experiment 7 aimed to address.

Experiment 7

In Experiment 6, it was suggested that the optimal sequence for training DRR in an intervention might involve training Same relations first, then More-Than/Less-Than relations, and finally training Opposite relations. The reason for this suggested sequence was that the experiments in this chapter so far have indicated that criterion level Same relational responding can be generated with few blocks of multiple exemplar training and may be present with a near criterion level of fluency for many members of the population, including children. This suggests that the relation of Same may well be the most basic relation examined in the current chapter. In contrast, criterion level More-Than/Less-Than relational responding is typically produced across a greater number of stimulus sets and training and testing blocks than criterion level Same relational responding is produced. This suggests that More-Than/Less-Than responding is a higher level relational skill, albeit one that may or may not be functionally dependent on Same relational responding repertoires in ways not yet understood. Perhaps most interestingly, criterion level Opposite relational responding has been the most difficult of all the relations to establish. This suggests that Opposite relational responding is the most complex of these three forms of derived relational responding studied thus far. Nevertheless, it is not clear from the current research, or any published research to date, whether or not easily acquired forms of relational responding necessarily serve as building blocks for more complex forms. It may well be the case that all forms of derived relational responding are independent units of behaviour with little or no functional overlap between them. In fact, the results of Experiment 6 seem to indicate that establishing More-Than/Less-Than relational responding is not functionally dependent on having an established repertoire of Same and Opposite relational responding. To further examine the issue of functional dependence between relational responding repertoires, and to cast further light on the matter of an optimal training sequence, Experiment 7 was conducted.

In Experiment 7, the naïve child participants from Experiment 6 who had only been only exposed to More-Than/Less-Than relational responding were re-recruited to form the subject pool of Experiment 7. (The participants in Experiment 7 were not exposed to Same relational responding. This was because earlier experiments have shown that Same relational responding seems to be already established in the repertoire of most young children who have taken part in this research). Participants who were part of an experimental/MET group (P60, P61, P62, P63, P64 and P65) in the previous experiment continued to be part of a MET group. Members of the control/non-MET group (P66, P67, P68, P69 and P70) from Experiment 6 remained in a control group for Experiment 7. All participants were exposed to Opposite relational training and testing to determine whether or not having a history of More-Than/Less-Than relational responding, either through a multiple exemplar intervention (MET group) or merely through unreinforced training and testing alone (control group), would have an impact on the acquisition of DRR in accordance with Opposite relations. In effect, this experiment aimed to shed further light on the previous suggestion that More-Than/Less-Than relational responding may not require criterion level Opposite relational responding for its development. This was achieved by examining the effect of previous MET interventions for More-Than/Less-Than relations (Experiment 6) on the acquisition of criterion level Opposite relational responding in the current experiment. If this supposition is supported, it may suggest that Opposite relational responding is a relatively functionally independent form of DRR, or at least not dependent on forms of relational responding studied in the current research thus far. Alternatively, it may be a form of responding that depends on multiple forms of relational responding to be already-established. Exploring this issue would be exhaustive and would require a dedicated programme of research in its own right. At the very least, however, we can say at this point that establishing Opposite relational responding may require a larger number of exemplars and training trials than had been previously expected.

Participants

All participant information was the same as the child participant information listed in Experiment 6.

Setting and Materials

Setting and material information was identical to the information listed for child participants in Experiment 6.

Ethics

All ethical information was identical to the previous experiments in this chapter.

General Experimental Sequence

Phase 1: Opposite Relational Pre-training

Opposite relational Pre-training in Experiment 7 was identical to pre-training in Experiments 4 and 5 of the current chapter.

Phase 2: Opposite Relational Training

Relational training for the establishment of responding in accordance with Opposite relations proceeded in an identical manner to Opposite relational training in Experiments 4 and 5.

Phase 3: Testing for Derived Opposite Relational Responding (Non-MET and MET)

Relational testing for the establishment of responding in accordance with Opposite relations proceeded in an identical manner to Opposite relational testing in Experiments 4 and 5. As before, only participants in the experimental (MET) group were exposed to multiple exemplar tests on alternating test phases if required. Members of the control (non-MET) group were exposed to standard relational probe tests only (i.e., without feedback).

Phase 4: Opposite Remedial Training and Testing

If a participant in the experimental group did not pass Phase 3 after seven exposures to standard training (Phase 2), three exposures to multiple exemplar testing and four exposures to standard relational testing, an experimental participant was exposed to Phase 4, remedial training and testing for Opposite relations. This phase was identical to the Opposite remedial phase in Experiments 4 and 5. The reader is reminded that members of the control group were not exposed to any remedial phase.

Results and Discussion

The data listed in this section consist of the total number of blocks required to pass Phase 1 pre-training, Phase 2 Opposite relational training, the total number of blocks required to pass the Phase 3 test for derivation of Opposite relations, and the Phase 4 remedial training and testing for Opposite relations, where necessary. In Table 18, the number of blocks required by each participant in the experimental group to meet criterion for each phase of Experiment 7 is shown. In Table 19, the number of blocks required by each participant in the control group (i.e., no MET or remedial levels) to meet criterion for each phase of Experiment 7 is shown.

Table 18. The number of pre-training, training (including re-training), relational testing, multiple exemplar testing, and remedial blocks and the number of novel stimulus sets required by each MET (experimental) child participant to meet criterion for Opposite relational responding on each of the phases in Experiment 7.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (Opposite relational training)	Blocks of Opposite relational training	Blocks of Opposite relational testing	Blocks of Opposite Multiple Exemplar Testing	Blocks of Opposite remedial training	Blocks of Opposite remedial testing
60	13	21	56	11	8	3	2
61	6	19	56	10	7	2	2
62	6	13	43	7	5	1	1
63	6	11	18	6	4	1	1
64	6	7	21	4	3	0	0
65	5	9	24	5	3	4	1

In Table 18 the number of pre-training, training (including re-training), relational testing, multiple exemplar testing, and remedial blocks and the number of novel stimulus sets required by each MET (experimental) child participant to meet criterion for Opposite relational responding on each of the phases in Experiment 7 are shown. On the whole, it is clear that participants required fewer exposures to training, multiple exemplar testing and testing across novel stimulus sets to meet criterion for relational responding in accordance with More-Than/Less-Than than participants required to meet criterion for relational responding in accordance with Opposite in Experiment 6. Participants also required fewer exposures to multiple exemplar testing and remedial training and testing cycles. In Experiment 6, P60, P61, P62, P63, P64 and P65 required a mean of 16.3 exposures to relational training blocks, a mean of 2.1 exposures to relational testing blocks and a mean of one exposure to More-Than/Less-Than multiple exemplar tests to meet the pre-established criterion for More-Than/Less-Than relational responding. As noted, only one participant (P65) required any exposure to remedial training and testing cycles (see Figure 31). In Experiment 7, P60, P61, P62, P63, P64 and P65 required a mean of 36.33 exposures to relational training blocks, a mean of 7.16 exposures to relational testing blocks, a mean of five exposures to blocks of More-Than/Less-Than multiple exemplar testing and a mean of three exposures to remedial training and testing blocks (see Figure 31). These findings clearly suggest that Opposite relations are more difficult to establish than More-Than/Less-Than relations, even for participants who have received a previous history of MET for More-Than/Less-Than relational responding.

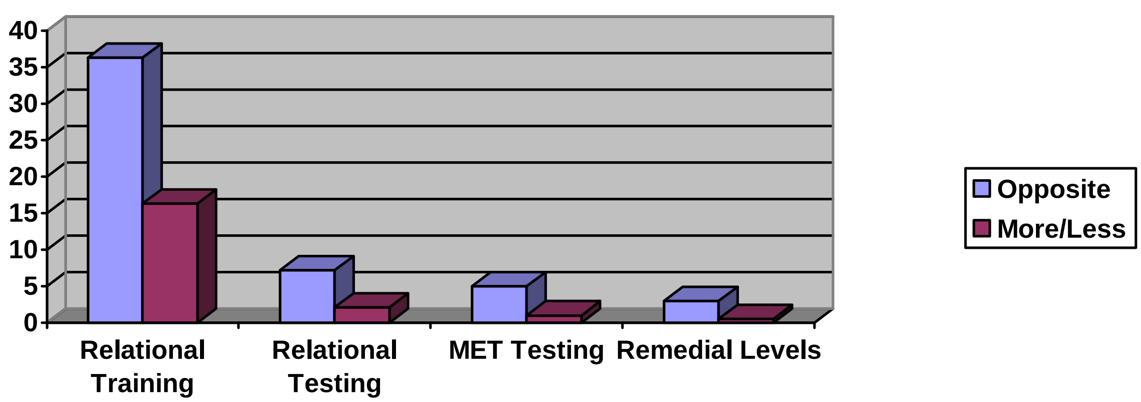


Figure 31. Number of blocks of Opposite and More-Than/Less-Than training and testing tasks required by MET child participants in Experiments 6 and 7.

Overall, it can be observed from Figure 31 and from Table 15 and Table 18, that participants required more training across a larger number of novel stimulus sets to meet criterion for Opposite relational responding than these participants required to meet criterion in accordance with More-Than/Less-Than relational responding in Experiment 6 (see Table 15). In other words, it would appear that Opposite relational responding is relatively difficult to establish compared to More-Than/Less-Than relational responding. Nevertheless, despite slow acquisition rates, all participants demonstrated DRR in accordance with Opposite within less than the maximum number of exposures to novel stimulus sets (i.e., 23).

The data obtained from the control group of child participants who had previous exposure to non-MET More-Than/Less-Than relational responding in Experiment 6 were also examined. As always, the participants in the control group were only exposed to relational training and probe testing. In other words, participants were trained to criterion and then tested. They did not receive any exposure to multiple exemplar testing (i.e., testing with feedback) or remedial training and testing. The control groups' performances are detailed in Table 20. The data listed in Table 20 consist of the total number of blocks required to pass the Phase 1 pre-training, the total number of novel stimulus sets each participant was exposed to, the total number of blocks of Phase 2 Opposite relational training received and the total number of exposures to testing blocks required to pass the Phase 3 test for derivation of Opposite relations.

Table 20. The number of pre-training, training (including re-training), relational testing blocks and the number of novel stimulus sets required by each experienced child non-MET (control) participant to meet criterion for Opposite relational responding on each of the phases in Experiment 7.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (Opposite relational training)	Blocks of Opposite relational training	Blocks of Opposite relational testing
66	10	23	74	23*
67	7	23	58	23*
68	8	23	134	23*
69	7	20	45	20
70	6	19	43	19

*Note. These participants failed to meet criterion.

Table 20 shows the number of pre-training, training (including re-training), relational testing blocks and the number of novel stimulus sets required by each experienced child non-MET (control) participant to meet criterion for Opposite relational responding on each of the phases in Experiment 7. In considering Table 20, it is clear that as in previous experiments, all participants produced derived relational responding (DRR) in accordance with Opposite relations across a larger number of training and testing blocks as compared to producing DRR in accordance with the relations of More-Than/Less-Than. In fact, three of the participants did not meet criterion for Opposite relational responding at all. Of the participants in this group, P69 required 45 exposures to blocks of relational training and 20 exposures to blocks of relational testing across 20 novel stimulus sets to meet the pre-established criterion. P70 required 43 exposures to blocks of relational training and 19 exposures to blocks of relational testing across 19 novel stimulus sets to meet the pre-established criterion. The other three participants (P66, P67, P68) did not demonstrate DRR in accordance with Opposite despite exposure to 58 (P67), 74 (P66) and 134 (P68) relational training blocks across the maximum number of exposures to novel stimulus sets (i.e., twenty-three) across twenty-three exposures to relational tests.

By comparing the non-MET group (P66, P67, P68, P69 and P70) with the MET group (P60, P61, P62, P63 and P64), it can be seen once again, that the multiple exemplar intervention does seem to have been of benefit to the acquisition of DRR in accordance with Opposite. This can be seen by the lower number of blocks of relational testing required to meet criterion for Opposite relational responding for participants in the MET group compared to the non-MET group. Further support for the facilitative role of MET on the acquisition of Opposite is provided by the fact that three of the participants in the control (non-MET) group failed to demonstrate the generalisation of DRR in accordance with the relation of Opposite within the pre-determined maximum number of exposures to novel stimulus sets (i.e., twenty-three). Thus, it appears that the multiple exemplar intervention employed here for Opposite relational responding was indeed more successful in establishing a repertoire of derived relational responding than was exposure to training and unreinforced testing alone.

Despite the clear effectiveness of the MET intervention, it should be noted that control participants did display some level of learning across the non-reinforced testing blocks. That is, practice alone in the absence of feedback during testing did lead to small improvements in the relational repertoire. This was seen in the performance of the participants who did demonstrate criterion level relational responding in accordance with Opposite relations despite not having access to multiple exemplars. Nevertheless, having exposure to multiple exemplar testing and remedial levels (to a limited extent) benefited the MET group to a great extent, in that all participants were able to demonstrate DRR in accordance with Opposite within the designated number of novel stimulus set blocks (i.e., 23).

One question that was raised at the outset of Experiment 7 related to the extent to which reaching criterion for DRR for More-Than/Less-Than relations (in Experiment 6) played a role in facilitating reaching criterion for DRR in accordance with the relation of Opposite. In order to help answer this question, data from Experiment 5 (see Table 9 and Table 10) were re-considered. In Experiment 5, the acquisition of Same and Opposite relational responding was examined using two groups of naïve children. One group was an experimental group, and thus received the MET intervention. The other group was a control group, and therefore, had no access to the MET intervention. The Same and Opposite acquisition rates of these two groups from Experiment 5 will be compared (in terms of their mean cycles of acquisition of Opposite relational responding) to those of the two groups already described here in Experiment 7.

In Experiment 5, MET participants (P40, P41, P42, P43 and P44) required a mean of 18 exposures to relational training blocks, a mean of 5.4 exposures to relational testing blocks and a mean of 3.8 exposures to Opposite multiple exemplar testing blocks to meet the pre-established criterion for Opposite relational responding. Three participants (P40, P41 and P43) required exposure to remedial training and testing cycles (i.e., one block of each). By comparison, in Experiment 7, MET participants (with a history of More/Less MET intervention) P60, P61, P62, P63, P64 and P65 required a mean of 36.33 exposures to relational training blocks, a mean of 7.16 exposures to relational testing blocks, a mean of five exposures to blocks of Opposite multiple exemplar testing and a mean of three exposures to remedial training and testing cycles. Thus, these findings do not suggest that training in More-Than/Less-Than relational responding assisted the Experiment 7 MET participants with the acquisition of Opposite relational responding. In other words, reaching criterion level responding in Opposite relations does not appear to depend functionally upon having attained criterion for More-Than/Less-Than. In fact, it appears that MET participants who attained criterion level responding for More-Than/Less-Than relations, required higher numbers of cycles of Opposite relational training and testing, as well as higher number of cycles of Opposite MET testing and exposures to remedial levels. Of course, this observation is highly tentative at this point, given the already present variability in performances across a small number of participants. Nevertheless, this observation may still serve as a rough guide to developing a first-step intervention for those with educational or intellectual deficits. At least from the current data, there would appear to be no benefit in establishing More-Than/Less-Than relational responding and Opposite relational responding in any particular order. However, it must be remembered that on the whole, Opposite relational responding does seem to require significantly larger numbers of training and testing blocks in order to be established to criterion level. Thus, Opposite appears to be a more complex relational repertoire than More Than/Less-Than, even if the development of the repertoire is not functionally dependent on a well established More-Than/Less-Than relational repertoire.

In Experiment 5, non-MET participants (P45, P46, P47, P48 and P49) required a mean of 37.40 exposures to relational training blocks to meet training criterion. P45, P48 and P49 required a mean of a mean of 14 exposures to relational testing blocks to meet the pre-established criterion for Opposite relational responding. Two participants (P46 and P47) were unable to meet testing criterion after exposure to the maximum number of relational training and testing cycles (i.e., 23 cycles). These participants were excused from further participation in the examination of Opposite relational responding. In Experiment 7, the five non-MET participants with a history of non-MET More-Than/Less-Than relational responding, (P66, P67, P68, P69 and P70) required a mean of 70.80 relational training blocks. Only P69 and P70 met criterion for testing in accordance with the relation of Opposite. These participants required a mean of 19.50 exposures to relational testing blocks in order to demonstrate DRR in accordance with the relation of Opposite. Three non-MET participants, P66, P67 and P68, were unable to meet testing criterion for the demonstration of Opposite relational responding after the pre-determined maximum number of exposures to cycles of training and testing (i.e., 23 cycles). These participants were excused from further participation in the examination of Opposite relational responding. Once again, these findings do not suggest that reaching criterion level relational responding for More-Than/Less-Than relations through relational training and unreinforced testing assisted non-MET participants with the acquisition of Opposite relational responding. In fact, it appears that the two non-MET participants from Experiment 7 who attained criterion level responding for More-Than/Less-Than relations required higher numbers of cycles of relational training and testing than their counterparts who met criterion for the demonstration of DRR in accordance with the relation of Opposite from Experiment 5 (P45, P48 and P49). Thus, having exposure to relational training and unreinforced testing with the relations of More-Than/Less-Than also does not seem to have benefited participants in demonstrating DRR in accordance with the relation of Opposite. However, given that only 50% of this very small participant pool met criterion in accordance with the relation of Opposite at all, these data need to be interpreted with extreme caution. That being said, a functional dependence of Opposite relational responding on More-Than/Less-Than relational responding is far from apparent in the current data.

To summarise, Experiment 7 established two main findings, albeit with a relatively small participant sample. Firstly, Opposite relational responding was slower to emerge than More-Than/Less-Than relational responding. Secondly, this was the case even when More-Than/Less-Than relational responding had been at criterion level first. It would appear, therefore, that More-Than/Less-Than relational responding does not necessarily facilitate Opposite relational responding in any significant way. It does seem plausible to suggest that successful Opposite relational responding seems to be much more difficult to generate than relations of symmetry and transitivity (in Experiment 2) and More-Than/Less-Than relations (in Experiment 6). Thus, it seems that training Opposite relational responding last in an intervention containing the relations examined so far in this thesis would be more suitable than training it earlier in such a sequence. Therefore, despite the seeming functional independence of Opposite relational responding, following a series of interventions for other relations, it is speculated that a participant would benefit from acquiring the generic skills involved in learning relational tasks with more basic relations before being exposed to relations of increasing difficulty (i.e., Opposition). More specifically, if a participant has already been exposed to the intervention format for more easily acquired relations, it is likely that other skills (e.g., perhaps skills such as attentional skills) may also be increased and that previous interventions would serve to familiarise a participant with the stimulus types and task format, and also that their behaviour would already have come under the control of the conditioned reinforcement involved in feedback and the social reinforcement provided by the experimenter for participation. These factors would likely improve the effectiveness of the MET intervention for Opposite relational responding.

General Discussion

Chapter 4 set out to examine a new MET procedure with various groups of adults and children, some of whom had prior exposure to multiple exemplar training and some of whom did not. This latter fact also allowed the experimenter to examine the issue of a possibly optimal sequence of relational training for a preliminary applied intervention.

The results of Experiment 4 suggested that the multiple exemplar intervention employed was successful in establishing a repertoire of derived relational responding that was previously absent or weak for all of the adult participants. The introduction of the extra control tasks to preclude extraneous sources of control now ensures that appropriate contextual control was established during that experiment. The addition of the remedial level possibly assisted one participant in producing DRR for Opposite relations. However, Experiment 4 also noted that adults, on the whole, did not require a great deal of intervention in order to establish DRR with all relations including the relation of Opposite.

Experiment 5 employed identical procedures with ten normally functioning children. Results indicated that child participants in Experiment 5 required exposure to training and testing across a larger number of novel stimulus sets than did the adult participants in Experiment 4 for both Same and Opposite relational responding. However, the differences between Same and Opposite DRR acquisition were not quite as noticeable for adults in Experiment 4 as they were for child participants in Experiment 5. Specifically, while Same relational responding emerged rapidly for both adults and children across Experiments 4 and 5, adults required little, if any, intervention to demonstrate DRR in accordance with Same relational responding. In addition, all members of the non-MET group produced derived relational responding (DRR) in accordance with Same relations in fewer relational training and testing cycles compared to producing criterion level DRR in accordance with the relation of Opposite. Two of the non-MET participants did not meet criterion for Opposite relational responding at all. However, there were no noticeable differences in the acquisition rates of Same DRR for MET and non-MET groups noted. Thus, it appeared that the multiple exemplar intervention employed was indeed more successful in establishing a repertoire of derived relational responding for Opposite relations than was training and unreinforced testing alone. For these reasons, it can be said that exposure to multiple exemplar testing and the availability of remedial levels of intervention can help to facilitate the acquisition of DRR in accordance with Opposite relational responding within the designated number of novel stimulus set blocks (i.e., 23).

Experiment 6 examined the issue of the optimal sequence of relational training, which was first raised in Chapters 2 and 3. Specifically, this experiment examined the role of a history of DRR with one relation or set of relations (i.e., Same and Opposite) in acquiring DRR with other relations (e.g., More-Than/Less-Than). Acquisition rate differences across child and adult participants were also examined. The final issue examined was whether or not exposure to training and unreinforced testing for More-Than/Less-Than relations alone could produce similar rates of acquisition for both child and adult participants as could be produced with exposure to a multiple exemplar intervention.

The results of Experiment 6 showed that adult participants who had no prior history of DRR in accordance with Same and Opposite relational responding and who were not exposed to the multiple exemplar intervention demonstrated slower rates of acquisition than the adult participants who had previous exposure to interventions for other forms of DRR. Adults with no prior history of DRR in accordance with Same and Opposite relations who were not exposed to the multiple exemplar intervention also demonstrated slower rates of acquisition than the adult participants who had no previous exposure to DRR, but who were exposed to the multiple exemplar intervention. Therefore, it was suggested that having exposure to a history of Same and Opposite relational responding may not determine acquisition rates of More-Than/Less-Than relational responding for adult participants.

Experiment 6 also showed that child participants with a previous history of DRR with Same and Opposite relations required more cycles of training (including re-training), relational testing and multiple exemplar testing with More-Than/Less-Than relations than did adults with the same history. Child participants with no previous history of DRR in accordance with Same and Opposite relational responding attained similar rates of acquisition of More-Than/Less-Than relational responding to their counterparts who had a laboratory history of DRR in accordance with Same and Opposite relations. Thus, it may be possible to say that having a laboratory history of DRR for Same and Opposite relational responding is not necessary for the establishment of More-Than/Less-Than relational responding. Of course, it may be the case that having a minimal and low-rate level of Same and Opposite relational skills below that of the current criteria levels, may play some role in the acquisition of all other relations. It is not feasible for the current research to gain access to children with no Same and Opposite relational skills whatsoever (i.e., pre-language) in order to test this idea. Nevertheless, it is likely that the pre-existing sub-criteria level Same and Opposite relational skills of these participants did in some way facilitate the emergence of More-Than/Less-Than relational skills. However, this is difficult to say with confidence given the limited scope of this individual research programme. Indeed, child participants who had no prior laboratory history of DRR in accordance with Same and Opposite relational responding and who were not exposed to training and unreinforced testing for More-Than/Less-Than relations demonstrated slower rates of acquisition for More/Less relational responding than all adult and child participants. It should be noted that the differences between the child participant groups were small and given the small sample size, not appropriate for statistical analyses. Thus, it appears that once again, having access to a multiple exemplar training intervention is helpful in generating DRR for More-Than/Less-Than relations. However, even in the absence of any MET intervention, it appears that DRR for More-Than/Less-Than relations can emerge, albeit at a slower rate.

Given the foregoing conclusions, the experimenter asked whether or not the relations of Same and Opposite need to necessarily be trained first in an intervention aimed at increasing relational skills. It was suggested in the discussion of Experiment 6 that training in Same relational responding should come first in an intervention with an applied population. However, it is now apparent that More-Than/Less-Than relations can emerge without intervention for adults, and even for children, without a history of multiple exemplar training with other relations in the laboratory setting. Thus, More-Than/Less-Than now appear to be relatively fundamental relational skills that are at a high level of proficiency for most adults and children entering the laboratory. These relations require little intervention to raise the fluency of those skills to criterion levels. In contrast, in Experiment 5, two participants failed to demonstrated DRR in accordance with Opposite at all. This was a surprising outcome because Same and Opposite have been generally considered to be established earlier in life than More-Than/Less-Than relations are established. These data suggested the reverse. Thus, it was suggested that an effective intervention might train participants to first acquire DRR in accordance with Same relations, then to acquire DRR in accordance with More-Than/Less-Than relations and then finally to acquire DRR in accordance with Opposite.

The naïve child participants from Experiment 6 who were only exposed to More-Than/Less-Than relational responding were re-recruited to form the subject pool of Experiment 7. All participants required larger amounts of training across more novel stimulus sets to meet criterion for Opposite than these participants required to meet criterion in accordance with More-Than/Less-Than relations in Experiment 5. Three of the non-MET participants did not meet criterion for Opposite relational responding at all. However, all MET participants demonstrated DRR in accordance with Opposite within less than the maximum number of exposures to novel stimulus sets (i.e., 23).

Despite relatively slow acquisition rates for Opposite relational responding, the multiple exemplar intervention seemed to have been of benefit to the acquisition of DRR in accordance with Opposite. This is evidenced by the lower number of blocks of relational testing required to meet criterion for participants in the MET group and by the fact that three of the participants in the control (non-MET) group failed to demonstrate the generalisation of DRR in accordance with the relation of Opposite within the pre-determined maximum number of exposures to novel stimulus sets (i.e., twenty-three).

Experiment 7 failed to provide empirical support for the earlier suggestion that More-Than/Less-Than relations must be trained before Opposite relations in a DRR intervention. More specifically, no functional dependence was shown. Indeed Experiment 7 suggested that in terms of functional relationships, it may not matter which of these relations is trained first in a derived relational responding intervention. In effect, Experiments 6 and 7, taken together, have demonstrated that relations of More-Than/Less-Than and Opposition are functionally independent. However, across the experiments in the current and previous chapter, Opposite relational responding has been consistently relatively difficult to establish. This supports the view that the relation of Opposite is more complex than other relations examined in this research. Given that the demonstration of Opposite relational responding has presented as difficult for several participants, it may be best suited to the end of an applied intervention rather than being delivered at the outset. This is because exposure to interventions presented in a similar format would likely increase the generic skills (e.g., attentional skills) involved in learning relational tasks more generally. Participants would also become increasingly familiar with the stimulus types and task format before being exposed to the most challenging relational tasks (i.e., Opposite). In effect, their behaviour would already have come under the control of the conditioned reinforcement involved in feedback and the social reinforcment provided by the experimenter for participation. These factors should at least partially improve the effectiveness on the MET intervention for Opposite relational responding. Given that this particular form of relational responding has been shown to be somewhat difficult to acquire for both children and adults across Chapters 3 and 4, it is speculated that introducing relations in increasing complexity throughout the intervention format may serve to maximise learning and reduce the aversive functions of repeated failure on a given training and testing block. It is hoped that such a strategy might also serve to reduce a possible high drop-out rate from the intervention. Drop-outs might result from presenting very demanding relational tasks to a population of educationally challenged children before copious reinforcement has been delivered for on-going participation and by their mastery of less challenging forms of relational responding than Opposite.

The current findings do not yield conclusive answers to the question of whether or not relations of Opposition functionally depend on the other trained relations which have shown to emerge more readily. In fact, the current data set suggests that the relations of More-Than/Less-Than and Opposition are functionally independent. Indeed, it may be important to note that unspecified relations were not tested for with the More-Than/Less-Than protocol. To give an example, the correct choice is unspecified during a task that leads to no answer. In effect, a test for the full range of possible more and less derived relations may in fact require larger number of examplars than were required here (see Whelan et al., 2006). Testing such derived relational responses is perhaps a superior test of the frames of More-Than and Less-Than. The absence of such probes in the current experiment, therefore, may account for why establishing fluency with Opposite relations was more difficult to establish than for comparative relations.

These are important basic research questions. However, as the current research had an applied goal as an end point, it was necessary to at least test the current intervention protocol first and allow the results of an intervention to inform further questions, rather than get trapped in a quagmire of purely basic research matters regarding a possible perfect MET training sequence. In any case, it is important to remember that the relationships between the various relations are not fixed in stone. The behavioural RFT view is not an organicist one (see Hayes & Brownstein, 1986) but a contextualistic one (see Hayes, 1993b). In other words, RFT does not see intellectual development as unfolding in a pre-determined direction for all humans, as in the Piagetian scheme. Rather, development acquires a unique trajectory for each human, depending on the entire gamut of biological and personal historical variables, including the effect of current context. Thus, RFT would place little emphasis on identifying the "correct" sequence in any case, but would emphasise the importance of finding a workable and maximally effective sequence, which may in turn vary across situations, and participant groups, depending on the particular deficits and relevant biological variables. Moreover, RFT would recognise that the relations between the various relational repertoires may not be a simple linear one. For instance, exhaustive research may demonstrate that Opposite relational responding is functionally dependent on relations of More-Than/Less-Than and Same, or on some other relation not yet studied in the current research. In particular, it is likely that the relation of Opposite is a specific instance of the difference relation. To explain this point, difference includes opposite (i.e., all opposites are different to samples, but not all different stimuli are opposite to a sample). This was first clarified in Roche and Barnes (1996a) where it was noted that in one sense, difference can be viewed as mere non-equivalence responding. However, Roche and Barnes (1996a) and Barnes and Roche (1996) refute this point on empirical grounds. Roche and Barnes (1996a) demonstrated that difference requires more complex forms of relational contextual control than non-equivalence (S-). As argued in both papers, opposite requires an even finer form of contextual control, and thus likely follows the emergence of difference relational responding in young children. What was not known at that time was precisely how much more training and testing might be required to establish opposite relational responding, as opposed to difference responding. The current data suggest anecdotally that while many children might display fluent difference relational repertoires, it may take considerably more multiple exemplar training in the natural environment for Opposite relational responding to emerge.

One further issue worth considering relates to the seemingly paradoxical militating effect of histories of More-Than/Less-Than responding on the acquisition of Opposite relational repertoires. It would appear that in many cases, such a history was associated with slower rates of acquisition of Opposite relational responding rather than faster rates. One possible explanation for this outcome points to the increasing complexity of the contextual control required to parse the various forms of relational responding into tight categories (i.e., functional classes). Put simply, as experimenters attempt to establish additional forms of contextual control with a given participant pool, the opportunity for confounded forms of stimulus control increases. For example, it is likely that when first presented with relational pre-training and training for Opposite relations, many participants with a laboratory history of More-Than/Less-Than responding may have responded in early trials in terms of More-Than/Less-Than relations, rather than demonstrating control by the Opposite contextual cue. Surely, over several trials the new form of control would be established. Nevertheless, an important part of the process of establishing conditionally discriminative control over responding is to “wash out” other forms of control across several trials (Sidman, 1994). This process takes time and becomes more challenging with each form of contextual control being established. This is because more and more established forms of control must be “washed out” upon each attempt to establish an additional form of contextual control. This process, while based on speculation, may help to explain the paradoxical effect of More-Than/Less-Than histories on the acquisition rates of Opposite relational responding repertoires.

The foregoing issues represent important basic research questions that require addressing in further research. However, in keeping with the pragmatic research tradition of RFT, it was decided to not pursue these basic questions at this point in the research. Rather it was decided to settle upon what appeared to be a “best fit” intervention for the current purposes. Given time constraints, this was done in order to at least make some positive impact on the lives of the participants as soon as possible. This method could then allow the data acquired to drive further empirical questions, as is typical in the inductive tradition in which behaviour analysis resides (Sidman, 1960).

Before a “best fit” intervention is delivered in an applied setting, one final issue needs to be examined. Specifically, it is not yet known whether or not a relational training intervention, as it has now been devised, would have any benefit to either a normally functioning or to an applied population. Thus, it was determined that it was important to test the general effectiveness of the intervention using participants selected from the general population. To achieve this, in Chapter 5, participants from Chapter 2, who had already taken part in training and testing for DRR in accordance with symmetry, transitivity (Experiment 2), Same and Opposite relational responding (identical to that described in Experiment 3), were re-recruited. The experiment was designed to establish whether or not the extensive history of multiple exemplar training provided to these participants thus far would lead to a measurable increase in IQ scores. First, however, it was necessary to provide these participants with a More-Than/Less-Than MET intervention in order that they would have been exposed to a full battery of Same, Opposite, More-Than/Less-Than relations, albeit with less effective procedures (i.e., conditional discrimination) than had been selected for use in the applied setting (i.e., relational evaluation.). Thus, in Experiment 8, MET participants from Experiment 2, Chapter 2 were administered the amended More-Than/Less-Than relational intervention from the current chapter. In Experiment 9, MET and non-MET participants from Experiment 2, Chapter 2, were re-administered IQ tests.

Chapter 5

A Preliminary Relational Training Battery and Its Effect on IQ

Experiments 8-9

Introduction

Chapter 4 examined a new procedure for generating DRR with Same and Opposite (Experiments 4 and 5) and More-Than/Less-Than (Experiments 6 and 7) relations, which served as an alternative to the standard Matching-to-Sample procedure. This new procedure employed a novel combination of the Relational Evaluation Procedure and the Yes/No Procedures, as described in Chapter 4. Chapter 4 also examined the issue of a possibly optimal sequence of relational training. More specifically, some participants who had a history of prior exposure to Same and Opposite relational responding (Experiment 6) were exposed to training and testing for More-Than/Less-Than relational responding. Other participants were exposed to More-Than/Less-Than relational responding first (Experiment 6) and to Opposite relational responding second (Experiment 7). The results of these experiments indicated that the novel procedure employed was successful in establishing a repertoire of derived relational responding that was previously absent or weak for all of the adult and child participants involved. Members of the control groups/non-MET groups who were exposed to training and unreinforced testing only (in Experiments 5, 6 and 7) performed more weakly than their counterparts in the MET groups, which showed that some form of multiple exemplar training may be necessary in order to establish DRR with more complex relations (e.g., Opposition).

It was also noted in Chapter 4 that adults, on the whole, did not require a great deal of intervention in order to establish DRR in accordance with relations of Same, Opposite and More-Than/Less-Than. Thus, it appeared that having exposure to multiple exemplar training for Same and Opposite and More-Than/Less-Than relational responding may not be as helpful to adult participants as it is for child participants in simply establishing relational repertoires.

Perhaps not surprisingly, child participants who had no prior history of DRR in accordance with Same and Opposite relational responding, and who were not exposed to the multiple exemplar intervention, demonstrated weaker levels of acquisition for More-Than/Less-Than relational responding than all other adult and child participants. However, the differences between the child participant groups were small for the acquisition of More-Than/Less-Than relations. A surprising outcome from Chapter 4 was that it appeared that More-Than/Less-Than relational responding could emerge without intervention for adults and to a lesser extent for children, even without a history of multiple exemplar training with other relations in the laboratory setting.

In addressing the issue of finding an optimal sequence of training in a relational intervention in Chapter 4, it was suggested that more basic relations (i.e., Same, More-than/Less-Than) may act as building blocks for apparently more complex relations (i.e., Opposite). However, Chapter 4 failed to demonstrate conclusively that training more easily acquired relations first would facilitate the speed of acquisition for less easily acquired forms of relational responding. In other words, it may not matter whether More-Than/Less-Than relations or Opposite relations are trained first in an applied intervention in terms of the rate of acquisition. Nevertheless, the findings from Chapters 3 and 4 together support the notion that the relations of Same, More-Than/Less-Than and Opposite range from basic to most complex respectively. It therefore makes intuitive sense to train relations in that order. This issue is revisited in the next chapter.

One exciting issue, which has not been examined thus far, relates to whether or not a more extensive history of multiple exemplar training, from symmetry and transitivity, Same and Opposite relational responding, (i.e., the formerly supposed "basic" forms of relational responding) together form a sort of behavioural basis for the acquisition of further relations of various kinds. In other words, it is not yet known if a basic behavioural repertoire involving multiple forms of DRR might be conducive to the acquisition of more complex relations, such as More-Than, Less-Than, hierarchy, et cetera. In other words, while relational repertoires for Same, More-Than, Less-Than and Opposite seem to have little functional dependence on each other, it may be the case that families of these relations may form functional units that impact upon the acquisition of novel relational repertoires. Of course, there is no way of knowing how such foundational repertoires might be composed in the natural environment. However, the research, thus far, has led to the establishment of such repertoires with several groups of participants. For instance, P1, P2, P3 and P8 have been trained to criterion levels of fluency in symmetry and transitivity (Experiment 2). While data was not reported, these participants were also trained to criterion on Same and Opposite relational responding in Chapter 3 using the same procedure as that reported in Experiment 3. The data gathered will now be outlined because doing so will provide us with an opportunity to study the effect of this extended repertoire on More-Than/Less-Than relational responding. In addition, establishing the relational repertoire of More-Than/Less-Than with this group provides an opportunity to assess for the first time, the impact of a longitudinal extended MET training intervention on IQ.

Given the foregoing, Chapter 5 addressed the issue of whether or not the new procedure for More-Than/Less-Than was successful in establishing a repertoire of More-Than/Less-Than relations in a group of children who had an extended history of multiple exemplar training over a long period of time (Experiment 8). This also served to provide these participants with the entire range of relations examined in this thesis thus far. This allowed for a subsequent assessment of the effect of the entire intervention on IQ (Experiment 9). Experimental participants from Experiment 2, Chapter 2 (P1, P2, P3, P8), were re-recruited to form the participant pool for Experiment 9. (The reader is reminded that these participants also had exposure tp an intervention which yielded criterion level Same and Opposite relational responding using the procedures described in Chapter 3). Thus, four normally functioning children with extensive histories of multiple exemplar training in derived relational responding in accordance with symmetry, transitivity, Same and Opposite made up the participant pool for Experiment 8. These participants were exposed to the identical sequence of More-Than/Less-Than multiple exemplar training and testing as described in Experiment 6, Chapter 4. Experimental and control participants from Experiment 2, Chapter 2 (P1, P2, P3, P4, P6, P7, P8, P9) were re-recruited to form the participant pool for Experiment 9. These participants were all administered IQ tests as in Experiments 1 and 2 of Chapter 2.

Method

Participants

All participants (P1, P2, P3 and P8) were re-recruited from Experiment 2. Participants had a mean age of 12.79 years. Participants' exact ages are listed in Table 21.

Table 21. Participants' ages at start of the Experiment 8.

Participant ID	Age
P1	14 years, 9 months
P2	12 years, 6 months
P3	12 years, 1 months
P8	11 years, 10 months

As noted in Experiments 1 and 2, of this thesis, these participants were recruited from friends and family of the researcher and from a local school. All participants had been identified by parents and teachers as students who were not presenting with any known or suspected learning difficulties.

Setting and Materials

All setting and material information was identical to that listed in Experiment 3.

Ethics

All ethical information was identical to that listed in Chapters 2, 3 and 4.

General Experimental Sequence

The experimental sequence for the More-Than/Less-Than relational training and testing protocol was identical to the sequence described in Experiment 6, Chapter 4. As all participants maintained group membership in the MET group, all participants were permitted access to multiple exemplar testing and remedial levels as needed. Thus, all participants received; Phase 1, pre-training to establish the arbitrary cues for relational responding in accordance with More-Than and Less-Than (see Figure 32), Phase 2 administration of a series of relational training tasks (see Figure 33) via a lap top computer, Phase 3, the administration of non-MET and MET tests for derived relational responding in accordance with relations of More-Than/Less-Than (see Figure 34) and finally, Phase 4, remedial training and testing. See Figure 35 for a schematic of all training and testing tasks used.

Figure 32 below shows the tasks used in Phase 1, pre-training to establish the arbitrary cues for relational responding in accordance with More-Than and Less-Than.

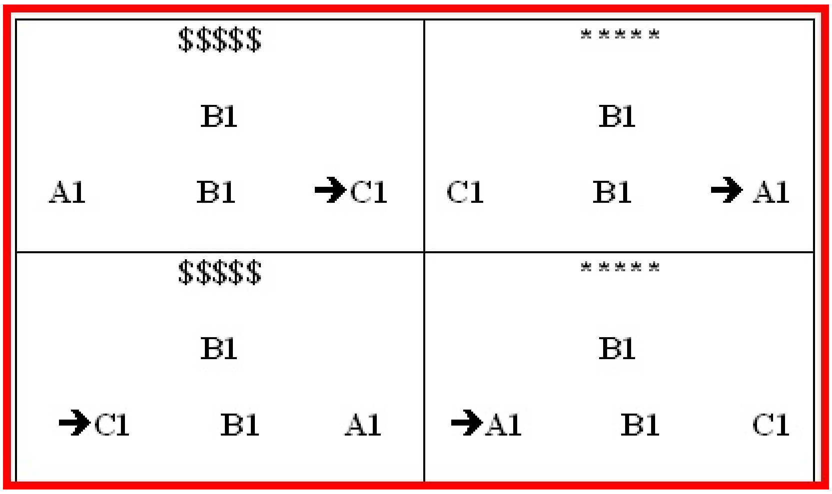


Figure 32. Tasks presented during More-Than/Less-Than relational pre-training in Phase 1. The alphanumerics represent non-arbitrary stimulus sets (see Appendix 21). The $$$$$ and ***** represent the cues for More-Than and Less-Than, respectively, which were established during this phase. An arrow indicates the correct choice in the presence (training phases) or absence (testing phases) of feedback.

Figure 33 below shows the tasks employed during Phase 2, More-Than/Less-Than relational training.

A1 $$$$$ B1 →Yes No	B1 $$$$$ C1 →Yes No	C1 $$$$$ D1 →Yes No
A1 ***** B1 Yes → No	N1 $$$$$ N2 Yes →No	N1 ***** N2 →Yes No

Figure 33. Tasks presented during More-Than/Less-Than relational training in Phase 2. The alphanumerics represent the nonsense syllables employed as stimuli. The $$$$$ and ***** represent the cues for More-Than and Less-Than, respectively which had been established during the Phase 1 Pre-training. An arrow indicates the correct choice in the presence of feedback.

Figure 34 below shows the tasks used in Phase 3, More-Than/Less-Than relational testing (non-MET and MET).

D1 $$$$$ A1 →Yes No	C1 $$$$$ A1 →Yes No
D1 ***** A1 Yes →No	C1 ***** A2 Yes →No

Figure 34. Tasks presented during More-Than/Less-Than relational testing in Phase 3. The alphanumerics represent the nonsense syllables employed as stimuli. The $$$$$ and ***** represent the cues for More-Than and Less-Than respectively which were established during the Phase 1 Pre-training. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Figure 35 shows a schematic of the training and testing tasks employed in Experiment 8. Pre-training and remedial phases are not shown here.

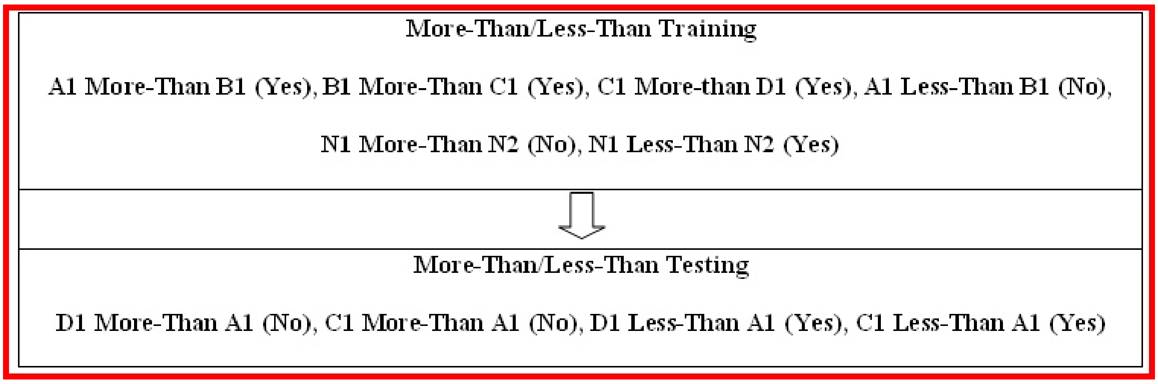


Figure 35. Schematic of the training and testing tasks employed during two of the four phases of Experiment 8 (where “Yes” and “No” in parentheses indicate correct responses). The remedial level (Phase 4) is identical to More-Than/Less-Than training and testing phases, with the exception that instead of using nonsense words (as represented by alphanumerics in the schematic), non-arbitrary stimuli (images of cats, dogs, houses, et cetera.) are used. These can be seen in more detail in Appendix 21.

Results and Discussion

The data listed in this section consist of the total number of blocks required to pass Phase 1 pre-training, Phase 2 More-Than/Less-Than relational training, the total number of blocks to pass the Phase 3 test for derivation of More-Than/Less-Than relations, and the total number of blocks required to pass the Phase 4 remedial training and testing, where applicable. This MET group had prior exposure to DRR training and testing in accordance with symmetry and transitivity (Experiment 2, Chapter 2) and DRR in accordance with Same and Opposite relations identical to that provided to participants in Experiment 3, Chapter 3.

Table 21. The number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each experienced child MET (experimental) participant to meet criterion for More-Than/Less-Than relations on each of the phases of Experiment 8.

Participant No.	Blocks of relational pre-training	Novel stimulus sets (More/Less relational training)	Blocks of More/Less relational training	Blocks of More/Less relational testing	Blocks of More/Less Multiple Exemplar Testing	Blocks of More/Less remedial training	Blocks of More/Less remedial testing
1	6	5	13	3	2	0	0
2	8	3	9	2	1	0	0
3	4	3	11	2	1	0	0
8	6	3	9	2	1	0	0
Mean	6	3.50	10.50	2.25	1.25	0	0

Note. Each child in this group also had a prior history of DRR with symmetry, transitivity, Same and Opposite relational responding.

Child participants with a previous history of DRR with symmetrical, transitive, Same and Opposite relations required few cycles of training (including re-training), relational testing and multiple exemplar testing overall. No participant from this group required any exposure to remedial levels of training and testing.

All participants reached criterion for More-Than/Less-Than relational responding within a low number of training and testing cycles (see Table 21). It is clear that the current procedure was successful in rapidly establishing the repertoire of More-Than/Less-Than relational responding for the participants who took part. The rates of acquisition cannot be directly compared to other child participant groups in Chapter 4 due to the fact that the actual time to acquisition was not measured. In addition, different training histories may also mean that direct comparisons are not meaningful. However, anecdotally, these participants were noted to attain criterion very quickly. This perhaps suggests that the whole package trained was of benefit to these participants.

There are many reasons why the participants in Experiment 8 may have acquired the skill of More-Than/Less-Than relational responding so quickly. Of course, it is simply not possible to tell from the current data set if it was access to the symmetry and transitivity MET intervention, or to the Same and Opposite MET intervention, in particular that generated these results. An extended study would be required to address this question. However, at the very least, it is speculated that the additional time spent responding under the contingencies presented by the MET intervention across multiple phases and sessions likely played some role. Disentangling the effect of exposure to the procedure itself, rather than the relational skills established, would require a separate controlled experiment in which control participants were exposed to an amount of similar, but non-MET, intervention relational training and testing, yoked to the average amount of contact with the training and testing protocols experienced by the experimental participants.

It is also possible, as was observed during the emergence of transitive (Chapter 2) and Same relational responding (Chapters 3 and 4), that the skill of deriving More-Than/Less-Than relations was already reasonably well developed in the repertoires of these participants or that this skill was a more basic level skill than others examined (e.g., Opposite) in the previous chapters. Participants in Experiment 8 (P1, P2, P3 and P8), cannot be directly compared to participants in Chapter 4, as the participants in Experiment 8 were exposed to a mixture of procedures, and were also exposed to symmetry and transitivity much earlier. In addition, the experimental participants in Experiment 8 (P1, P2, P3 and P8), were also more experienced at training and testing procedures due to involvement over a longer period of time than other participants involved in the other experiments in this thesis. However, making some casual observations may be useful.

Given the foregoing, if the reader is to look back to Experiment 6, Chapter 4, it can be seen that the randomly selected MET child participant groups described (P40, P41, P42, P43, P44, P60, P61, P62, P63, P64, P65 and P66) tended to require more relational training, and in many instances, more relational testing to acquire criterion level More-Than/Less-Than responding than did the longitudinal MET group (P1, P2, P3 and P8) in Experiment 8. In addition, the random MET child participant groups described in Experiment 6, Chapter 4 (P40, P41, P42, P43, P44, P60, P61, P62, P63, P64, P65 and P66) also tended to require more relational pre-training and two participants (P42 and P65) even required exposure to the remedial levels of training and testing. Thus, it appears that the extended repertoire of DRR did help acquisition of More-Than/Less-Than relational responding to some extent. It is therefore possible to speculate that families or groups of relational skills together may lead to increasingly faster rates of acquisition of novel relational repertoires. As discussed above, this cannot be said with any certainty, and the particular groups of relational repertoires that may be optimal for training early in life in order to hasten acquisition of further relations is impossible to speculate on at this point. But the in-principle observation can be made that longer interventions with given relations may improve acquisition of novel relations. However, given that all groups (including non-MET groups) demonstrated criterion level More-Than/Less-Than relational responding with relative ease, group differences are small and difficult to interpret.

These issues notwithstanding, the MET procedure for generating DRR in accordance with More-Than/Less-Than relations has now been tested across several experiments and has been shown to reliably lead to the rapid emergence of More-Than/Less-Than relational repertoires. In addition, with a small cohort of participants now trained on the entire battery of relations examined thus far, and across multiple sessions and across a large period of time, we may now turn our attention to the effect of this extended intervention on IQ. Experiment 9 examined how exposure to an entire intervention, albeit over a long period of time, and using various procedures, can demonstrate an in-principle effect on IQ scores.

To this end, participants from Experiment 8 were re-recruited and administered the WISC-III^{UK}. (The reader is reminded that these participants also formed part of the participant pool for Experiments 1 and 2 in Chapter 2, and that these participants were also administered the identical procedures as in Experiment 3, Chapter 3). The control participants from Experiment 2, Chapter 2 were also re-recruited to act as a comparison group. (The reader is reminded that these participants also formed part of the participant pool for Experiment 1, Chapter 2).

Experiment 9

The results from Experiment 1, Chapter 2, showed correlations between basic DRR skills and IQ subtest scores. It was also pointed out at that time that these correlations did not constitute proof of functional dependence of DRR skills on IQ or the converse. It was noted that it seemed likely that a myriad of variables are at work in determining both IQ and DRR skill levels, even though from a RFT perspective, these two variables still likely bear an important functional relationship to each other. In effect, it was suggested that DRR was likely a necessary but not sufficient variable for determining IQ. Indeed findings from Experiment 2, Chapter 2, are consistent with findings by O'Toole and Barnes-Holmes (in press), as described in Chapter 1 of this thesis, whereby correlations were found between measures of relational flexibility (Implicit Relational Assessment Procedure or IRAP) and IQ scores for adults, as measured by the Kaufman Brief Intelligence Test. In that body of work, O'Toole et al. found that faster responding on the IRAP predicted higher IQ. In Experiment 2, Chapter 2, of the current research, the experimenter attempted to move beyond correlational research by checking if IQ scores could actually be raised by manipulating relational skill levels. More specifically, a brief intervention was conducted to examine whether or not an increase in DRR skill would lead to an increase in full scale IQ score or its various component subtest and scale scores. In the introduction to Experiment 2, Chapter 2, it was predicted that the presence of such a relational intervention would in turn lead to differences in these various IQ measures. These predictions were correct in that the presence of an intervention seems to have had some measurable effect on overall IQ scores. Those rises, while impressive, were modest and would benefit from replication. Experiment 9 will provide an opportunity to see if an extended intervention has an even more dramatic effect on IQ than the minimal intervention employed across Experiments 1 and 2.

Method

Participants

Participants P1, P2, P3 and P8 were re-recruited from Experiment 8 in the current chapter. Participants P4, P6, P7 and P9 were re-recruited from the non-MET/control group from Experiment 2, Chapter 2. As noted in Experiments 1, 2 and 8 of this thesis, children were recruited from friends and family of the researcher, and from local schools, and also on having been identified by parents and teachers as students who were not presenting with any known or suspected learning difficulties. Participants had a mean age of 12.8 years. Participants' individual exact ages at the beginning of Experiment 9 are listed in Table 22.

Participant ID	Age
P1	14 years, 10 months
P2	12 years, 7 months
P3	12 years, 2 months
P4	11 years, 1 month
P6	12 years, 8 months
P7	12 years, 6 months
P8	11 years, 11 months

Table 22. Participants' ages at start of Experiment 9.

Setting and Materials

All participants were administered the Wechsler Intelligence Scale for Children, Third Edition, UK (WISC-III^{UK}) in their own home. (See the Setting and Materials section of Chapter 2 for a detailed description of the WISC-III^{UK}). In each instance, a quiet private room was used to minimise distractibility. As always, times for sessions were chosen based on when there were least likely to be external distractions (such as the presence of other family members) in the house. A room was also chosen in each house based on the presence of a large table where each child could sit and work comfortably for the duration of the IQ test administration. The experimenter sat opposing the participant. No one else was present in the room, although in every instance, the participants' parent(s) were either in the next room or in another room in the house. Parents were also informed that they should feel free to come in and out of the room as they chose. As mentioned in previous experiments, no parent did this.

Ethics

Information regarding ethical considerations was identical to that described in Chapters 2, 3 and 4.

General Experimental Sequence

IQ testing in every instance took place in one session. This session took approximately two hours for each participant, including time for set-up, clear-up and any breaks requested by participants. Each IQ test was administered on an individual basis, as is standard procedure. A session consisted of the administration of the entire battery of the WISC-III^{UK}. The session started with the experimenter asking the participant if they were willing to take part in the research. If they responded, “yes” (e.g., that they were willing), then the experimenter proceeded to administer the WISC-III^{UK}. Every participant agreed to continue. The experimenter then proceeded to ask a series of questions about everyday situations, about general word usage, about basic computation and about abstract concepts, among other items. The experimenter also asked the participant to complete several timed tasks involving the manipulation of concrete objects (i.e., blocks, picture cards, jigsaw puzzle pieces, et cetera). See Wechsler (1992) for more information about the procedures for administration of the WISC-III^{UK}, as well as actual subtest items. (See Appendix 3 for a description of each subtest). As is standard procedure, each child was permitted to take as many breaks as needed. However, it should be noted that no child took extended breaks. Specifically, breaks were only taken for bathroom and drinks.

Results and Discussion

Experimental participants’ individual subtest and scale scores at Time 1, Time 2 (from Experiment 1 and 2 of Chapter 2) and Time 3 (Experiment 9, current chapter) are listed in this section in Table 23. All three IQ test administrations have been listed for the purposes of comparison. Control participants’ individual subtest and scale scores at Time 1, Time 2 (from Experiment 1 and 2 of Chapter 2) and Time 3 (Experiment 9, current chapter) are listed in this section in Table 24. (See Appendix 27 for all individual participant subtest and scale scores at Time 3).

Table 23. IQ Subtest scores, Full scale IQ (FSIQ) scores, Verbal IQ (VIQ) scores, Performance IQ (PIQ) scores, for each experimental participant (P1, P2, P3 and P8) at Time 1, Time 2 and Time 3.

Experimental participant (P1).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Score at Time 3	Difference from Time 1 to Time 3
Picture Completion				**10**	**10**	**13**	**+3**
Information	**12**	**12**	**16**				**+4**
Coding				**11**	**12**	**12**	**+1**
Similarities	**12**	**11**	**15**				**+3**
Picture Arrangement				**8**	**12**	**19**	**+11**
Arithmetic	**12**	**8**	**12**				**0**
Block Design				**9**	**11**	**14**	**+5**
Vocabulary	**9**	**10**	**11**				**+2**
Object Assembly				**7**	**10**	**13**	**+6**
Comprehension	**8**	**11**	**14**				**+6**
(Symbol Search)				**(7)**	**(13)**	**(16)**	**+9**
(Digit Span)	**(10)**	**(11)**	**(9)**				**-1**
Sum of Scaled Scores (Verbal)	**53**	**52**	**68**				**+15**
Sum of Scaled Scores (Perf)				**45**	**55**	**71**	**+26**
Verbal IQ	**103**	**102**	**122**				**+19**
Performance IQ				**92**	**107**	**132**	**+40**
Full Scale IQ	**98**	**105**	**131**				**+33**

Note. IQ scores at Time 1 (in black font), Time 2 (in blue font) and Time 3 (in red font). Differences in each score from Time 1 to Time 3 are listed in green.

Experimental participant (P2).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance. Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				**15**	**12**	**17**	**+2**
Information	**14**	**16**	**15**				**+1**
Coding				**10**	**12**	**12**	**+2**
Similarities	**13**	**15**	**18**				**+5**
Picture Arrangement				**14**	**16**	**18**	**+4**
Arithmetic	**12**	**10**	**17**				**+5**
Block Design				**10**	**8**	**12**	**+2**
Vocabulary	**12**	**12**	**13**				**+1**
Object Assembly				**11**	**10**	**12**	**+1**
Comprehension	**15**	**13**	**19**				**+4**
(Symbol Search)				**(12)**	**(11)**	**(14)**	**+2**
(Digit Span)	**(13)**	**(11)**	**(17)**				**+3**
Sum of Scaled Scores	**66**	**66**	**82**				**+16**
Sum of Scaled Scores				**60**	**58**	**71**	**+11**
Verbal IQ	**120**	**120**	**139**				**+19**
Performance IQ				**115**	**112**	**132**	**+17**
Full Scale IQ	**119**	**118**	**137**				**+18**

Experimental participant (P3).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				**13**	**14**	**16**	**+3**
Information	**8**	**9**	**15**				**+7**
Coding				**9**	**10**	**11**	**+1**
Similarities	**12**	**12**	**15**				**+3**
Picture Arrangement				**10**	**11**	**19**	**+9**
Arithmetic	**13**	**10**	**16**				**+3**
Block Design				**9**	**11**	**10**	**+1**
Vocabulary	**13**	**13**	**14**				**+1**
Object Assembly				**11**	**12**	**13**	**+2**
Comprehension	**15**	**14**	**19**				**+4**
(Symbol Search)				**(8)**	**(11)**	**(11)**	**+3**
(Digit Span)	**(10)**	**(10)**	**(9)**				**-1**
Sum of Scaled Scores (Verbal)	**61**	**58**	**79**				**+18**
Sum of Scaled Scores (Perf)				**52**	**58**	**69**	**+17**
Verbal IQ	**113**	**109**	**136**				**+23**
Performance IQ				**103**	**112**	**130**	**+27**
Full Scale IQ	**109**	**111**	**135**				**+26**

Experimental participant (P8).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				**7**	**7**	**11**	**+4**
Information	**7**	**9**	**12**				**+5**
Coding				**9**	**14**	**13**	**+4**
Similarities	**13**	**12**	**11**				**-2**
Picture Arrangement				**13**	**19**	**18**	**+5**
Arithmetic	**10**	**8**	**12**				**+2**
Block Design				**6**	**7**	**9**	**+3**
Vocabulary	**9**	**9**	**9**				**0**
Object Assembly				**9**	**13**	**12**	**+3**
Comprehension	**12**	**12**	**16**				**+4**
(Symbol Search)				**(12)**	**(18)**	**(13)**	**+1**
(Digit Span)	**(12)**	**(12)**	**(11)**				**-1**
Sum of Scaled Scores (Verbal)	**51**	**50**	**60**				**+9**
Sum of Scaled Scores (Perf)				**54**	**60**	**76**	**+22**
Verbal IQ	**101**	**100**	**111**				**+10**
Performance IQ				**91**	**115**	**137**	**+46**
Full Scale IQ	**96**	**107**	**128**				**+32**

The successive panels of Table 23 show the experimental participants' IQ subtest and scale scores at Time 1, Time 2 and Time 3. Table 24 shows the control participants' IQ subtest and scale scores at Time 1, Time 2 and Time 3. As in Table 23, in Table 24 Time 1 scores are shown in black, Time 2 scores are shown in blue and Time 3 scores are shown in red. Differences in each score from Time 1 to Time 3 are listed in green.

Table 24. IQ Subtest scores, Full scale IQ scores, Verbal IQ scores, Performance IQ scores, for each control participant (P4, P6, P7 and P9) at Time 1, Time 2 and Time 3.

Control participant (P4).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				**8**	**8**	**6**	**-2**
Information	**11**	**12**	**14**				**+3**
Coding				**9**	**10**	**14**	**+5**
Similarities	**12**	**9**	**9**				**-3**
Picture Arrangement				**11**	**16**	**12**	**+1**
Arithmetic	**13**	**9**	**9**				**-4**
Block Design				**8**	**9**	**8**	**0**
Vocabulary	**10**	**7**	**7**				**-3**
Object Assembly				**10**	**10**	**10**	**0**
Comprehension	**14**	**11**	**10**				**-4**
(Symbol Search)				**(11)**	**(5)**	**(11)**	**0**
(Digit Span)	**(13)**	**(14)**	**(11)**				**-2**
Sum of Scaled Scores (Verbal)	**60**	**48**	**49**				**-11**
Sum of Scaled Scores (Perf)				**46**	**53**	**50**	**+4**
Verbal IQ	**111**	**98**	**99**				**-12**
Performance IQ				**94**	**104**	**99**	**+5**
Full Scale IQ	**104**	**101**	**99**				**-5**

Note. IQ scores at Time 1 (in black font), Time 2 (in blue font) and Time 3 (in red font). Differences in each score from Time 1 to Time 3 are listed in green.

Control participant (P6).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				10	13	8	-2
Information	14	15	16				+2
Coding				14	14	15	+1
Similarities	10	11	13				+3
Picture Arrangement				6	10	9	+3
Arithmetic	10	11	13				+3
Block Design				14	8	8	-6
Vocabulary	13	12	12				-1
Object Assembly				11	10	10	-1
Comprehension	13	11	8				-5
(Symbol Search)				(12)	(13)	(9)	-3
(Digit Span)	(10)	(10)	(14)				+4
Sum of Scaled Scores	60	60	62				+2
Sum of Scaled Scores				55	55	50	-5
Verbal IQ	111	111	115				+4
Performance IQ				107	107	99	-8
Full Scale IQ	111	111	108				-3

Control participant (P7).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				**11**	**11**	**9**	**-2**
Information	**10**	**11**	**13**				**+3**
Coding				**13**	**10**	**10**	**-3**
Similarities	**9**	**12**	**11**				**+2**
Picture Arrangement				**12**	**16**	**15**	**-3**
Arithmetic	**10**	**13**	**9**				**-1**
Block Design				**9**	**10**	**12**	**+4**
Vocabulary	**11**	**10**	**12**				**+1**
Object Assembly				**10**	**12**	**10**	**0**
Comprehension	**11**	**11**	**8**				**-3**
(Symbol Search)				**(10)**	**(11)**	**(11)**	**+1**
(Digit Span)	**(10)**	**(10)**	**(9)**				**-1**
Sum of Scaled Scores (Verbal)	**51**	**57**	**53**				**+2**
Sum of Scaled Scores (Perf)				**55**	**59**	**56**	**+1**
Verbal IQ	**101**	**108**	**103**				**+2**
Performance IQ				**107**	**113**	**109**	**+2**
Full Scale IQ	**104**	**111**	**106**				**+2**

Control participant (P9).

Subtests	Verbal Scaled Scores at Time 1	Verbal Scaled Scores at Time 2	Verbal Scaled Scores at Time 3	Performance Scaled Scores at Time 1	Performance Scaled Scores at Time 2	Performance Scaled Scores at Time 3	Difference from Time 1 to Time 3
Picture Completion				**10**	**10**	**8**	**-2**
Information	**15**	**15**	**15**				**0**
Coding				**14**	**16**	**12**	**-2**
Similarities	**9**	**9**	**12**				**+3**
Picture Arrangement				**10**	**7**	**7**	**-3**
Arithmetic	**13**	**13**	**14**				**+1**
Block Design				**9**	**6**	**7**	**-2**
Vocabulary	**12**	**13**	**13**				**+1**
Object Assembly				**9**	**9**	**8**	**-1**
Comprehension	**10**	**11**	**10**				**0**
(Symbol Search)				**(15)**	**(14)**	**(6)**	**-9**
(Digit Span)	**(7)**	**(10)**	**(8)**				**+1**
Sum of Scaled Scores (Verbal)	**59**	**61**	**64**				**+5**
Sum of Scaled Scores (Perf)				**52**	**48**	**42**	**-10**
Verbal IQ	**110**	**113**	**117**				**+7**
Performance IQ				**103**	**96**	**88**	**-15**
Full Scale IQ	**107**	**106**	**104**				**-3**

The successive panels of Table 24 show the control participants' IQ subtest and scale scores at Time 1, Time 2 and Time 3.

It was expected that there would be mild fluctuation of subtest scores from Time 1 to Time 3 across participants and between the two groups. For this reason, no small fluctuations (i.e., any change of four points or less) have been discussed here. For our purposes, we have considered changes of more than four scaled points to be large fluctuations. Therefore, changes of four points or more on any individual subtest have been considered briefly. Full scale IQ, verbal IQ and performance IQ are discussed separately.

The experimental participants are considered first. Experimental participant, P1, presented with the most inter-subtest change from Time 1 to Time 3. She presented with changes of four points or more in six subtests (i.e., Information, Picture Arrangement, Block

Design, Object Assembly, Comprehension, and Symbol Search). All of these changes were in a positive direction (i.e., rises). Experimental participant, P8, presented with the next highest amount of inter-subtest change. P8 presented with changes of four points or more in five subtests (i.e., Picture Completion, Information, Coding, Picture Arrangement and Comprehension). These changes were also all in a positive direction. Experimental participant, P2, presented with the next highest amount of inter-subtest change. P2 presented with changes in four subtests (i.e., Similarities, Picture Arrangement, Arithmetic and Comprehension). Once again, these changes were all in a positive direction. Experimental participant, P3, presented with the least amount of inter-subtest change. P3 presented with positive changes in three subtests (i.e., Information, Picture Arrangement and Comprehension). It is interesting to note that all four experimental participants presented with rises in the Picture Arrangement subtest and the Comprehension subtest. It is also of interest to point out that three out of four of these participants presented with rises in the Information subtest.

The control participants are considered next. Control participants, P4 and P6, both presented with large inter-subtest change in three subtests. P4 presented with positive change in the Coding subtest and negative changes in Arithmetic and in Comprehension. P6 presented with negative change in Block Design and Comprehension and with positive change in Digit Span. P7 and P9 presented with large change in only one subtest. P7 presented with a positive change in Block Design. P9 presented with a negative change in Symbol Search. Commonalities between the inter-subtest change are present for P4 and P6 in that they both presented with a negative change in Comprehension score. P6 and P7 both presented with changes in Block Design score. However, one of these changes was in a positive direction and one was in a negative direction.

In the WISC-III^{UK}, all inter-subtest scores contribute to the either the verbal IQ scale or the performance IQ scale. All subtests together contribute to the full scale IQ score. As noted earlier in this chapter (and in Chapter 1 of this thesis), a certain amount of inter-subtest fluctuation is expected. Small variations from one test take to another are not noteworthy. The larger changes that have been seen here, particularly for the experimental group, can be said to be unexpected. How these various inter-subtest changes impact on verbal IQ, performance IQ and full scale IQ is considered next in Table 25.

Table 25. All participants' Full Scale IQ scores, Verbal IQ scores and Performance IQ scores at Time 1, Time 2 and Time 3.

P#	Group	Time 1 FSIQ	Time 2 FSIQ	Time 3 FSIQ	Time 1 VIQ	Time 2 VIQ	Time 3 VIQ	Time 1 PIQ	Time 2 PIQ	Time 3 PIQ
1	EXP	98	105	131	103	102	122	92	107	132
2	EXP	119	118	137	120	120	139	115	112	132
3	EXP	109	111	135	113	109	136	103	112	130
8	EXP	96	107	128	101	100	111	91	115	137
4	CT	104	101	99	111	98	99	94	104	99
6	CT	111	111	108	111	111	115	107	107	99
7	CT	104	111	106	101	108	103	107	113	109
9	CT	107	106	104	110	113	117	103	96	88

Note. Each participant in this table is listed by their participant number. "EXP" and "CT" are used to indicate whether a participant was a member of the experimental or the control group. IQ scores at Time 1, Time 2 and Time 3 are shown in black, blue and red respectively. (The individual subtest scores which contributed to these overall scores for each participant can be viewed in Appendix 26 and 27).

Overall, experimental participants' IQ scores tended to rise from Time 1 to Time 3. Experimental participant, P1, presented with a full scale IQ score rise of 33 full scale IQ points, comprised of a rise of 19 verbal IQ points, and a rise of 40 performance IQ points from Time 1 to Time 3. Experimental participant, P2, presented with a full scale IQ score rise of 18 full scale IQ points from Time 1 to Time 3, comprised of a rise of 19 verbal IQ points, and 17 performance IQ points. Experimental participant, P3, presented with a full scale IQ rise of 26 full scale IQ points, comprised of a rise of 23 verbal IQ points, and 27 performance IQ points (from Time 1 to Time 3). Experimental participant, P8, presented with a full scale IQ rise of 32 full scale IQ points (from Time 1 to Time 3), comprised of a rise of 10 verbal IQ points, and 46 performance IQ points.

Overall, control participants' IQ scores did not present with the same trend of large rises as the experimental participants did. Control participant, P4, presented with a full scale IQ drop of five points from Time 1 to Time 3. This was comprised of a drop of 11 verbal IQ points, and a rise of five performance IQ points. Control participant, P6, presented with full scale IQ drop of three points from Time 1 to Time 3. This was comprised of a rise of four verbal IQ points, and a fall of eight performance IQ points. Control participant, P7, presented with a full scale IQ rise of two points from Time 1 to Time 3. This was comprised of a two point rise of verbal IQ score, and a two point rise in performance IQ. Control participant, P9, presented with a full scale IQ fall of three points from Time 1 to Time 3. This was comprised of a seven point rise in verbal IQ, and a 15 point drop in performance IQ score.

Means for the differences between groups in full scale IQ score, verbal IQ score and performance IQ score from Time 1 to Time 3 have been calculated and can be seen in Table 26. The experimental group showed a mean rise of 27.25 full scale IQ points from Time 1 to Time 3. The control group showed a mean fall of 2.25 full scale IQ points from Time 1 to Time 3. The experimental group showed a mean rise of 17.75 verbal IQ points, and the control group showed a mean rise of .25 verbal IQ points from Time 1 to Time 3. The experimental group showed a mean performance IQ rise of 32.50 points and the control group showed a mean performance IQ fall of 4.50 points.

Table 26. Mean Full Scale IQ, Verbal IQ and Performance IQ changes for each group from Time 1 to Time 3.

Group	FSIQ	VIQ	PIQ
Experimental	+27.25	+17.75	+32.50
Control	-2.25	+.25	-4.50

While the number of participants employed here was relatively small, the findings from Experiment 9 provide further evidence that rises in IQ scores were generated using the DRR intervention. The reader is reminded that this conclusion was also drawn in Experiment 2, Chapter 2. Thus, this is the second occasion on which this observation has been made using a RFT-based MET intervention with the same group of participants. In effect, the current findings, while gathered using only eight participants, would appear more robust when considered in the light of previous findings. An inferential analysis is not appropriate using this small data set. However, simple observations of the probabilities of these outcomes occurring by chance are worth considering. As an example, let us leave aside the very impressive absolute rises in IQ observed for experimental over control participants. Considering only increases and decreases in IQ in binary form, and given that IQ scores are subject to natural fluctuations within known ranges, the odds of observing an increase in FSIQ, as opposed to a non-increase from Time 1 to Time 3 for any one participant should be 50% (i.e., 1/2). Therefore, the chances of observing four rises in FSIQ from Time 1 to Time 3 for all four experimental participants while also observing four non-increases in FSIQ for all four control participants, precisely as predicted, was 1/2 X 1/2 X 1/2 X 1/2 X 1/2 X 1/2 X 1/2 X 1/2 = 0.004, or 4 chances in a thousand. Therefore, the odds of observing four rises in FSIQ from Time 1 to Time 3 for all four experimental participants, as well as the three non-increases in FSIQ for control participants, actually observed in this experiment (see Table 25) was 1/2 X 1/2 X 1/2 X 1/2 X 1/2 X 1/2 X 1/2 = 0.008, or 8 chances in a thousand. Of course, these observations and many similar ones that can be made are rather anecdotal. Nevertheless, the fact that the current data set points so obviously towards significant effects on IQ of the current longitudinal relational training intervention strongly supports the central tenets of this thesis. Perhaps more importantly, the large and significant IQ rises which were displayed by the experimental group have important implications for clinical populations in need of interventions of this type.

One potential criticism of the data gathered in Experiment 9 is that the procedures which apparently led to the IQ rises seen were varying. However, the reason for this variation in protocols was that the total intervention for experimental participants, P1, P2, P3, and P8 was developed across time and across several experiments. Thus, the longitudinal participants necessarily received varying training and testing formats, as well as symmetry and transitivity, as the training and testing procedures were improved across time. Of course, in a real intervention with an applied population, it would be unwise and inefficient to vary the training and testing protocols across phases. For this reason, the protocols will be streamlined for an applied population (see Chapter 6). Nevertheless, this current study gives us a picture of the types of IQ rises that might be possible given the development of these repertoires. In other words, in following research functionally similar repertoires will be created for an applied population with special educational needs even if the actual training and testing formats differ somewhat from the MTS procedure and the procedure employed for Same and Opposite and More-Than/Less-Than relational training and testing with the current group.

General Discussion

The findings from Experiment 8 showed that the procedure for More-Than/Less-Than relational responding was successful in generating a repertoire of those relations or at least in increasing an already present repertoire of More-Than/Less-Than relations for participants (P1, P2, P3 and P8). Thus, it can now be said that this procedure was successfully employed across both adults (Experiment 6) and children (Experiment 6) with varying levels of exposure to DRR and varying histories of exposure to relational interventions (Chapters 4 and 5).

In Experiment 9, experimental participants (P1, P2, P3 and P8) were re-recruited. These participants now had exposure to a relational intervention which consisted of an entire battery of relations (i.e., symmetry, transitivity, Same, Opposite and More-Than/Less-Than). Control participants (P4, P6, P7 and P9) were also re-recruited to serve as longitudinal comparisons to the experimental or intervention group. The administration of the WISC-III^{UK} to both participant groups revealed that participants in the experimental group showed large FSIQ rises from Time 1 (Experiment 2, Chapter 2) to Time 3 (Experiment 9, current chapter), and that similar rises were not seen for the control group from Time 1 (Experiment 2, Chapter 2) to Time 3 (Experiment 9, current chapter).

The procedures developed and tested across Experiments 3, 4, 5, 6, 7 8 and 9 (in Chapters 3, 4 and 5) have gone some way towards providing a challenging intervention that appears to be increasing relational skills levels beyond those observed during baseline phases. However, a relatively high skill level pre-intervention has been observed for several of the relations examined (i.e., Same relational responding in Chapters 3 and 4 and More-Than/Less-Than relational responding across many adult and child participants in Chapter 4). Thus, a high skill level at baseline remains an issue. That is, high pre-experimental fluency in DRR narrows improvements in DRR across experimental and control groups and makes the effectiveness of DRR interventions more difficult to observe. However, in contrast to Same and More-Than/Less-Than relational responding, all participants, regardless of age, required intervention for Opposite relations. In other words, Opposite relational responding did not seem to be at a high level of fluency at baseline for either adults or children across Experiments 3, 4, 5, 6 and 7. In fact, all participants required some level of MET to meet criterion for Opposite relational responding. Several control participants across Experiments

5 and 7 (e.g., P30, P31, P32, P33, P46, P47, P65, P66 and P67), who did not have access to MET, failed to meet criterion for DRR in accordance with the relation of Opposite.

Moreover, the relational intervention employed seems to have affected substantial IQ rises, as observed in Experiment 9. Thus, it has been shown that an extensive history of DRR, in accordance with various relations over an extended period of time, can have a significant impact on full scale IQ score for normally functioning children. Individual subtests which may contribute importantly to these rises may be the Picture Arrangement subtest, the Comprehension subtest, and the Information Subtest. Interestingly, correlations were seen for all three of these subtests with the combined measure of stimulus equivalence in Experiment 1, Chapter 2. These correlations were not necessarily the strongest correlations seen in Experiment 1. Despite that, it is interesting to note that these subtests, at least at a cursory glance, appear to remain important in terms of their relationship to relational skill in that they both correlated with DRR in accordance with very basic level relational skills (symmetry and transitivity combined), and in that exposure to more extensive relational interventions seems to have had an impact on these subtests for normally developing child participants.

The recorded rises, in what has often been called a psychological invariant, should be of great interest to educational psychologists, psychometricians, and specialist educators working in a wide variety of applied areas. While clinically relevant rises in IQ were neither expected nor required in this investigative research endeavour, they were nevertheless observed for experimental participants in many cases. That is, small rises were witnessed for all members of the experimental group from Time 1 to Time 2 (in Experiment 2, Chapter 2) and only for one member of the control group (Experiment 2, Chapter 2). In Experiment 9, large rises in full scale IQ were observed for all members of the experimental group, but not for the control group.

For the purposes of the current research, a clinically relevant rise in full scale IQ score may be considered a rise that would move a participant out of his or her qualitative range. Many examiners utilise a qualitative system, as well as a quantitative system, to describe a child's performance (see Wechsler, 1992, 1991). Table 27 presents specific IQ score ranges and their corresponding qualitative diagnostic categories using the WISC-III^{UK}. This table has been adapted from Wechsler (1992, 1991).

Table 27. Quantitative full scale IQ scores and their corresponding qualitative diagnostic classifications for the WISC-III^{UK}.

Quantitative IQ Score	Qualitative Diagnostic Classification
130 and above	**Exceptionally high**
120-129	**High**
110-119	**High Average**
90-109	**Average**
80-89	**Low Average**
70-79	**Low**
69 and below	**Exceptionally Low**

In Table 27, the quantitative IQ score ranges and the corresponding qualitative classifications for each IQ score range, using the WISC-III^{UK}, are shown. (These ranges are also used on earlier and later versions of the Wechsler scales for both adults and children). According to Wechsler (1992, 1991, p. 33), about 95% of all children obtain scores within the 70-130 range. Thus, only 5% of children can be said to score in the exceptionally high or the exceptionally low qualitative range. Of the remaining 95% of children, the ranges that children generally score within are made up of ten points. This is excepting the average score of 100, which includes scores from 90 to 109 within that average range. Each of the distributions of the verbal, performance and full scale IQ scores has a mean of 100 and a standard deviation of 15 (Wechsler, 1992, 1991, p. 33).

Of course, it is recognised that children can score anywhere within each range in order to be classified within a qualitative diagnostic range. So a child who attained a full scale IQ score of 108 on a first IQ test and a 110 on a second IQ test would not be noteworthy, even though this child would have technically "moved" from the average to the high average range. What would be noteworthy, however, would be a child who attained a score of 108 on a first IQ test and a 118 on a second IQ test, as this is a larger rise than would typically be seen from test to re-test administrations (see Wechsler, 1992, 1991, p. 63; see also discussion in Experiment 2, Chapter 2; see also Juliano, Haddad, & Carroll, 1988 for discussions on practice effects).

The rises seen in Experiment 9 of the current chapter are noteworthy because participants showed large IQ rises, and also moved out of their diagnostic qualitative ranges. Two of the experimental participants (P1 and P3) "moved" from the qualitative average range to the qualitative exceptionally high range. One participant (P2) moved from the qualitative high average range to the qualitative exceptionally high range. Finally, one participant (P8) moved from the qualitative average range to the qualitative high range. These moves are much higher than would typically be seen in test-re-test situations. Control group participants (P4, P7 and P9) stayed within the average range from first test administration to third test administration. Control group participant (P6) moved from the high average range to the average range from first test administration to third test administration. (This change was only comprised of three full scale IQ points, but nevertheless, a change in qualitative range). Comparing the two groups makes the rises seen in the experimental group's IQ scores even more noteworthy. Thus, at this point in the research, it is clear that by utilising a DRR intervention, clinically relevant IQ rises can be demonstrated with normally functioning children. What remains unclear is whether or not these IQ rises can be demonstrated with a population of participants of sub-normal IQ or sub-normal ability.

The data presented in the current chapter lend further support to a relational intervention being of value to a population with intellectual deficit. At a full scale IQ level, it appears that normally functioning participants would benefit from this type of intervention. It also appears that the relational intervention has had a stronger impact on performance IQ than on verbal IQ. This is consistent with the stability data for the WISC-III^{UK} (Wechsler, 1992, 1991), which shows that discrepancies in scores due to practice effects are higher for performance IQ scores than for verbal IQ scores. Perhaps part of the reason for such an effect is related to practice with the format of test taking. Thus, it is possible that, as suggested earlier in this chapter, part of the reason the intervention protocol used in this research has been so successful is that in addition to increasing relational skills, it may be increasing the attentional skills of the participant and increasing control over the format by the experimenter using social reinforcement. As the subtests which contribute to the performance IQ are thought to measure an individual's "perceptual organization", it is not surprising that when the level of control over the testing environment is increased that such testing skills might also increase. The subtests in the performance domain are all timed subtests, so it is perhaps not surprising that the increase in speculated test taking skills, such as attention, would yield rises in overall performance IQ. However, it is important to note that verbal IQ scores also rose across all four participants which may provide further support for the relational nature of the subtests involved.

It is important to note that both the performance and verbal domains have been impacted to a large degree. At the inter-subtest level, large positive changes were seen for ten of the twelve subtests administered. All four experimental participants presented with large positive change in the Comprehension subtest (verbal domain) and in the Picture

Arrangement subtest (performance domain). Three out of four experimental participants presented with positive change in the Information subtest (verbal domain). This suggests that children that score weakly on these subtests may benefit from a relational intervention such as this. Of course, any subtests not rising to such a large degree may be targeted in other types of interventions for various other skills important in the classroom environment (i.e., task completion, motivation, et cetera), or perhaps a relational intervention involving relations that were not included in this intervention so far. Although the overall full scale IQ rises seem to be the most impressive, the individual subtest scores may be more important from a functional point of view.

The foregoing theoretical functional analysis of the relationship between the intervention employed thus far, and IQ, may point us in the direction of important relationships that can be examined more carefully in experimentation. However, the main focus of the current research remains to test such an intervention at a macro level with a population in need, rather than to tease these relationships apart purely for conceptual argument. Thus, while many important questions remain, they can be pursued in future research. At this point in the current research, it is clear to see that the relational intervention developed so far has the potential to have impact on full scale IQ, as well the many component parts which contribute to a full scale IQ score. Thus, we are now armed with a "best fit" intervention that can be brought to an applied population that are known to be struggling in the mainstream academic classroom.

It is hoped that the effectiveness of these new procedures will in fact be greater than those seen for the intervention employed across Experiment 2, Chapter 2, Experiment 3, Chapter 3, and Experiment 8 of the current chapter, insofar as they are expected to establish greater fluency of relational responding across larger number of stimulus sets with an applied population. Some further modifications will have to be made, in line with suggestions raised by the research thus far. That is, the matter of the need for a time-based fluency criterion will also be addressed for the applied intervention. This is to maximise the effectiveness of the intervention beyond what might be expected from the results of Experiments 8 and 9.

Chapter 6

Examining the Effect of a Same, More -Than/Less-Than and Opposite Relational Responding Multiple Exemplar Intervention on Full Scale IQ

Experiments 10-11

Introduction

In Chapter 5, Experiment 9, it was seen that the entire developed intervention for equivalence, Same, Opposite, and More-Than/Less-Than was successful, in principle, in effecting substantial IQ rises for an experimental longitudinal group of normally functioning children (i.e., P1, P2, P3 and P8), who were trained over an extended period of time. Similar IQ rises were not seen for a matched control group (P4, P6, P7 and P9), who had no access to the developed intervention, but who had been exposed to repeated training and testing for equivalence relations without multiple exemplars. The question remained whether or not this developed intervention would have similar effects with a population of children experiencing learning difficulties within a mainstream Irish primary school.

Before the developed intervention was administered to children experiencing learning difficulties, one outstanding issue needed to be addressed. This issue relates to a measurement of baseline levels of skills being taken at the outset of interventions. While previous experiments in this thesis have generally shown improved relational skills in a series of domains, they have not attempted to explicitly employ a traditional behavioural A-B-A-type design. Such a design generally involves a "three-phase experimental design consisting of an initial baseline phase (A) until steady state responding is obtained, an intervention phase in which the treatment condition (B) is implemented until the behavior has changed and steady state responding is obtained, and a return to baseline conditions (A) by withdrawing the independent variable to see whether responding "reverses" to levels observed in the initial baseline phase" (Cooper, Heron, & Heward, 2007, p. 689). Including the reintroduction of the independent variable in an A-B-A-B design is often preferred because reintroducing the B conditions enables the replications of treatment effects, which can, in turn, strengthen the demonstration of experimental control (see Cooper, et al., 2007, p. 177). In the current research, there were many reasons for not employing the traditional A-B-A or A-B-A-B designs up to this point. However, the two most important reasons related to practicalities and to ethics. More specifically, in the current research, agreements had been made with a local primary school to engage a certain number of their students in the research throughout the course of a school year. In return for this student participation, the researcher agreed to provide free comprehensive psychometric assessments for these students, and psycho-educational reports for the school, and for each child’s parents. Included in the assessment process were various consultations with teachers, parents and any other involved professional (e.g., speech therapists, occupational therapists, the child guidance clinic) where appropriate. Thus, given the intensive and extensive nature of both the research itself, and the assessment process, it was estimated early on that full A-B-A designs would not be feasible given the time and labour constraints involved in conducting this research. It was also deemed that once relational interventions had been administered, they were likely to lead to relatively permanent changes in relational skills. Thus, a return to the baseline state may not have even been possible due to the irreversibility of behavioural effects. With regard to ethical considerations, it was deemed unethical by the experimenter to withdraw interventions over an extended period in an attempt to measure the deterioration of recorded improvements in the absence of multiple exemplar contingencies. This consideration also impeded the use of an A-B-A or similar research design. Consequently, the only appropriate single participant design may be described as A-B-C, where C refers to a baseline state different to that at the pre-intervention time.

Despite the foregoing, it was important to reconsider the baseline measures employed during early experiments. In particular, the assessments of relational skills pre intervention, required exposure to the very contingencies to be employed during the experimental phase. For instance, in order to assess a participant's skill levels in deriving equivalence relations, it was necessary in Experiments 1 and 2 to expose the participant to equivalence training and testing to criterion. The intervention then consisted of an attempt to improve the fluency of this relational skill. However, the baseline measurement was found in itself to radically impact upon the skill. Thus, what was required was an independent measure at the outset, and completion of an intervention, that would impact to as little an extent as possible on the skills under analysis. Such a procedure should also allow for a rapid assessment of pre and post intervention relational skills in contrast to the laborious process used in Experiments 1 and 2. The first experiment of this chapter, therefore, focused on developing a baseline relational abilities index.

In employing a population with special needs, it is especially important to take objective baseline measures that can be used to fully assess the impact of an intervention. While the current research does not constitute a traditional applied intervention, and is still exploratory and experimental in nature, it nevertheless would speak more clearly to researchers in the applied field if a simple baseline measure and a clear A-B-A type design was employed using a baseline. Furthermore, employing baseline measures may, in principle, allow the experimenters to determine that a particular participant may not be in need of a relational intervention or may have deficits in particular domains only, rather than across the range of relational skills being targeted by the intervention.

Given the foregoing, the relational abilities measure or index used in the current chapter will now be outlined. For the sake of simplicity, this measure will be referred to as the relational abilities index, or with the acronym, RAI. The RAI consisted of four separate baseline measures of Same, More-Than, Less-Than and Opposite relational responding. Each measure consisted of three sets of twenty presentations of stimuli (arbitrary nonsense syllables, see Appendix 28) in the presence of an English phrase which functioned as a pre-experimentally established contextual cue. In effect, these presentations functioned like arbitrary statements which could be read from left to right. Participants were required to respond “yes” or “no” to each ‘statement’ by clicking on the relevant word on the screen. More specifically, following detailed instructions (see Appendix 29), a participant was presented with twenty statements and questions such as; “A is MORE THAN B. B is MORE THAN C. Is A MORE THAN C?”. The participant was then required to click on either “yes” or “no” to indicate their choice. No feedback was provided following this response and another similar statement using different nonsense syllables was then presented. After twenty such presentations of statements and questions, the order of the statements was reversed. An example of a reversal statement is; “B is MORE THAN C. A is MORE THAN B. Is C MORE THAN A?”. Following twenty such presentations of reversed statements and questions, the participant was then presented with twenty new statements and questions exactly like the reversed sample, but with a strict response time criterion on each trial. Identical protocols were followed for Same, Opposite and Less-Than relations. In this manner, the experimenter was able to obtain a baseline skill level for Same, More-Than, Less-Than and Opposite without lengthy pre-training cycles to establish arbitrary contextual cues and relational networks. Also, given that each participant was only presented with a total of 240 statements each, and that no feedback was given, these baseline skill level measurements could be taken quickly and without interfering with the intervention process.

One final issue remained to be addressed prior to commencing the recruitment and employment of population with learning difficulties. This issue was that the intervention developed across this thesis so far had necessarily only been delivered to a normally functioning population. Therefore, it was as yet unknown whether or not a different population would have difficulty meeting criterion on the individual measures of relational responding. It was also unknown whether or not a population of children with learning difficulties would require more sessions to reach criterion than their normally developing peers. Thus, two changes were made to the developed intervention. The first change was that, as RFT views the relation of Same as the functional equivalent of equivalence relations (as discussed in Chapter 4), participants were only administered an intervention for derived relational responding in accordance with Same relations (and not for equivalence), More-Than/Less-Than relations and Opposite relations. This served as a time saving measure. However, it also allowed for the use of a single training and testing format across the intervention given that the procedures had been modified since the administration of the equivalence relational responding intervention. It was speculated that this would in turn reduce the learning demands placed on participants to reach criterion on each phase.

The second change to the intervention related to the expected lower acquisition rates that might be observed among the population of interest. Recall that a remedial level was added to the relational intervention procedure in Experiments 4, 5, 6, 7 and 8. The reason for adding this remedial level (as described in detail in Chapter 4) was that some normally functioning participants taking part in previous experiments had struggled to meet criterion for Opposite relational responding. Berens and Hayes (2007) suggested that one possible way of facilitating the derivation of arbitrary relational responding when it is absent or weak would be to make use of non-arbitrary stimuli as well as arbitrary stimuli in training and testing protocols (see also Y. Barnes-Holmes, D. Barnes-Holmes, Roche, & Smeets, 2001; Steele, & Hayes, 1991). The remedial level described in Chapter 4, and employed across all experiments in Chapter 4 and in Experiment 8, Chapter 5, seemed to be helpful in assisting normally functioning participants to reach criterion for the various types of relational responding involved. However, even given exposure to the remedial levels after exposure to seven cycles of relational training, four cycles of relational testing and three cycles of multiple exemplar testing, some participants in Chapter 4 still required more than one exposure to the remedial levels of intervention. Thus, it might be expected that a population presenting with learning difficulties might require an increased amount of exposure to remedial levels of training and testing. For this reason, participants with learning difficulties, who took part in the experiments in the current chapter, were permitted extra access to remedial levels if needed. Specifically, if a participant did not meet criterion on the first probe test following the standard exposure to seven cycles of relational training, four cycles of relational testing, three cycles of multiple exemplar testing, and one exposure to remedial training and testing, the training and testing procedure was modified slightly. More specifically, if a participant did not reach criterion on the first probe test after a remedial level, the participant was exposed to more relational training, a multiple exemplar test and finally, a new remedial level. The protocol then proceeded in this manner (i.e., relational training, probe test, more relational training, multiple exemplar test, remedial training and remedial testing), until a participant met criterion on the first probe test with a novel stimulus set following a remedial level. It was predicted that this would reduce any possible punishing effects of participation on learning if a participant persistently failed to meet criterion.

In Experiment 10a, the RAI (Relational Abilities Index) was first piloted with three normally developing twelve-year-olds (P71, P72 and P73) from a local primary school. These participants were described by their teacher as “average range, sixth class students, who did not present with any known or suspected learning difficulties”. Following the successful completion of the RAI by the three normally developing participants, a population of children presenting with various learning difficulties was recruited for Experiment 10b. This population included eight sixth class students (P74, P75, P76, P77, P78, P79, P80 and P81) from the same local primary school as the children who took part in Experiment 10a. All members of the population group were identified as struggling to maintain progress with the mainstream curriculum. The participants in the educationally challenged group were first administered a standard psycho-educational assessment that followed all National Educational Psychological Services (NEPS) guidelines. This assessment included the Wechsler Intelligence Scale for Children, Fourth UK edition (WISC-IV^{UK}), as well as a battery of other assessment tools, which are not discussed in this thesis, as these other tools were necessary for diagnostic purposes, but were not related to the current research. Following administration of the IQ test, these eight participants were administered the RAI. Spearman's Rho correlational analyses were then carried out between subtests and indices of the WISC-IV^{UK} and scores of relational ability on the RAI. Experiment 11 then proceeded with the administration of the intervention for Same, More-Than/Less-Than and Opposite relational responding. Following the intervention, follow-up IQ tests and RAI tests were administered to assess any changes that may have occurred across the intervention.

Experiment 10a

Method

Participants

Three normally developing 6th class students were recruited from a local primary school to pilot test the RAI. These participants (P71, P72 and P73) were all twelve years of age and were identified by both parents and teachers as "not presenting with any known or suspected learning difficulties". Therefore, these participants were not experiencing any difficulties with the mainstream curriculum, and were not attending learning support (i.e., remedial) classes for any subject. These participants were identified by teachers as having scored within the average or above average range on standardised tests administered within the school setting. These participants had not taken part in any previous research of this type. As such, all participants were considered "naïve".

Setting and Materials

Each participant took part in a classroom set aside by the primary school for the research. As in previous experiments, times for sessions were chosen based on when this classroom was available. The RAI was administered to each child individually via Macintosh™ *ibook* lap top computer. The RAI was controlled by software written by the author using *Psyscope*. The stimuli were composed of 120 three letter nonsense syllables in a consonant-vowel-consonant sequence. All nonsense syllables used for the RAI are listed in Appendix 28.

Ethics

Experiments 1 and 2 of this thesis received ethical approval by the NUIM Ethics Committee. As the format of Experiment 10a was similar to Experiments 1 and 2, ethical approval was not sought specifically for Experiment 10a. However, at all times, ethical guidelines established by the Psychological Society of Ireland and the British Psychological Society were observed. In addition, as in previous and subsequent experiments in this thesis, consent was obtained from all participants in verbal and written form.

General Experimental Sequence

The experiment took place over the course of one ninety-minute session per participant. Research commenced at the convenience of the school and the classroom teacher and upon the signing of consent forms. Participants were also comprehensively debriefed on the exact nature of the procedure and purpose of the experiment when all participants had completed the research. Participants were not ever informed as to the precise details of the RAI testing protocol while they were engaged in the experiment.

This was done in order to prevent confounding of the measure by participants discussing possible strategies for increasing their scores on the RAI in between one student's completion of the RAI and the next student's commencement of the RAI. Sessions in this experiment consisted of four phases in total. Phase 1 included a RAI for Same relational responding. Phase 2 consisted of a RAI for More-Than relational responding. Phase 3 consisted of a RAI for Less-Than relational responding, and Phase 4 consisted of a RAI for Opposite relational responding. Overall, this experiment served as a pilot test of the relational abilities index which was to be used throughout the rest of this chapter if no difficulties presented.

Phase 1: RAI for Same Relational Responding

Level 1

Following detailed instructions (see Appendix 29a) a participant was presented with twenty statements and questions. Initially, one statement such as, "A is the SAME AS B" appeared on the screen in black letters. The alphanumerics represent nonsense syllables chosen randomly from a list of possible A and B nonsense syllables. The first statement remained on the screen until the end of the trial. One second after the first statement appeared on the screen, a second statement such as, "B is the SAME AS C", in black letters appeared on the screen. This also remained on screen until the end of the trial. After another one second interval, a question such as, "Is A the SAME AS C?" appeared in red on the screen. On an equal number of trials the question "Is A the OPPOSITE OF C?" was presented. This served to ensure that the participant could not attain a score of 20 out of 20 by merely responding "Yes" to every question, and thus they were required to attend to both the contextual cue words and the relevant stimuli. At the same time as the question appeared, “No” and “Yes” also appeared in blue letters on the lower left and right hand of the screen. The two statements, the question and the possible response choices (i.e., “Yes” and “No”) remained on the screen until such time as the participant had indicated a choice by clicking on it with the mouse. No feedback was provided following the participant’s response.

Level 2

Following the presentation of 20 Level 1 trials, a set of 20 further trials were presented in which the order of the statements presented was reversed. That is, participants could not read the statements sequentially to derive a relation. Rather, the relation was more easily derived if the participant first read the statement that was presented second. An example of a reversed presentation is; “B is the SAME AS C. A is the SAME AS B. Is C the SAME AS A?”. The presentation format was identical to that for Level 1.

Level 3

Immediately following Level 2, the participant was presented with new instructions (see Appendix 29b) and twenty new trials identical to Level 2 trials (but with novel stimuli). However, a strict response window of seven seconds was employed on each trial. If the participant failed to respond within seven seconds, the non-response was recorded as an incorrect response and the next statement was presented.

Across all sixty trials of Levels 1, 2 and 3 combined, the “Yes” and “No” response key positions were counterbalanced across trials. None of the nonsense syllables used as stimuli was employed across more than one trial.

Phase 2: RAI for More-Than Relational Responding

The RAI for More-Than relational responding proceeded in an identical manner to Phase 1, with the exception that the statements included the words “More-Than” instead of “the Same as”. In addition, the questions presented included the words “More-Than” instead of “the Same as” and “Less-Than” instead of “Opposite of”.

Phase 3: RAI for Less-Than Relational Responding

The RAI for Less-Than relational responding proceeded in an identical manner to Phase 2, with the exception that the statements included the words “Less-Than” instead of the words “More-Than”. In addition, the questions included the words “Less-Than” instead of “More-Than” and “More-Than” instead of “Less-Than”.

Phase 4: RAI for Opposite Relational Responding

The RAI for Opposite relational responding proceeded in an identical manner to Phase 1, with the exception that the statements included the words “Opposite of” instead of “the Same as”. In addition, the questions presented included the words “Opposite of” instead of “the Same as” and “Same as” instead of “Opposite of”.

Results and Discussion

The raw data for each participant’s pilot test of the RAI are listed in Table 28. Levels 1, 2 and 3 are then combined and averaged for each participant to produce a mean score out of 20 of baseline DRR for Same, More-Than, Less-Than and Opposite relations.

Table 28. Raw RAI scores out of 20 across all three RAI levels for each pilot participant for each relation measured.

Participant	RAI Level	Same	More	Less	Opposite
71	1	16	17	20	20
	2	20	19	20	20
	3	18	16	19	20
	Mean	**18**	**17.33**	**19.67**	**20**
72	1	18	20	20	20
	2	20	17	18	18
	3	17	8	17	18
	Mean	**18.33**	**15**	**18.33**	**18.67**
73	1	20	20	17	19
	2	20	20	18	20
	3	15	15	16	17
	Mean	**18.33**	**18.33**	**17**	**18.67**

In Table 28, each participant is shown by their participant number. Participants' baseline scores for each of the three RAI levels of each relational phase are then shown. It can be seen from this table that normally developing child participants displayed a high level of proficiency of relational responding without any experimental intervention. The reader is reminded that a score of 18 would be responding at a rate of 90% correct. Responding at this rate might normally be considered criterion level relational ability in a MTS relational network training context. This is not the criterion that is applied here but it nevertheless serves as a guide for assessing fluency of performances. For our purposes here, any score of 15 or more may be considered well above chance levels. Scores of 10-14 may be considered near chance levels. Scores of 9 or below may be considered lower than chance levels. It can be seen from Table 28 that participants scored somewhat lower on Level 2 and Level 3 RAI tasks than they scored on Level 1 RAI tasks. It was not surprising that participants might achieve lower scores when the presentation of relational statements is not linear and when time limits on responding are introduced.

Thus, the RAI does appear to be distinguishing broadly between levels of behavioural demand across the various relational repertoires. However, with the exception of the Level 3 (i.e., including a time limit on completing a block of testing) of the More-Than tasks for P72, all participant's scored above chance level during the baseline phase. Thus, it would not appear that these normally developing participants would require a relational intervention for these relational skills. That is not to say, of course, that such an intervention might not be of some benefit to them. Even though participants regularly scored above chance level, participants often scored with less than 100% accuracy on some relational tests, and particularly for those presented with time constraints on responding. It also might benefit participants in other ways. For example, such an intervention might increase attentional skills or positively impact full scale IQ scores or individual subtest scores. Imposing the time constraints might also be expected to increase the fluency of responding. Thus, even normally developing participants might benefit from some amount of relational intervention. However, the purpose of the pilot test was merely to see that the RAI could, in principle, establish a child's baseline level of relational abilities against which improvements in relational repertoires could be assessed. It was also to ensure that the experimental instructions could be clearly understood by children, and to generally beta-test the specifically developed software. It would appear that the RAI provides enough inter-subject and inter-relation variability above zero and below 100% accuracy to be useful as a measure of assessing relational skills pre-intervention in a group of children with various educational disadvantages. This is the issue to which Experiment 10b now turns.

Experiment 10b

Method

Participants

Eight 6th class participants (P74, P75, P76, P77, P78, P79, P80 and P81) were put forward by the School Principal of a local school, in conjunction with the Deputy Principal/Resource Teacher, and the class teacher, to take part in the experimental intervention programme. All eight participants were identified by their classroom teacher as students who often struggled to maintain progress with their normally developing 6th class peers for a variety of reasons. These eight participants also frequently received the lowest scores in their classroom on standardised tests. Seven of the eight participants had received learning support assistance at some time during their academic careers as a result of below average attainment. Children are only eligible to receive learning support assistance[1] when they have scored below the 10th percentile on school based standardised tests for reading or mathematics in any given school year. Some of these participants were also in receipt of specialised resource assistance[1] for diagnosed learning disabilities. Others presented with no diagnosed learning disability, but were merely described as "struggling". For our purposes, all students who presented as "struggling" with the mainstream curriculum are described as presenting with a learning difficulty. These

[1] The reader should note that under the "General Allocation Model", the Department of Education provides small group setting remedial tuition to children struggling in school who present below the 10th percentile in attainment areas of literacy or numeracy on standardised tests, or who have been diagnosed by a Psychologist as presenting with a high incidence learning disability. More intensive learning assistance can be applied for on the recommendation of a Psychologist if a student is presenting with formally diagnosed multiple disabilities, a low incidence disability, such as ADHD, Autistic Spectrum Disorder, severe behaviour difficulties, et cetera. The allocation of specialised resource teaching is provided only when students' needs fall within certain clearly defined parameters and only when these needs cannot be met within the regular mainstream classroom. For more information of the allocation of learning support or resource teaching hours, the reader can refer to the NCSE Guidelines, 2008. (See also the Epsen Act, 2003).

participants will also be referred to in this thesis as “educationally challenged”. Specific participant information provided by parents, teachers, school psychologists and speech and language therapists is listed below. The reader should note that the descriptions provided by parents and teachers are presented in unedited form to as great an extent as possible. These descriptions contain obvious mentalisms and non functional terminology. They do, however, provide the reader with a picture of the range of difficulties experienced by these various participants in the school setting. Participants’ ages at the outset of the intervention are listed in Table 29. A summary of participants’ involvement with specialist teachers, psychologists and speech and language therapists is also provided in Table 30.

P74 was a twelve year old boy who had attended learning support assistance for mathematics in 2003 and 2004 due to standardised test (Sigma-T) scores below the 10th percentile. His teacher reported that while his Sigma-T scores had improved, that he frequently struggled with the class curriculum. She described his main difficulties as “related to his short attention span, poor concentration, being easily distracted and frequently talkative”. P74 was also described by his teacher as a generally well behaved child who liked to succeed and was popular with his peers.

P75 was an 11 year old female. She had received learning support assistance from Senior Infants up until the end of 5th class due to standardised test scores below the 10th percentile (Sigma-T and Micra-T). She had received this assistance in both literacy and numeracy. P75 had received a previous psycho-educational assessment in June 2002. The Psychologist’s report from this assessment indicated that P75 presented with normal intelligence, but that she also presented with signs of a specific learning difficulty in the area of reading. A psychometric review of P75's literacy attainments in November of 2002 by the same Psychologist, indicated that P75 was presenting significantly below what would be predicted, given her full scale IQ, in the area of literacy. Special resource teaching assistance was recommended to increase P75's literacy attainments. At the time of the commencement of the intervention programme, P75's teacher reported that P75 had difficulties with most academic areas, but that she was a cooperative student who worked well in group situations and who had generally good relationships with peers and teachers.

P76 was a twelve year old boy with a well documented history of learning disability. P76 was reported to have received resource and learning support assistance throughout his academic career. P76 received a psychometric assessment in 2001 after repeating Junior Infants. The Psychologist who administered this assessment reported that P76 was functioning within the exceptionally low range of ability (0.1st percentile) for both verbal and performance scales of the WISC-III^{UK}. P76 was also reported to have some difficulty recognising word and letter sounds at that time. It was also noted that P76 had difficulties with expressive and receptive language, for which a speech and language therapy (SLT) assessment was recommended. The follow-up SLT assessment took place in 2002. The results of this assessment indicated a mild to moderate delay of both comprehension and expression of language, along with a mild to moderate delay of vocabulary development. Blocked sessions of SLT were attended by P76 for several weeks. After that time, P76 showed some improvement in these areas and was discharged from the public SLT services. P76 received a further psychometric assessment in 2004 administered by the same Psychologist, who again utilised the WISC-

III^{UK}. Results of that assessment indicated that P76's full scale IQ of 65 continued to place him in the exceptionally low range of intellectual ability. This assessment report also noted that P76 presented with significant difficulties in the areas of reading, writing and spelling and that P76 would require continued resource and learning support assistance. P76's classroom teacher noted difficulties in all areas of curriculum at the time of recruitment. P76's teacher also noted that he had great difficulties with long-term memory, with a very limited vocabulary, with reading, with hand-writing, with remembering learned spellings, with consolidating mathematics, with motor development, with peer relations and with adult relations.

P77 was the only participant in the applied group who had not had access to learning support assistance. Thus, P77 scored above the 10th percentile on the Sigma-T and Micra-T tests each school year. However, his teacher reported that P77 often struggled to maintain progress with his same age peers. P77 was 11 years old at the start of the research programme. P77's parents and teachers reported major concerns with his concentration and motivation.

P78 was a twelve year old girl who had received learning support assistance throughout most of her academic career due to Micra-T and Sigma-T scores below the 10th percentile. She had not accessed learning support during the current academic career, despite qualifying for it, as her parents did not give permission for her to attend. P78's classroom teacher described her as a likeable, shy and nervous child who got along well with all of her peers and struggled with most academic tasks.

P79 was an 11 year old girl who had received learning support assistance throughout her academic career due to Sigma-T and Micra-T scores below the 10th percentile. P79 received a diagnosis of ADHD in Junior Infants from the local Child and Adolescent Mental Health Team. Unfortunately, no copy of this report was available at the school or from P79's mother. The teacher reported that P79 had poor concentration skills, was difficult to keep on task, tended to rush through tasks so as to finish quickly, that she loved to talk, but did not have a wide vocabulary, that she was generally mannerly and well-behaved. Difficulties with reading, spelling and hand-writing were also reported by P79's teacher.

P80 was an 11 year old girl who had received learning support assistance throughout her academic career as a result of Micra-T and Sigma-T scores below the 10th percentile. P80 received a psychometric assessment due to on-going concerns with academic progress in 2004. The Psychologist who administered the assessment reported that P80 was functioning within the lower end of the low average range (FSIQ 81), that she presented with specific learning difficulties in both literacy and numeracy, and that she presented with significant attention difficulties. This Psychologist recommended a referral to the local Child and Adolescent Mental Health team. She also recommended that P80 continue to receive learning support assistance and that her progress needed to be monitored closely. P80 was subsequently diagnosed with emotional disturbance (ADHD) in 2005 by the local Child and Adolescent Mental Health team. On the basis of this diagnosis, P80 began to receive resource teaching hours in addition to the learning support teaching that she had already been receiving. P80's classroom teacher reported great difficulties with most academic areas consistent with P80's diagnosis of ADHD. Reading, mathematics, spelling and handwriting difficulties were also reported by P80's classroom teacher. P80 was reported to be friendly, cooperative and confident with adults, as well as having good relationships with her peers.

P81 was an 11 year old girl who had received learning support assistance for most of her academic career as a result of Micra-T and Sigma-T scores below the 10th percentile. P81's classroom teacher reported that P81 had difficulties with listening skills, with long-term memory and with concentration. The teacher also reported that P81 had poor spelling ability, poor mathematics concepts and that P81 was mannerly, but very socially immature.

Table 29. Participants' ages at start of the experiment.

Participant ID	Age
P74	12 years, 0 months
P75	11 years, 10 months
P76	12 years, 5 months
P77	11 years, 11 months
P78	12 years, 11 months
P79	11 years, 6 months
P80	11 years, 11 months
P81	12 years, 2 months

It can be seen from the participant information and from Table 29 that all participants were roughly matched for age. All participants were in 6th class at a local school. All eight participants can be described as "struggling" in school, despite the fact that not all were presenting with diagnosed *learning disabilities*. For the purposes of this research, these participants were all considered to have *learning difficulties*. As stated earlier, these participants will also be described as "educationally challenged".

Table 30. Participant's educational profiles at the commencement of the research programme.

Participant ID	Specialist teaching received	Involvement with other professionals	Formal Diagnosis
P74	Learning support	n/a	n/a
P75	Learning support	Psychologist	Specific Learning Difficulty in reading
P76	Learning support, resource, Resource teaching, Speech and Language Therapy	Psychologist, Speech and Language Therapist	Mild General Learning Disability, Mild-Moderate Expressive and Receptive Language Delay
P77	n/a	n/a	n/a
P78	Learning support	n/a	n/a
P79	Learning support	Psychologist, Psychiatrist	Specific Learning Difficulties in reading, ADHD
P80	Learning support, Resource teaching	Psychologist, Psychiatrist	Specific Learning Difficulties in reading and mathematics, ADHD
P81	Learning support	n/a	n/a

Setting and Materials

Each participant took part in a classroom set aside by the primary school for the research. As in previous experiments, times for sessions were chosen based on when this classroom was available. Each participant was first administered a comprehensive and standard psychometric assessment to determine their full scale IQ and also to fulfil the agreement made with the school. Although various sections of the Wechsler Individual Achievement Test (WIAT-II^{UK}) were administered for literacy, numeracy, written language and oral language for diagnostic purposes required by the school, the current research programme used only the IQ measure as a dependent variable of the relational intervention. The measurement of full scale IQ was taken using the WISC-IV^{UK}[2] (described in detail below). Following the assessment, the RAI was administered to each child individually via a Macintosh™ *ibook* lap top computer. The RAI was controlled by software created by the author using *Psyscope* (Cohen et al., 1993).

The revised and updated version of the WISC-III^{UK} is the WISC-IV^{UK}. The WISC-IV^{UK} was employed as the measure if IQ in the current study. Like its predecessor, it is an individually administered, comprehensive clinical instrument for assessing the intelligence of children ages 6 years 0 months through 16 years 11 months (see Wechsler, 2003). The WISC-IV^{UK} provides composite scores that represent intellectual functioning in specified cognitive domains, (i.e., Verbal Comprehension Index, Perceptual Reasoning Index, Working Memory Index, and Processing Speed Index), as well as providing a composite score that represents a child's general intellectual ability or full scale IQ. Although the WISC-IV^{UK} does not map on to the WISC-III^{UK} exactly, there is sufficient overlap for the current research purposes. It should be noted, however that there were some changes to the subtests employed as core subtests from the WISC-III^{UK} to the WISC-IV^{UK} (see Wechsler, 2003). Only these core subtests were employed in the current research programme, and they are described in detail in Appendix 30. The core subtests administered all contribute to the four cognitive domains. The domains are described in Appendix 31. The Verbal Comprehension Index

[2] In the research that follows, it was necessary to use the WISC-IV^{UK} rather than the WISC-III^{UK} as the WISC-III^{UK} was no longer being used by NEPS as an assessment tool since the WISC-IV^{UK} became available. Because providing a standard assessment was part of the terms of the access to participants, it was necessary to no longer use the WISC-III^{UK}, but to use the WISC-IV^{UK}. While the WISC-IV^{UK} is not identical to the WISC-III^{UK}, it may be considered the functional equivalent for present purposes.

(VCI) and the Perceptual Reasoning Index (PRI) in the WISC-IV^{UK} are roughly analogous to the Verbal IQ (VIQ) and the Performance IQ (PIQ) in the WISC-III^{UK}. In fact, Wechsler states that; "With the change in structure and nomenclature, the VCI and PRI should be substituted for the VIQ and PIQ in clinical decision-making and other situations where the VIQ and PIQ were previously required" (Wechsler, 2003, p. 6). Thus, the VCI and PRI received similar analysis as that conducted in Chapter 2 with the VIQ and PIQ. The full scale IQ, individual subtests and the two other indices (Working Memory Scale and Processing Speed Scale) were also addressed (see Wechsler, 2004, for reliability and validity information).

Ethics

Experiments 1 and 2 of this thesis received ethical approval by the NUIM Ethics Committee. As the format of Experiment 10b was similar to Experiments 1 and 2, ethical approval was not sought specifically for Experiment 10b. However, at all times, ethical guidelines established by the Psychological Society of Ireland and the British Psychological Society were observed. In addition, as in previous and subsequent experiments in this thesis, consent was obtained from all participants (or their parents) in verbal and written form.

General Experimental Sequence

The experiment took place over the course of two or three ninety-minute sessions per participant. (Administrations of WIAT-II^{UK} took an additional one to two ninety-minute sessions per participant). Research commenced at the convenience of the school and the classroom teacher and upon the signing of consent forms. Participants were also comprehensively debriefed on the exact nature of the procedure and purpose of the experiment when all participants completed the total research programme (including Experiment 11). Participants were not ever informed as to the precise details of the IQ testing or the RAI testing protocols while they were engaged in the experiment. This was done in order to prevent confounding of the measure by participants discussing answers to test items on the WISC-IV^{UK} or any possible strategies for taking the RAI in between one student's completion of the WISC-IV^{UK} and RAI and the next student's commencement of the WISC-IV^{UK} and the RAI. Sessions in this experiment consisted of five phases in total. Phase 1 included the administration of a psychometric assessment using the WISC-IV^{UK}. Following the completion of the psychometric assessment, the administration of the RAI began. Phase 2 included three sets of twenty relational statements and an equal number of relational questions to establish each student's baseline skill levels at Same relational responding. Phase 3 included three sets of twenty relational statements and questions to establish a student's skill at More-Than relational responding. Phase 4 included three sets of twenty relational statements and questions to establish a student's skill at Less-Than relational responding. Phase 5 included three sets of twenty relational statements and questions to establish a student's skill at Opposite relational responding. Following the administration of all five phases, a Spearman's Rho correlational analysis was conducted to determine if there were any correlations between individual IQ subtests and RAI indices or between IQ indices and RAI indices.

Phase 1: IQ testing. Each participant was seated comfortably at a large table in the room designated by the school. The experimenter sat opposing the participant. No one else was present in the room, although in every instance, the school secretary was across the hall. The participant was then asked if they were willing to participate in the research. If they responded, “yes” (e.g., that they were willing), then the experimenter proceeded to administer the WISC-IV^{UK}. Every participant agreed to continue. The experimenter then proceeded to ask a series of questions about everyday situations, about general word usage, about basic computation and about abstract concepts, among other items. The experimenter also asked the participant to complete several timed tasks involving the manipulation of concrete objects (e.g., blocks). See Wechsler (2004) for more information about the procedures for administration of the WISC-IV^{UK} as well as actual subtest items. Each IQ test administration took between 1 and 2 hours and each child was permitted to take as many breaks as they needed. Following the IQ test administration, another appointment was set up with each child’s teacher for a date within one week of the IQ test administration to proceed with the next phase.

Phase 2: RAI for Same relational responding

This phase was identical to Phase 1 in Experiment 10a.

Phase 3: RAI for More-Than relational responding

This phase was identical to Phase 2 in Experiment 10a.

Phase4: RAI for Less-Than relational responding

This phase was identical to Phase 3 in Experiment 10a.

Phase 5: RAI for Opposite relational responding

This phase was identical to Phase 4 in Experiment 10a.

Results and Discussion

Each participant's IQ subtest scores, as well as full scale IQ scores are listed in this section. These IQ scores are listed exactly as they would be listed on the Wechsler record forms for the WISCIV[UK]. Following each participant's IQ test information, RAI raw scores and mean scores for Phases 2-5 for all eight participants are listed. Finally, correlational data between these two measures are presented.

P74. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	**Raw Scores**	**Scaled Scores**					
Block Design	14	4		4			4
Similarities	17	7	7				7
Digit Span	17	10			10		10
Picture Concepts	12	4		4			4
Coding	34	5				5	5
Vocabulary	33	8	8				8
Letter-Number sequencing	18	10			10		10
Matrix Reasoning	15	5		5			5
Comprehension	21	7	7				7
Symbol Search	24	9				9	9
Sum of Scaled Scores			22	13	20	14	69
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

In column 1 each subtest administered is shown. In column 2, the participant's raw score (or number correct) is shown. In column 3, the raw score is standardised so that it can be compared to an average score of 10. Columns 4-7 show each scaled score within the index to which it contributes (listed in the bottom row). Column 8 shows all the standard scores, which taken together, contribute to the full scale IQ.

The scaled scores for P74 are shown below with the composite score conversions. The standardisation procedure involves the re-calculation of scores around a mean of 100 (or, more correctly, the 90-109 range).

P74. Sum of Scaled Scores to Composite Score Conversions

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	22	85	16	79-93
Perceptual Reasoning (PRI)	13	65	1	60-75
Working Memory (WMI)	20	99	47	91-107
Processing Speed (PSI)	14	83	13	76-94
Full Scale (FSIQ)	69	77	6	73-83

In column 1, each scale or index is presented. In column 2, the sum of scaled scores for each index is displayed. Column 3 displays the composite scores which have been standardised so that an average range score falls between 90 and 109. In column 4, the percentile ranks for each index is presented. Finally, column 5 displays the range within which a participants "true" score is likely to fall at the 95th percentile confidence interval for each index.

The foregoing pairs of tables are also presented below for each remaining participant.

P75. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	22	6		6			6
Similarities	21	9	9				9
Digit Span	18	11			11		11
Picture Concepts	17	10		10			10
Coding	56	12				12	12
Vocabulary	34	8	8				8
Letter-Number sequencing	18	10			10		10
Matrix Reasoning	19	7		7			7
Comprehension	22	8	8				8
Symbol Search	23	9				9	9
Sum of Scaled Scores			25	23	21	21	90
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P75. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	25	91	27	85-98
Perceptual Reasoning (PRI)	23	86	18	79-95
Working Memory (WMI)	21	102	55	94-109
Processing Speed (PSI)	21	103	58	94-112
Full Scale (FSIQ)	90	92	30	87-97

P76. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	25	7		7			7
Similarities	12	5	5				5
Digit Span	19	11			11		11
Picture Concepts	12	4		4			4
Coding	40	6				6	6
Vocabulary	29	7	7				7
Letter-Number sequencing	15	6			6		6
Matrix Reasoning	19	7		7			7
Comprehension	15	4	4				4
Symbol Search	21	8				8	8
Sum of Scaled Scores			16	18	17	14	65
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P76. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	16	73	4	68-81
Perceptual Reasoning (PRI)	18	75	5	69-85
Working Memory (WMI)	17	91	27	84-99
Processing Speed (PSI)	14	83	13	76-94
Full Scale (FSIQ)	65	74	4	70-80

P77. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	28	7		7			7
Similarities	17	7	7				7
Digit Span	18	11			11		11
Picture Concepts	16	9		9			9
Coding	44	8				8	8
Vocabulary	40	10	10				10
Letter-Number sequencing	19	11			11		11
Matrix Reasoning	22	8		8			8
Comprehension	24	9	9				9
Symbol Search	22	8				8	8
Sum of Scaled Scores			26	24	22	16	88
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P77. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	26	93	32	87-100
Perceptual Reasoning (PRI)	24	88	21	81-97
Working Memory (WMI)	22	104	61	96-111
Processing Speed (PSI)	16	88	21	80-98
Full Scale (FSIQ)	88	90	25	85-95

P78. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	26	7		7			7
Similarities	16	6	6				6
Digit Span	10	3			3		3
Picture Concepts	18	9		9			9
Coding	44	7				7	7
Vocabulary	25	5	5				5
Letter-Number sequencing	12	3			3		3
Matrix Reasoning	18	6		6			6
Comprehension	18	5	5				5
Symbol Search	19	7				7	7
Sum of Scaled Scores			16	22	6	14	58
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P78. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	16	73	4	68-81
Perceptual Reasoning (PRI)	22	84	14	78-93
Working Memory (WMI)	6	59	0.3	55-70
Processing Speed (PSI)	14	83	13	76-94
Full Scale (FSIQ)	58	70	2	66-76

P79. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	26	8		8			8
Similarities	18	8	8				8
Digit Span	25	16			16		16
Picture Concepts	18	10		10			10
Coding	51	10				10	10
Vocabulary	31	7	7				7
Letter-Number sequencing	18	10			10		10
Matrix Reasoning	18	7		7			7
Comprehension	16	5	5				5
Symbol Search	18	7				7	7
Sum of Scaled Scores			20	25	26	17	88
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P79. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	20	81	10	75-89
Perceptual Reasoning (PRI)	25	90	25	83-98
Working Memory (WMI)	26	116	86	107-122
Processing Speed (PSI)	17	91	27	83-101
Full Scale (FSIQ)	88	90	25	85-95

P80. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	20	5		5			5
Similarities	15	6	6				6
Digit Span	16	9			9		9
Picture Concepts	17	10		10			10
Coding	58	12				12	12
Vocabulary	27	6	6				6
Letter-Number sequencing	16	8			8		8
Matrix Reasoning	10	2		2			2
Comprehension	23	8	8				8
Symbol Search	28	11				11	11
Sum of Scaled Scores			20	17	17	23	77
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P80. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	20	81	10	75-89
Perceptual Reasoning (PRI)	17	73	4	68-83
Working Memory (WMI)	17	91	27	84-99
Processing Speed (PSI)	23	109	73	99-117
Full Scale (FSIQ)	77	83	13	79-88

P81. Total Raw Score to Scaled Score Conversion for IQ test data.

Subtest	Raw Scores	Scaled Scores					
Block Design	46	11		11			11
Similarities	17	7	7				7
Digit Span	17	10			10		10
Picture Concepts	15	8		8			8
Coding	49	9				9	9
Vocabulary	27	6	6				6
Letter-Number sequencing	17	9			9		9
Matrix Reasoning	23	9		9			
Comprehension	20	7	7				7
Symbol Search	19	7				7	7
Sum of Scaled Scores			20	28	19	16	83
			Verbal Comprehension	Perceptual Reasoning	Working Memory	Processing Speed	Full Scale

P81. Sum of Scaled Scores to Composite Score Conversions.

Scale	Sum of Scaled Scores	Composite Score	Percentile Rank	Range at 95% Confidence Interval
Verbal Comprehension (VCI)	20	81	10	75-89
Perceptual Reasoning (PRI)	28	96	39	89-103
Working Memory (WMI)	19	97	42	90-105
Processing Speed (PSI)	16	88	21	80-98
Full Scale (FSIQ)	83	87	19	82-92

The foregoing table pairs show the subtest and index scores which all contribute to full scale IQ scores. It can be seen that the educationally challenged group of participants presented with full scale IQ scores ranging from 70 (P78) to 92 (P75). These baseline IQ scores will be retained for comparison post intervention so that any potential improvements can be assessed. Baseline RAI scores were also taken so that potential improvements in relational skill could also be assessed post intervention. Taking baseline RAI scores also allowed for a correlational analysis of relational abilities as measured by RAI scores and the various intelligence measures. Baseline raw and mean RAI scores are presented next in Table 31 for all participants.

The raw data for each participant's performance on the RAI are listed by RAI level for each Same, More-Than, Less-Than and Opposite relations. RAI levels 1, 2 and 3 are then combined and averaged for each participant to produce a mean score of baseline DRR for Same, More-Than, Less-Than and Opposite relations.

Table 31. Raw RAI scores out of 20 across all three RAI levels for each participant for each relation.

Participant	RAI Level	Same	More	Less	Opposite
74	1	9	12	13	12
	2	19	11	2	12
	3	8	11	3	18
	Mean	**15.33**	**11.33**	**6**	**14**
75	1	8	12	16	16
	2	3	10	20	17
	3	7	6	14	16
	Mean	**6**	**9.33**	**16.67**	**16.33**
76	1	9	8	14	10
	2	14	9	18	11
	3	11	6	12	11
	Mean	**11.33**	**7.67**	**14.67**	**10.67**
77	1	8	17	16	15
	2	5	20	20	18
	3	9	18	17	19
	Mean	**7.33**	**18.33**	**17.67**	**17.33**
78	1	20	20	18	19
	2	2	16	19	19
	3	1	16	18	8
	Mean	**7.67**	**17.33**	**18.33**	**15.33**
79	1	7	11	7	7
	2	5	18	10	10
	3	15	18	12	12
	Mean	**9**	**15.67**	**9.67**	**9.67**
80	1	8	9	12	13
	2	11	12	10	11
	3	9	10	12	10
	Mean	**9.33**	**10.33**	**11.33**	**11.33**
81	1	3	9	18	13
	2	11	2	0	20
	3	13	0	1	2
	Mean	**9**	**2**	**6.33**	**11.67**

In Table 31, raw RAI scores out of 20 across all three RAI levels for each participant for each relation are presented. Mean RAI scores for each relation are also presented for each participant. It can be seen from this table that educationally challenged children presented with more variation in their levels of proficiency of relational responding for Same, More-Than, Less-Than and Opposite relations than did their normally functioning same-age peers in the pilot group (Experiment 10a). The reader is again reminded that a score of 18 would be responding at a rate of 90% correct. Responding at this rate might normally be considered criterion level relational ability in a MTS relational network training context. This is not the criterion that is applied here but it nevertheless continues to serve as a guide for assessing fluency of performances. For our purposes here, the same system of describing scores will be employed as was used in Experiment 10a. Thus, any score of 15 or more is considered well above chance level. Scores of 10-14 may be considered near chance levels. Scores of 9 or below may be considered below chance levels.

Using the foregoing system, P74 attained above chance level for Same responding, at chance level for Opposite and More-Than responding and below chance level for Less-Than responding. P75 attained below chance levels for Same responding, near chance level for More-Than responding and above chance level for Less-Than and Opposite responding. P76 attained near chance levels for Same responding, below chance levels for More-Than responding and near chance levels for Less-Than responding and for Opposite responding. P77 attained below chance levels for Same responding and above chance levels for More-Than, Less-Than and Opposite relational responding. P78 also attained below chance levels for Same responding and above chance levels for More-Than, Less-Than and Opposite relational responding. P79 attained below chance levels for Same responding, above chance levels for More-Than responding, and near chance levels for Less-Than and Opposite responding. P80 attained near chance levels for Same, More-Than, Less-Than and Opposite relational responding. P81 attained below chance levels for Same responding, More-Than and Less-Than responding and near chance levels for Opposite responding. Given the foregoing mean relational abilities scores, it appears that most of these participants would benefit from some type of a relational intervention.

The current data allowed for a re-examination of the issue of whether or not any subtest or index of the individual participant's IQ scores is meaningfully related to their relational abilities as measured by the RAI. A preliminary correlational analysis using scatterplots revealed that there appeared to be many correlations between subtest and full scale scores on the $WISCIV^{UK}$ with Same, More-Than, Less-Than and Opposite relational responding. The preliminary analysis suggested that some of these correlations may be statistically significant. Thus, the foregoing correlations were then statistically examined using a Spearman's Rho Correlational analysis. Specifically, Spearman's Rho correlational analyses examined the relationships between; 1) mean RAI scores for Same relational responding with all IQ subtests and scales administered, 2) mean RAI scores for More-Than relational responding with all IQ subtests and scales administered, 3) mean RAI scores for Less-Than relational responding with all IQ subtests and scales administered, and 4) mean RAI scores for Opposite relational responding with all IQ subtests and scales administered. Table 32 indicates the correlation coefficients for each of these analyses for each participant in this experiment.

Table 32: Spearman's RHO Correlations for mean RAI scores for Same relational responding, More-Than relational responding, Less-Than relational responding and Opposite relational responding with all IQ measures.

IQ Measure/ Subtest	Mean Same RAI DRR r values	P values	Mean More-Than RAI DRR r values	P values	Mean Less-Than RAI DRR r values	P values	Mean Opposite RAI DRR r values	P values
Block Design	-.282*	.249	-.122*	.387	.049	.454	-.268*	.260
Similarities	-.506***	.100	.086	.420	-.196*	.321	.258*	.269
Digit Span	-.179*	.336	.000	.500	-.086	.420	-.209*	.310
Picture Concepts	-.572***	.069	.247*	.278	.284*	.248	.012	.488
Coding	-.506***	.100	-.156*	.356	.096	.411	-.048	.455
Vocabulary	-.159*	.354	.230*	.292	-.036	.466	.412**	.155
Letter Number Sequencing	-.295*	.239	.317**	.222	-.195*	.322	.366**	.186
Matrix Reasoning	-.454**	.129	-.195*	.322	.073	.432	.171*	.343
Comprehension	-.451**	.131	.218*	.302	.121*	.387	.655***	.039
Symbol Search	.224*	.297	-.124*	.385	-.111*	.397	.173*	.341
Verbal Comprehension Scale	-.395**	.166	.258*	.269	-.061	.443	.614***	.053
Perceptual Reasoning Scale	-.587***	.063	-.048	.455	.071	.433	.024	.478
Working Memory Scale	-.301**	.234	.263*	.264	-.228*	.294	.084	.422
Processing Speed Scale	-.364**	.187	-.110*	.397	.037	.466	-.074	.431
Full Scale IQ	-.584***	.064	.000	.500	-.048	.455	.228**	.294

Note. The numbers in this table are raw "r" scores, or correlations. One asterisk indicates a mild correlation. Two asterisks indicate a moderate correlation and three asterisks indicate a strong correlation. Strong correlations which are also significant at the 0.05 level (1-tailed) are listed in red.

Cohen's (1988) system of interpreting values between 0 and 1 has been used here. In this interpretation r=.10 to .29 or r = -.10 to -.29 indicates a small or mild correlation. R= .30 to .49 or r= -.30 to -.49 indicates a medium or moderate correlation. R= .50 to 1.0 or -.50 to -.1.0 indicates a large correlation. (See Pallant, 2001 for detailed discussion of this system).

From Table 32, it can be seen that there are small, medium and large (or mild, moderate and strong) correlations between most subtests of the WISC-IV^{UK} and one or more types of mean relational ability as measured by the mean score on the Relational Abilities Index (RAI) for Same relations, More-Than relations, Less-Than relations and

Opposite relations. More specifically, there were mild correlations for the Block Design subtest with mean Same RAI scores, mean More-Than RAI scores and mean Opposite RAI scores. There were mild correlations for the Similarities subtest and mean Less-Than RAI scores and mean Opposite RAI scores. Mild correlations were also seen for the Digit Span subtest with mean Same RAI scores and mean Opposite RAI scores. There were also mild correlations noted for the Picture Concepts subtest with mean More-Than and Less-Than RAI scores. Mild correlations were seen for the Coding subtest with mean More-Than RAI scores. Mild correlations were also observed for the Vocabulary subtest with mean Same and More-Than RAI scores. There were also mild correlations seen for the Letter Number Sequencing subtest with mean Same and Less-Than RAI scores. Mild correlations were also observed for the Matrix Reasoning subtest with mean More-Than and Opposite RAI scores. Mild correlations were seen for the Symbol Search subtest with all four mean RAI scores. Of the four factor index scales, mild correlations were seen with the Verbal Comprehension scale, the Working Memory Scale and the Processing Speed scale and the mean More-Than RAI scores. Mild correlations were also evident for the Working Memory Scale with mean Less-Than RAI scores.

Fewer moderate than mild correlations were seen between subtest and scale scores with mean RAI scores. More specifically, medium or moderate correlations were seen for the Vocabulary subtest with mean Opposite RAI scores. Moderate correlations were also seen for the Letter Number Sequencing subtest with mean More-Than and mean Opposite RAI scores. There were also moderate correlations for the Matrix Reasoning subtest with mean Same RAI scores. Similarly, moderate correlations were seen for the Comprehension subtest, the Verbal Comprehension Scale, the Working Memory Scale and the Processing Speed Scale with mean Same RAI scores. One final moderate correlation was seen for Full Scale IQ with mean Opposite RAI scores.

Large or strong correlations were also seen with various subtest or scale scores with mean RAI scores. Specifically, strong correlations were seen for the Similarities subtest with mean Same RAI scores. Similarly, strong correlations were also seen for the Picture Concepts subtest, the Coding subtest, the Perceptual Reasoning Scale and Full Scale IQ with mean Same RAI scores. A strong and significant at the 0.05 level (1 tailed) correlation was seen for the Comprehension subtest with mean Opposite DRR scores. Finally, a strong correlation approaching significance level ($p = .053$) was also seen for the Verbal Comprehension Scale with mean Opposite RAI scores.

Similar to the findings in Experiment 1, Chapter 2, many correlations were found between many of the subtest scores and the relational measure used. It is also interesting to note that there were only 17 analyses run (out of sixty) for which no correlation was found for IQ subtest and any one of the four relational tasks.

The matter of the surprising direction of some of the observed correlation was also discussed in Chapter 2, in which several unexpected correlations between relational abilities and IQ scores and subtests were found. In the current chapter, it can be seen that most of the correlations observed here were in an unexpected direction. While this may be surprising, it is entirely possible that even with a population of participants who are struggling in school, that a Same relational repertoire is already well established, and therefore, sufficiently well developed for most individuals that it bears little relevance to IQ. Interestingly, fewer correlations in an unexpected direction were observed between the individual subtest and scale scores with mean More-Than RAI scores. This supports the foregoing suggestion insofar as this relational repertoire is less well developed among school children, and skills levels in this domain might be expected to be roughly distributed across this population in a manner related to general intellectual ability. It is perhaps relevant also that the positive correlations between More-Than RAI scores were amongst the strongest of all the correlations observed (i.e., stronger than the negative correlations).

The pattern is less clear for correlations seen between the individual subtest and scale scores with mean Less-Than RAI scores. Overall, fewer correlations were seen between the individual subtest and scale scores with mean Less-Than RAI scores than with any other mean RAI score. The correlations seen were all mild. Two of them were in an expected direction (i.e., positive correlations) and four of them were in an unexpected direction (i.e., negative correlations). At this point it is very difficult to speculate on the lack of correlations observed. However, it is entirely possible that relations of Less-Than simply do not correlate meaningfully with IQ subtests or scales.

Finally, more clear patterns of correlations were observed between the individual subtest and scale scores with mean Opposite RAI scores. More specifically, out of ten correlations seen, eight were in an expected direction. The two correlations which were not in an expected direction (i.e., Block Design and Digit Span) were mild correlations and do not contribute to the Verbal Comprehension Scale. In addition, there was one correlation here which was strong and significant (Comprehension and mean Opposite RAI score). Comprehension is one of the subtests which contribute to the Verbal

Comprehension Scale. In addition, the Verbal Comprehension Scale scores were also strongly correlated with mean Opposite RAI scores.

It is important to note that several previous experiments in the current thesis pointed to the relational complexity of Opposite relational responding for both adults and children, compared to Same, More-Than and Less-Than relations. Thus, it is not surprising that the large variances in fluency in Opposite relational responding typically observed across the population should be correlated with a measure of general relational ability such as that presented by standard IQ tests. In addition, the fact that the strongest correlations found in this domain were in an expected direction and were observed between RAI scores and *verbal task* performance provides further support for the RFT suggestion that there exists a functional relationship between verbal skills, as measured by standardised IQ tests items, and relational skills.

Given the foregoing, it must be considered at this stage that only relational skill levels that show a large variance across the population and which are at a relatively low level of development for most individuals can be used as relational ability indices. These very relational abilities may also be those that correlate best with various intelligence measures. Of course, these correlations may be arrived at by painstaking correlational studies across an infinite number of relational skills and an almost infinite number of intelligence measures, across various populations. However, the current research, as a basic research enterprise, first attempted to establish the functional relationships between various relations, and this in turn could now be said to have pointed to the most likely relational skill of those studies to be used as a viable RAI index for general intelligence. Of course, the correlations between Opposite RAI scores and IQ are not ubiquitous and maximal. Nevertheless, those that have been observed are rendered more meaningful for the various relations examined to have been studied in relation to each other at a more basic level in advance of the current study.

Several interpretive challenges still remain for the current data. These are magnified by the necessarily small sample size, the young age of the participants, and the fluctuations which may occur in performances due to age and also due to the individual differences within the small applied population employed. These types of fluctuations can make for perhaps less than perfectly controlled experimental designs. These issues will be returned to in the General Discussion. What is important to understand at this point however, is that there is no logical requirement for the observation of a correlation between RAI scores and IQ scores for one to functionally interact with the other. In other words, it may emerge that behavioural repertoires not considered in the current research have a larger than expected influence on IQ scores. In addition, it may be that relational repertoires and these other, yet to be discovered, hypothetical repertoires have arisen via somewhat different sets of contingencies. In effect, we might in this case, not expect to observe large, or very frequent, correlations between these two repertoires. However, even in this scenario in which two behavioural repertoires have different sources, it may still be possible for the relevant behavioural repertoires to interact. More specifically, it may still be possible to raise the IQ of a child using a relational responding intervention, even if general IQ scores and relational ability scores fail to correlate highly. As an example, the development of general attending skills and reading skills may be established in different but related ways in a school setting, and may be considered relatively separate repertoires. Indeed, poor correlations may be observed between measures of each. Nevertheless, it does not follow, in this case, that improving general attending skills will not improve reading. Thus, while RFT argues that the source of relational skills and most intellectual skills is shared, this is not a necessary precondition for functional relationships to exist between IQ and relational ability. Consequently, the endeavour to raise IQ reported in the following experiment is in no way undermined or threatened by relatively ambiguous correlations found between IQ and Same, More-Than and Less-Than relational abilities in the current study.

Experiment 11

Experiment 10 has shown that there were some mild/moderate correlations between various IQ subtests and scales with derived relational abilities as measured by mean RAI scores for Same, More-Than, Less-Than and Opposite relations. There were also some strong and significant correlations noted which were of particular interest. The mild/moderate corrections were, in many cases somewhat ambiguous. Nevertheless, as noted in the General Discussion of Experiment 10b there is no logical requirement for the observation of a correlation between RAI scores and IQ scores for one to functionally interact with the other. It may still be possible to raise the IQ of a child using a relational responding intervention, even if general IQ scores and relational ability scores fail to correlate highly. Thus, these same participants were retained for Experiment 11, to examine whether or not the developed DRR intervention would improve relational responding skills (as measured by the RAI) and in turn, have an impact on IQ subtests and index scores.

Method

Participants

Participant information was identical to that listed in Experiment 10b.

Setting and Materials

The materials used were identical to those described in Experiment 10b.

Ethics

Information regarding ethical considerations was identical to that described in Chapters 2, 3, 4 and 5.

General Experimental Sequence

The experiment took place over the course of the majority of the school year for bi-weekly 70-90 minute sessions. Following the completion of Experiment 10b, participants took part twice weekly until they met criterion on all measures (except when absent from school or engaged in conflicting school activities). As in previous experiments, the research for Experiment 11 commenced at the convenience of the participants and the school administration. As before, participants were also comprehensively debriefed on the nature of the procedure and purpose of the experiment when all participants completed the research. Participants were not ever informed as to the precise details of the training and testing protocols while they were still engaged in the experiment. This was done in order to prevent confounding of the measure by participants discussing possible strategies for meeting criterion in between training and testing sessions outside of the experimenter's awareness. Sessions in this experiment consisted of fourteen phases in total. Three of the phases (i.e., remedial levels for Same, More-Than/Less-Than and Opposite) were made available to participants only as needed. Not all participants required these extra phases for all relations.

Following the administration of the WISC-IV^{UK} IQ test and the RAI to each participant in Experiment 10b, Experiment 11 commenced the intervention. Phase 1 included pre-training to establish the arbitrary cues for relational responding in accordance with Same relations. Phase 2 included the administration of a series of relational training tasks via a lap top computer. These relational training tasks formed the training procedure for the establishment of responding in accordance with Same (i.e., Same training). Phase 3 included the administration of a test for derived or emergent Same relations (i.e., Same testing). Phase 4 was a Same remedial phase which was administered to participants only if they were unable to pass Phase 3 after four cycles. More specifically, if a participant failed to meet criterion on a Phase 3 probe test following a remedial level, then remedial phases were made available after every MET test. In this manner, if a participant failed to meet criterion on the first exposure to a non-MET test following a remedial level, the intervention changed to include more access to remedial levels such that a participant received cycles of relational training, relational (non-MET) testing, further relational training, multiple exemplar testing, remedial training and remedial testing as standard until such time as a participant met criterion on a non-MET test. Phase 5 included pre-training to establish the arbitrary cues for relational responding in accordance with More-Than and Less-Than relations. Phase 6 included the administration of relational training tasks for the establishment of relational responding in accordance with More-Than/Less-Than relations. Phase 7 included the administration of a test for derived relational responding (DRR) in accordance with emergent Opposite relations (i.e., Opposite testing). Phase 8 was a remedial phase which was administered to participants only if they were unable to pass Phase 7 after four cycles. As described in Phase 4 with Same relational responding, if a participant failed to meet criterion on the first exposure to a non-MET test following a remedial level, the remedial levels for More-Than/Less-Than relational responding were included as standard. Phase 9 included pre-training to establish the arbitrary cues for relational responding in accordance with Opposite relations. Phase 10 included the administration of relational training tasks for the establishment of relational responding in accordance with Opposite relations (i.e., Opposite training). Phase 11 included the administration of a test for derived relational responding (DRR) in accordance with emergent Opposite relations (i.e., Opposite testing). Phase 12 was an Opposite remedial phase, which was administered to participants only if they were unable to pass Phase 6 after four cycles. In other words, if a participant failed to meet criterion on the first exposure to a non-MET test following a remedial level, then the remedial levels for Opposite relational responding were included as described in Phase 4. Phase 13 involved the administration of a follow-up RAI for Same, More-Than, Less-Than and Opposite relations. Phase 14 involved the administration of a follow-up WISC-IV^{UK}.

Phase 1: Same Relational Pre-training

Pre-training in this experiment was identical to pre-training in Experiments 3, 4, 5 and 6 of the previous chapters. (The reader should note that as before, cues for Opposite were also trained to control for the possibility of a participant being inadvertently trained to respond “Yes” every time the Same cue appeared on the screen. This served to further ensure contextual control over the cues and the relata). See Figure 36.

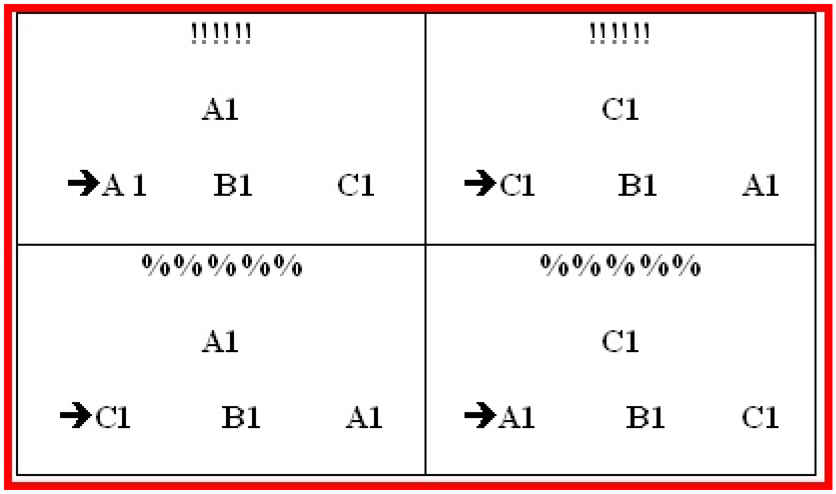


Figure 36. Tasks presented during Same relational pre-training in Phase 1. The alphanumerics represent the non-arbitrary stimuli employed as stimuli. The "!!!!!" and "%%%%%" are cues representing Same and Opposite. An arrow indicates the correct choice in the presence of feedback.

Phase 2: Same Relational Training

Same relational training in this experiment was identical to Same relational training in Experiment 4 and 5 in Chapter 4. See Figure 37.

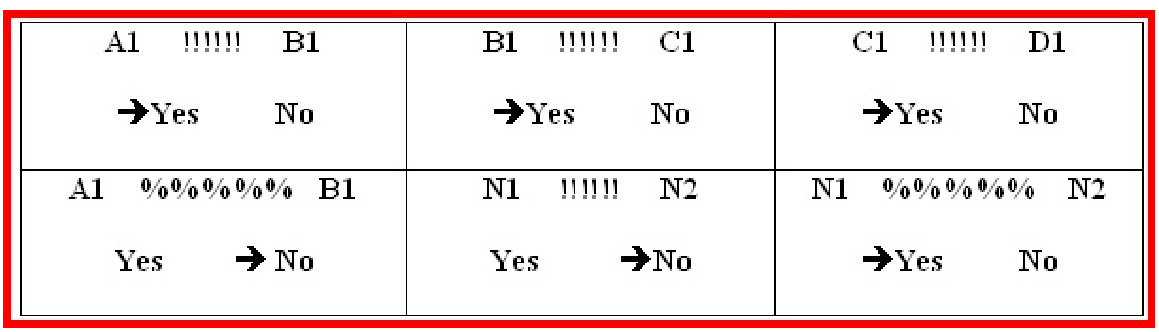


Figure 37. Tasks presented during Same relational training in Phase 2. The alphanumerics represent the arbitrary nonsense syllables employed as stimuli. The "!!!!!" and "%%%%%" are cues representing Same and Opposite which had been established during the pre-training phase. An arrow indicates the correct choice in the presence of feedback.

Phase 3: Testing for Derived Same Relational Responding (Non-MET and MET)

Non-MET and MET testing in Phase 3 proceeded in an identical manner to that described for experimental participants in Experiments 4 and 5 in Chapter 4. See Figure 38.

D1 !!!!!! A1 →Yes No	C1 !!!!!! A1 →Yes No
D1 %%%%%% A1 Yes →No	C1 %%%%%% A2 Yes →No

Figure 38. Tasks presented during Same Relational Testing in Phase 3. The alphanumerics represent the arbitrary nonsense syllables employed as stimuli. The "!!!!!" and "%%%%%" are the cues representing Same and Opposite which had been established during the pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback

Phase 4: Same Remedial Training and Testing

If a participant was not able to pass Phase 3 after seven exposures to standard training, three exposures to multiple exemplar testing and four exposures to regular testing, then a participant was exposed to Phase 4: remedial training and testing for Same relations. This phase was identical to remedial training and testing for Same relations in Chapter 4. However, this phase was made available to participants more regularly if participants did not meet criterion on the first non-MET test for derived Same relations following this phase. More specifically, if a participant failed to meet criterion on the first exposure to a relational test following a remedial level, the remedial level then became a standard part of the intervention cycle. For example, such a participant who had not met criterion on a relational test following a remedial level would then be exposed to a cycle of further relational training with a novel stimulus set, a multiple exemplar test, further remedial relational training with a non-arbitrary novel stimulus set, followed by a remedial multiple exemplar test employing the same non-arbitrary stimulus set, followed by more relational training with a new novel stimulus set, followed by a relational test (i.e., without feedback). This cycle, including the remedial levels as standard, continued until such time as a participant met criterion on the first relational test following a remedial level. (See Figure 39 for training tasks. See Figure 40 for testing tasks). Upon doing so, a participant was considered to have demonstrated the generalisation of relational responding in accordance with Same relations and proceeded to the next phase, Phase 5.

A1 !!!!!! B1 →Yes No	B1 !!!!!! C1 →Yes No	C1 !!!!! D1 →Yes No
A1 %%%%%% B1 Yes → No	N1 !!!!!! N2 Yes →No	N1 %%%%%% N2 →Yes No

Figure 39. Tasks presented during remedial Same relational training in Phase 4. The alphanumerics represent the non-arbitrary stimuli used. The “!!!!!” and “%%%%%” are cues representing Same and Opposite which had been established during the pre-training phase. An arrow indicates the correct choice in the presence of feedback.

Figure 40 shows a sample of the remedial testing tasks employed in Phase 4.

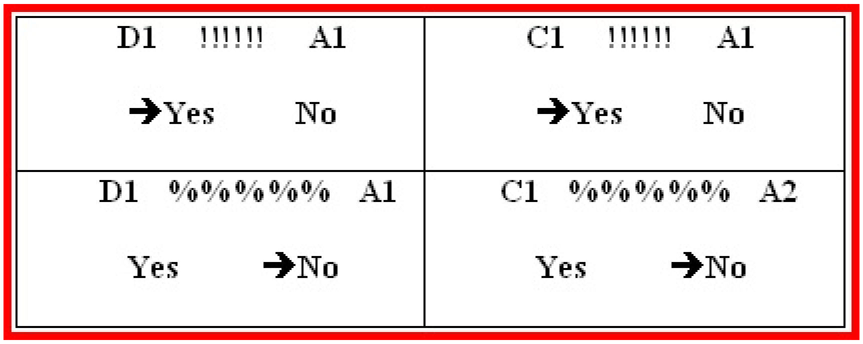


Figure 40. Tasks presented during Same remedial relational testing in Phase 4. The alphanumerics represent the non-arbitrary stimuli used. The "!!!!!" and "%%%%%" are the cues representing Same and Opposite which had been established during the pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) of presence (MET phases) of feedback.

Phase 5: More-Than/Less-Than Relational Pre-training

More-Than/Less-Than relational pre-training proceeded in an identical manner to that described in Experiment 6, Chapter 4. Figure 41 displays a sample of the tasks used during this phase.

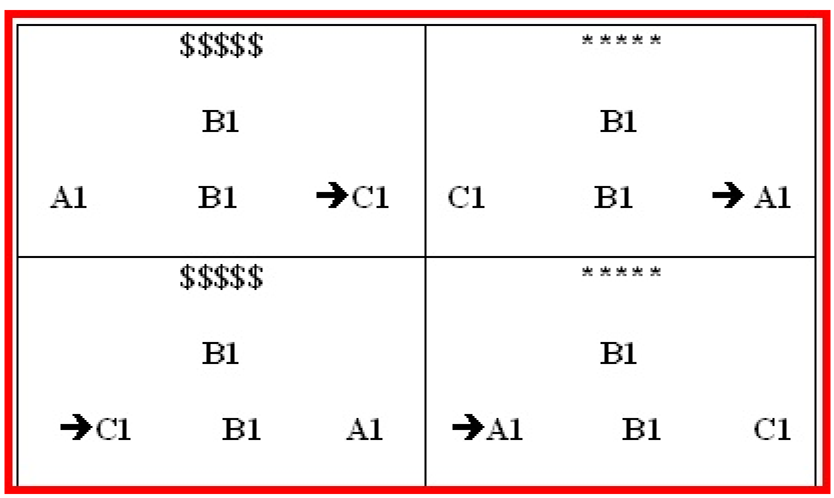


Figure 41. Tasks presented during More-Than/Less-Than relational pre-training in Phase 5. The alphanumerics represent non-arbitrary stimulus sets (see Appendix 21). The $$$$$ and ***** represent the cues for More-Than and Less-Than, respectively, which were established during this phase. An arrow indicates the correct choice in the presence (training phases) or absence (testing phases) of feedback.

Phase 6: More-Than/Less-Than Relational Training

The training procedure employed in Phase 6 was identical to the More-Than/Less-Than training procedure described in Experiment 6, Chapter 4. See Figure 42.

A1 $$$$$ B1 →Yes No	B1 $$$$$ C1 →Yes No	C1 $$$$$ D1 →Yes No
A1 ***** B1 Yes →No	N1 $$$$$ N2 Yes →No	N1 ***** N2 →Yes No

Figure 42. Tasks presented during More-Than/Less-Than Relational Training in Phase 6. The alphanumerics represent the arbitrary nonsense syllables employed as stimuli. The $$$$$ and ***** represent the cues for More-Than and Less-Than respectively which were established during the Phase 5 Pre-training. An arrow indicates the correct choice in the presence of feedback.

Phase 7: More-Than/Less-Than Relational Testing (MET and non-MET):

The testing procedure employed in Phase 7 was identical to the More-Than/Less-Than testing procedure described for experimental (MET) participants in Experiment 6, Chapter 4. See Figure 43.

D1 $$$$$ A1 →Yes No	C1 $$$$$ A1 →Yes No
D1 ***** A1 Yes →No	C1 ***** A2 Yes →No

Figure 43. Tasks presented during More-Than/Less-Than relational testing in Phase 7. The alphanumerics represent the nonsense syllables employed as stimuli. The $$$$$ and ***** represent the cues for More-Than and Less-Than respectively which were established during the Phase 5 pre-training. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Phase 8: More-Than/Less-Than Remedial Training and Testing:

The More-Than/Less-Than remedial training and testing phase proceeded in an identical manner to the More-than/Less-Than remedial phase described for MET participants in Experiment 6, Chapter 4. As was seen with Phase 4, Same remedial training and testing, Phase 8 was made available to participants on a more regular basis if needed. More specifically, if a participant failed to meet criterion on the first exposure to a relational test following a remedial level, the remedial level then became a standard part of the intervention cycle. For example, such a participant who had not met criterion on a relational test following a remedial level would then be exposed to a cycle of further relational training with a novel stimulus set, a multiple exemplar test, further remedial relational training with a non-arbitrary novel stimulus set, followed by a remedial multiple exemplar test employing the same non-arbitrary stimulus set, followed by more relational training with a new novel arbitrary stimulus set, followed by a relational test (i.e., without feedback). This cycle, including the remedial levels as standard, continued until such time as a participant met criterion on the first relational test following a remedial level for More-Than/Less-Than relations. (See Figure 44 for training tasks. See Figure 45 for testing tasks). Upon doing so, a participant had demonstrated the generalisation of relational responding in accordance with More-Than/Less-Than relations and proceeded to Phase 9.

A1 $$$$$ B1 →Yes No	B1 $$$$$ C1 →Yes No	C1 $$$$$ D1 →Yes No
A1 ***** B1 Yes → No	N1 ***** N2 Yes →No	N1 ***** N2 →Yes No

Figure 44. Tasks presented during remedial More-Than/Less-Than relational training in Phase 4. The alphanumerics represent the non-arbitrary stimuli employed. The “$$$$$” and “*****” are cues representing More-Than and Less-Than which had been established during the pre-training phase. An arrow indicates the correct choice in the presence of feedback.

See Figure 45 for a sample of the remedial testing tasks employed during this phase.

D1 ***** A1 →Yes No	C1 ***** A1 →Yes No
D1 $$$$$ A1 Yes →No	C1 $$$$$ A2 Yes →No

Figure 45. Tasks presented during Same remedial relational testing in Phase 4. The alphanumerics represent the non-arbitrary stimuli employed. The “*****” and “$$$$$” are the cues representing Less-Than and More-Than which had been established during the pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Phase 9: Opposite Relational Pre-training.

Opposite relational pre-training proceeded in an identical manner to Same relational pre-training described in Phase 1 of the current experiment and also described in Chapters 3 and 4. (See Figure 36 in Phase 1).

Phase 10: Opposite Relational Training.

The training procedure used in Phase 10 was identical to that described for MET (experimental) participants in Experiments 4 and 5 in Chapter 4. See Figure 46.

A1 %%%%% B1 →Yes No	B1 %%%%% C1 →Yes No	C1 %%%%% D1 →Yes No
A1 !!!!!! B1 Yes → No	N1 %%%%% N2 Yes →No	N1 !!!!!! N2 →Yes No

Figure 46. Tasks presented during Opposite Relational Training in Phase 5. The alphanumerics represent the arbitrary nonsense syllables employed as stimuli. The "!!!!!" and "%%%%%" are the arbitrary cues for Same and Opposite respectively which had been established during pre-training. An arrow indicates the correct choice in the presence of feedback.

Phase 11: Testing for Derived Opposite Relations (Non-MET and MET).

Phase 11 was identical to testing for the MET participants in Experiments 4 and 5 of Chapter 4 for derived Opposite relations. (See Figure 47 for the tasks presented during Phase 11).

D1 %%%%% A1 →Yes No	C1 %%%%% A1 Yes →No
D1 !!!!!! A1 Yes →No	C1 !!!!!! A1 →Yes No

Figure 47. Tasks presented during Opposite relational testing in Phase 11. The alphanumerics represent the arbitrary nonsense syllables employed as stimuli. The "!!!!!" and "%%%%%" are the arbitrary cues for Same and Opposite respectively which had been established during pre-training. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

Phase 12: Opposite Remedial Training and Testing

The Opposite remedial training and testing phase proceeded in an identical manner to those described in Phase 4 (Same remedial phase) and Phase 8 (More-Than/Less-Than remedial phase), employing non-arbitrary stimuli to train Opposite relations. Also as described in Phases 4 and 8, Opposite remedial training and testing was made available to participants on a more regular basis if needed. More specifically, if a participant failed to meet criterion on the first exposure to a relational test following a remedial level, the remedial level then became a standard part of the intervention cycle. For example, such a participant who had not met criterion on a relational test following a remedial level would then be exposed to a cycle of further relational training with a novel stimulus set, a multiple exemplar test, further remedial relational training with a non-arbitrary novel stimulus set, followed by a remedial multiple exemplar test employing the same non-arbitrary stimulus set, followed by more relational training with a new novel stimulus set, followed by a relational test (i.e., without feedback). This cycle, including the remedial levels as standard, continued until such time as a participant met criterion on the first relational test following a remedial level for Opposite relations (see Figure 48 for training tasks. See Figure 49 for testing tasks). Upon doing so, a participant had demonstrated the generalisation of relational responding in accordance with Opposite relations and was considered to be finished with the intervention and thereby proceeded to the follow-up measures in Phases 13 and 14.

A1 %%%%%% B1 →Yes No	B1 %%%%%% C1 →Yes No	C1 %%%%%% D1 →Yes No
A1 !!!!!! B1 Yes → No	N1 %%%%%% N2 Yes →No	N1 !!!!!! N2 →Yes No

Figure 48. Tasks presented during remedial Opposite relational training in Phase 12. The alphanumerics represent the non-arbitrary stimuli employed. The "!!!!!" and "%%%%%" are cues representing Same and Opposite which had been established during the pre-training phase. An arrow indicates the correct choice in the presence of feedback.

D1 !!!!!! A1 →Yes No	C1 !!!!!! A1 →Yes No
D1 %%%%%% A1 Yes →No	C1 %%%%%% A2 Yes →No

Figure 49. Tasks presented during Opposite remedial relational testing in Phase 3. The alphanumerics represent the non-arbitrary stimuli employed. The "!!!!!" and "%%%%%" are the cues representing Same and Opposite which had been established during the pre-training phase. An arrow indicates the correct choice in the absence (non-MET phases) or presence (MET phases) of feedback.

See also Figure 50 below for a schematic of all training and testing tasks employed in Phases 2, 3, 4, 6, 7, 8, 10, 11 and 12 of Experiment 11).

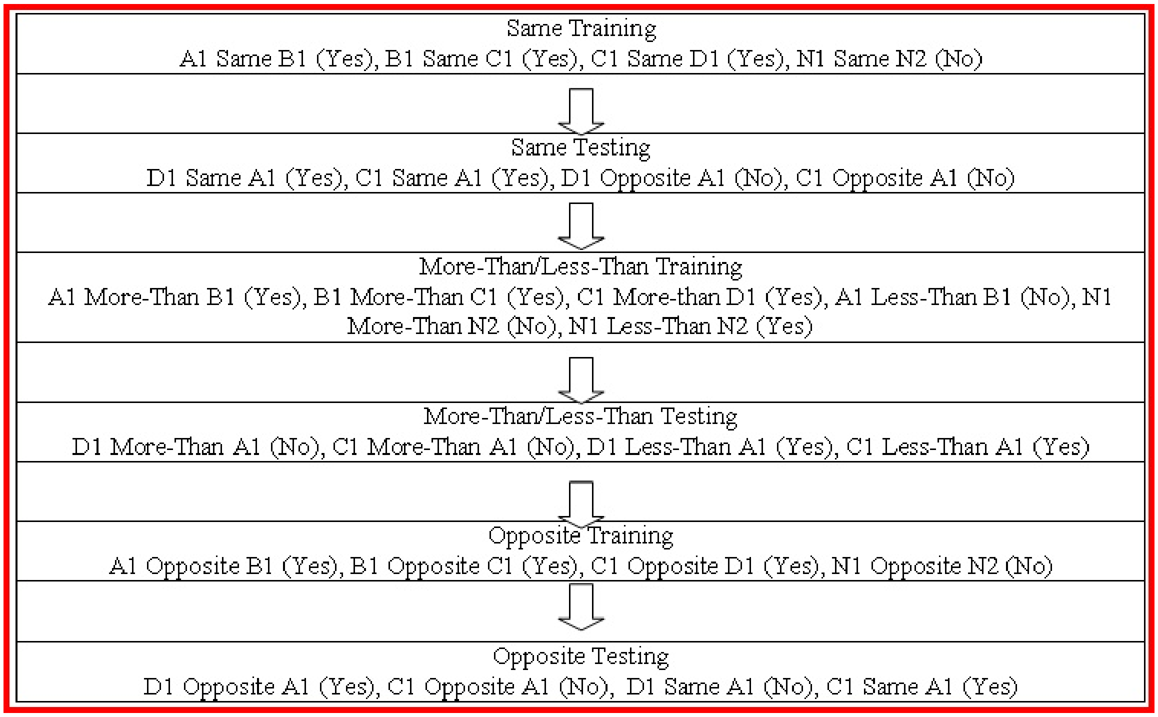


Figure 50. Schematic of the training and testing tasks employed during the intervention in Experiment 11 where Yes and No in parentheses indicate correct responses. The remedial levels (Phases 4, 8 and 12) are identical to standard training and testing phases, with the exception that instead of using nonsense words (as represented by alphanumerics in the schematic), pictures of cats, dogs, houses, et cetera are used. Pre-training tasks are not included in this schematic.

Phase 13: Follow-up Relational Abilities Index (RAI).

The follow-up RAI proceeded in an identical manner to the RAI procedures described in Experiment 10a and 10b. Specifically, four follow-up RAI tests were conducted with each participant. The first RAI test (described as Phase 1 in Experiment 10a) included three sets of twenty statements and questions to establish a student's skill at Same relational responding. The second RAI test (described as Phase 2 in Experiment 10a) included three sets of twenty statements and questions to establish a student's skill at More-Than relational responding. The third RAI test (described as Phase 3 in Experiment 10a) included three sets of twenty statements and questions to establish a student's skill at Less-Than relational responding. The final RAI test (described as Phase 4 in Experiment 10a) included three sets of twenty statements and questions to establish a student's skill at Opposite relational responding.

Phase 14: Follow-up IQ test.

The WISC-IV^{UK} was administered to each participant in an identical manner to that described in Experiment 10b. No additional assessment tools were administered at this time.

Results and Discussion

The results section details each participant's individual progress throughout the intervention and follow-up measures. Table 31 displays each participant's total requirements for the number of blocks of pre-training, relational training, relational testing, multiple exemplar testing and remedial levels in order to meet criterion for Same relational responding. This table also shows how many novel stimulus sets were required across the training and testing cycles. Table 32 displays each participant's total requirements for the number of blocks of pre-training, relational training, relational testing, multiple exemplar testing and remedial levels in order to meet criterion for More-Than/Less-Than relational responding. This table also shows how many novel stimulus sets were required across the training and testing cycles for More-Than/Less-Than relations. Table 33 shows each participant's total requirements for the number of blocks of pre-training, relational training, relational testing, multiple exemplar testing and remedial levels in order to meet criterion for Opposite relational responding. This table also shows how many novel stimulus sets were required across the training and testing cycles for Opposite relations. Table 34 shows all participants' raw RAI scores prior to and following the intervention. Table 35 shows all participants' mean RAI scores pre and post intervention. Table 36 shows each individual participant's scores on the WISC-IV^{UK} prior to and following the intervention. All results are discussed.

Table 31. The number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion for Same relations on each of the phases in Experiment 11.

Participant Number	Blocks of Same relational pre-training	Novel stimulus sets (not including remedial levels)	Blocks of Same relational training	Blocks of Same relational testing	Blocks of Same Multiple Exemplar testing	Blocks of Same remedial training	Blocks of Same remedial testing
74	5	1	11	1	0	0	0
75	6	3	10	2	1	0	0
76	15	1	18	1	0	0	0
77	13	1	6	1	0	0	0
78	11	3	13	2	1	0	0
79	7	3	39	2	1	0	0
80	12	1	10	1	0	0	0
81	22	1	13	1	0	0	0

Table 31 shows the number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion for Same relations on each of the phases in Experiment 11. It can be seen from this table that there was a good deal of variation among the participants regarding number of blocks of training required to meet criterion for Same responding, but almost no variation among participants regarding number of blocks of testing required to meet criterion for Same responding.

As has been noted in previous experiments in Chapters 3 and 4, there was little variation among participants in the number of blocks of testing required to meet criterion for Same responding. It is clear that even participants with learning difficulties required little intervention across exposures to training, multiple exemplar testing and testing across novel stimulus sets to meet criterion for relational responding in accordance with Same (i.e., Same relational responding). More specifically, five participants (P74, P76, P77, P80 and P81 met criterion for Same relational responding on their first exposure to a test block. These participants required 11, 18, six, ten and 13 blocks of training respectively to meet criterion for training. The remaining participants, P75, P78 and P79, required ten, 13 and 39 blocks respectively of relational training across three novel stimulus sets. All three of these participants required three blocks of testing before reaching criterion for Same relational responding. In all three cases, two of these blocks were relational tests with no feedback (i.e., probe tests), and one was a multiple exemplar test (i.e., with feedback). Taken together, P74, P75, P76, P77, P78, P79, P80 and P81 required a mean of 15 blocks of relational training, a mean of 1.38 blocks of relational testing and a mean of 0.38 blocks of multiple exemplar testing before meeting criterion for Same relational responding.

Participants' performances across the intervention for More-Than/Less-Than relational responding can be seen in Table 32 below.

Table 32. The number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion for More-Than/Less-Than relations on each of the phases in Experiment 11.

Participant Number	Blocks of More/Less pre-training	Novel stimulus sets (not including remedial levels)	Blocks of More/Less relational training	Blocks of More/Less relational testing	Blocks of More/Less Multiple Exemplar testing	Blocks of More/Less remedial training	Blocks of More/Less remedial testing
74	16	3	52	1	0	0	0
75	9	18	43	10	8	8	6
76	8	10	48	6	4	2	2
77	9	14	36	8	6	7	4
78	6	18	36	10	9	8	6
79	10	5	15	3	2	0	0
80	6	7	49	4	3	0	0
81	5	14	41	8	6	10	4

Table 32 shows the number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion for More-Than/Less-Than relations on each of the phases in Experiment 11. It can be seen from this table that there was greater variance among participants' performances for More-Than/Less-Than relational responding than was seen for the same participants for Same relational responding.

Of the participants in this group, one (P74) did not require any multiple exemplar testing or remedial levels. However, it should be noted that P74 was exposed to relational training across three novel stimulus sets without any access to multiple exemplar sets because he reached the maximum allowable exposures to any one stimulus set (i.e., 20 blocks) during training without meeting criterion on two occasions. This participant was also exposed to additional pre-training to ensure that he had control over the contextual cues. Two others (P79 and P80) did not require any remedial levels, but did require access to two and three blocks of multiple exemplar testing respectively. Taken together, after establishing the contextual cues for More-Than/Less-Than relational responding, P74, P75, P76, P77, P78, P79, P80 and P81 required a mean of 40 exposures to relational training blocks and a mean of 6.25 exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding. In addition, P75, P76, P77, P78, P79, P80 and P81 also required a mean of 5.43 multiple exemplar testing blocks to meet this criterion. P75, P76, P77, P78 and P81 were also exposed to a mean of seven remedial training levels and a mean of 4.40 remedial testing levels to meet criterion.

Participants' performances across the intervention for Opposite relational responding can be seen in Table 33 below.

Table 33. The number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion for Opposite relations on each of the phases in Experiment 11.

Participant Number	Blocks of Opposite pre-training	Novel stimulus sets (not including remedial levels)	Blocks of Opposite relational training	Blocks of Opposite relational testing	Blocks of Opposite Multiple Exemplar testing	Blocks of Opposite remedial training	Blocks of Opposite remedial testing
74	6	1	5	1	0	0	0
75	5	10	24	6	4	4	2
76	5	16	49	9	7	6	5
77	6	10	19	6	4	6	2
78	4	7	12	4	3	0	0
79	9	12	30	7	5	5	3
80	17	14	53	8	6	5	4
81	8	16	52	9	7	8	5

Table 33 shows the number of pre-training, training, relational testing and multiple exemplar testing blocks and the number of novel stimulus sets required by each participant to meet criterion for Opposite relations on each of the phases in Experiment 11. It can be seen from this table that there was a large amount of variance in the amount of training to reach criterion that participants required. Less variance was seen in the number of blocks of testing which participants required.

Of the participants in this group, one (P74) did not require any multiple exemplar testing or remedial levels. One other participant (P78) did not require any remedial levels, but did require access to three blocks of multiple exemplar testing. Taken together, after establishing the contextual cues for Opposite and Same relational responding, P74, P75, P76, P77, P78, P79, P80 and P81 required a mean of 30.50 exposures to relational training blocks and a mean of 6.25 exposures to relational testing blocks to meet the pre-established criterion for More-Than/Less-Than relational responding. In addition, P75, P76, P77, P78, P79, P80 and P81 also required a mean of 5.14 multiple exemplar testing blocks to meet this criterion. P75, P76, P77, P78, P79, P80 and P81 were also exposed to a mean of 5.67 remedial training levels and a mean of 3.50 remedial testing levels to meet criterion.

As was seen in previous chapters of this thesis, most participants in the current experiment did not struggle to meet criterion for Same relational responding. In fact some participants (P74, P76, P77, P80 and P81) did not require any multiple exemplar intervention to meet criterion. The remaining participants (P75, P78 and P79) only required one multiple exemplar testing block to meet criterion. Thus, it would appear that even in a population of children with learning difficulties, responding to the relation of Same is already present in their repertoires. This further supports the idea presented in earlier chapters, and indeed widely accepted by RFT theorists, that the relation of Same is a basic relation which is established early in the human verbal repertoire.

More-Than/Less-Than relational responding presented more difficulty for the population of children with learning difficulties than did Same relational responding. This can be seen by the higher number of cycles required by participants to meet criterion for training and testing during the intervention, and also by the fact that more participants required access to multiple exemplar sets and to remedial levels. Performance on Opposite relational responding during the intervention for participants is roughly comparable to their performance for More-Than/Less-Than relational responding. This strongly suggested that the acquisition of Opposite relational responding may have been facilitated by the More-Than/Less-Than relational responding intervention. More specifically, normally functioning participants in earlier chapters all showed greater differences between their More-Than/Less-Than relational responding skill and Opposite relational responding skill levels. It was speculated, therefore, that Opposite relational skills should be targeted last in any intervention so that other factors involved in fluent relational responding (e.g., attention, control by social reinforcers) could be established first across less demanding relational tasks. The finding that so many children functioned almost equally well on Opposite and More-Than/Less-Than relational tasks suggests that this strategy may have been worthwhile. Of course, we do not yet know all the component skills that might be required to acquire Opposite relational responding repertoires. Some requisite skills may involve attention and responding to task format and even relating in its most basic sense (i.e., A-->B). It is presumed that all of these skills are required to acquire and develop relational repertoires. Thus, the current participant pool may experience deficits in these general repertoires, as well as their specific relational skill repertoires. Furthermore, it is possible that some of these pre-requisite skills were established or strengthened during the Same and More-Than/Less-Than interventions. Table 34 and Table 35 below show participants' total raw and mean RAI scores for Same, More-Than, Less-Than and Opposite both before and after the intervention.

Table 34. Participant's total raw RAI scores for Same, More-Than, Less-Than and Opposite relational responding at Time 1 and Time 2.

P	RAI Level	Time 1 Same	Time 2 Same	Time 1 More	Time 2 More	Time 1 Less	Time 2 Less	Time 1 Opposite	Time 2 Opposite
74	1	9	**20**	12	**13**	13	**13**	12	**20**
	2	19	**20**	11	**10**	2	**17**	12	**20**
	3	18	**20**	11	**11**	3	**15**	18	**17**
75	1	8	**16**	12	**20**	16	**15**	16	**20**
	2	3	**20**	10	**17**	20	**18**	17	**20**
	3	7	**20**	6	**15**	14	**20**	16	**20**
76	1	9	**20**	8	**16**	14	**11**	10	**19**
	2	14	**20**	9	**18**	18	**17**	11	**18**
	3	11	**20**	6	**19**	12	**19**	11	**12**
77	1	8	**20**	17	**20**	16	**19**	15	**17**
	2	5	**20**	20	**20**	20	**20**	18	**20**
	3	9	**19**	18	**19**	17	**19**	19	**19**
78	1	20	**18**	20	**20**	18	**20**	19	**18**
	2	2	**20**	16	**16**	19	**20**	19	**20**
	3	1	**19**	16	**19**	18	**19**	8	**20**
79	1	7	**18**	11	**20**	7	**20**	7	**20**
	2	5	**20**	18	**20**	10	**16**	10	**19**
	3	15	**20**	18	**17**	12	**15**	12	**20**
80	1	8	**18**	9	**20**	12	**18**	13	**20**
	2	11	**20**	12	**19**	10	**19**	11	**20**
	3	9	**20**	10	**20**	12	**14**	10	**19**
81	1	3	**19**	4	**20**	18	**20**	13	**20**
	2	11	**20**	2	**19**	0	**19**	20	**20**
	3	13	**20**	0	**18**	1	**19**	2	**20**

Note. The letter "P" in the first column represents "participant number". Time 1 scores are listed in black font. Time 2 scores are listed in red font.

Table 34 shows participant's total raw RAI scores out of 20 for Same, More-Than, Less-Than and Opposite relational responding at Time 1 and Time 2. It can be seen from this table that most participants' relational skills for Same, More-Than/Less-Than and Opposite at each RAI level have risen from Time 1 to Time 2. This can be seen even more clearly in Table 35 in which participant's raw RAI scores have been averaged to provide one RAI score for each relation. Mean RAI scores at Time 1 and Time 2 for each relation examined are shown in Table 35 below.

Table 35. Participants' mean baseline RAI scores for Same, More. Less and Opposite relations at Time 1 and Time 2.

P	Time 1 Mean Same	Time 2 Mean Same	Time 1 Mean More	Time 2 Mean More	Time 1 Mean Less	Time 2 Mean Less	Time 1 Mean Opposite	Time 2 Mean Opposite
74	15.33	**20**	11.33	**11.33**	6	**15**	14	**19**
75	6	**18.67**	9.33	**17.33**	16.67	**17.67**	16.33	**20**
76	11.33	**20**	7.67	**17.67**	14.67	**15.67**	10.67	**16.33**
77	7.33	**19.67**	18.33	**19.67**	17.67	**19.33**	17.33	**18.67**
78	7.67	**19**	17.33	**18.33**	18.33	**19.67**	15.33	**19.33**
79	9	**19.33**	15.67	**19**	9.67	**17**	9.67	**19.67**
80	9.33	**19.33**	10.33	**19.67**	11.33	**17**	11.33	**19.67**
81	9	**19.67**	2	**19**	6.33	**1933**	11.67	**20**

Note. The letter "P" in the first column represents "participant number". Time 1 scores are listed in black font. Time 2 scores are listed in red font.

Table 35 shows participants' mean baseline RAI scores for Same, More, Less and Opposite relations at Time 1 and Time 2. The data in this table clearly show that the intervention has been successful in improving participant's derived relational responding skill as measured by the RAI. With the exception of P74, whose mean RAI score for More-Than relations stayed the same, every other RAI score for every participant for each relation measured rose from Time 1 to Time 2. For the relation of Same, an average mean rise of 10.09 RAI points was calculated. For the relation of More-Than, an average mean rise of 6.25 RAI points was calculated. For the relation of Less-Than, an average mean rise of five RAI points was calculated. Finally, for the relation of Opposite, an average mean rise of 5.79 RAI points was calculated. (See Appendix 32 for the mean rises for each participant for each relation which contributed to the average mean scores). It remained to be seen whether or not improving participants' relational abilities would have an impact on their full scale IQ scores or on any other the subtest or index measures.

Table 36 shows each individual participant's IQ Index scores and the subtests which contribute to them at Time 1 and at Time 2. Table 36 also shows the differences in each subtest and index from Time 1 to Time 2.

Table 36: IQ subtest scores, Full scale IQ scores, Verbal IQ scores and Performance IQ scores, for each participant at baseline and follow-up.

Participant: 74

Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	**Difference from Time 1 to Time 2**
Block Design			4	**9**					**+5**
Similarities	7	**8**							**+1**
Digit Span					10	**12**			**+2**
Picture Concepts			4	**11**					**+7**
Coding							5	**9**	**+4**
Vocabulary	8	**9**							**+1**
Letter-Number Sequencing					10	**2**			**-8**
Matrix Reasoning			5	**7**					**+2**
Comprehension	7	**8**							**+1**
Symbol Search							9	**11**	**+2**
Sum of Scaled Scores	22	**25**	13	**27**	20	**14**	14	**20**	
VC Scale Composite Scores	85	**91**							**+6**
PR Scale Composite Score			65	**94**					**+29**
WM Scale Composite Score					99	**83**			**-16**
PS Scale Composite Score							83	**100**	**+17**
FSIQ Composite Score	77	**89**							**+12**

Note. Time 1 scores are listed in black font. Time 2 scores are listed in red font. The differences in scores from Time 1 to Time 2 are listed in blue font. In this table, "VC" represents "Verbal Comprehension", "PR" represents "Perceptual Reasoning", "WM" represents "Working Memory", "PS" represents "Processing Speed" and "FSIQ" represents "Full Scale IQ".

The remaining participants are now presented by their respective participant numbers. The table continues with the same format as was presented for P74.

Participant: 75

Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	**Difference from Time 1 to Time 2**
Block Design			6	**9**					**+3**
Similarities	9	**11**							**+2**
Digit Span					11	**11**			**0**
Picture Concepts			10	**11**					**+1**
Coding							12	**12**	**0**
Vocabulary	8	**9**							**+1**
Letter-Number Sequencing					10	**9**			**-1**
Matrix Reasoning			7	**9**					**+2**
Comprehension	8	**11**							**+3**
Symbol Search							9	**14**	**+5**
Sum of Scaled Scores	25	**31**	23	**29**	21	**21**	21	**26**	
VC Scale Composite Scores	91	**100**							**+9**
PR Scale Composite Score			86	**98**					**+12**
WM Scale Composite Score					102	**102**			**0**
PS Scale Composite Score							103	**118**	**+15**
FSIQ Composite Score	92	**105**							**+13**

Participant: 76									
Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	Difference from Time 1 to Time 2
Block Design			7	**9**					+2
Similarities	5	**7**							+2
Digit Span					11	**7**			-4
Picture Concepts			4	**6**					+1
Coding							6	**5**	-1
Vocabulary	7	**8**							+1
Letter-Number Sequencing					6	**5**			-1
Matrix Reasoning			7	**7**					0
Comprehension	4	**7**							+3
Symbol Search							8	7	-1
Sum of Scaled Scores	16	**22**	18	**22**	17	**12**	14	**12**	
VC Scale Composite Scores	73	**85**							+12
PR Scale Composite Score			75	**84**					+9
WM Scale Composite Score					91	**77**			-14
PS Scale Composite Score							83	**78**	-5
FSIQ Composite Score	74	**76**							+2

Participant: 77

Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	Difference from Time 1 to Time 2
Block Design			7	**7**					0
Similarities	7	**12**							+5
Digit Span					11	**15**			+4
Picture Concepts			9	**11**					+1
Coding							8	**14**	+6
Vocabulary	10	**12**							+2
Letter-Number Sequencing					11	**11**			0
Matrix Reasoning			8	**9**					+1
Comprehension	9	**12**							+3
Symbol Search							8	**12**	+4
Sum of Scaled Scores	26	**36**	24	**27**	22	**26**	16	**26**	
VC Scale Composite Scores	93	**110**							+17
PR Scale Composite Score			88	**94**					+6
WM Scale Composite Score					104	**116**			+8
PS Scale Composite Score							88	**118**	+30
FSIQ Composite Score	90	**111**							+21

Participant: 78

Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	Difference from Time 1 to Time 2
Block Design			7	**8**					+1
Similarities	6	**8**							+2
Digit Span					3	**11**			+8
Picture Concepts			9	**13**					+4
Coding							7	**9**	+2
Vocabulary	5	**6**							+1
Letter-Number Sequencing					3	**10**			+7
Matrix Reasoning			6	**8**					+2
Comprehension	5	**10**							+5
Symbol Search							7	**8**	+1
Sum of Scaled Scores	16	**24**	22	**29**	6	**21**	14	**17**	
VC Scale Composite Scores	73	**89**							+16
PR Scale Composite Score			84	**98**					+14
WM Scale Composite Score					59	**102**			+43
PS Scale Composite Score							83	**91**	+8
FSIQ Composite Score	70	**92**							+22

Participant: 79									
Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	Difference from Time 1 to Time 2
Block Design			8	**8**					0
Similarities	8	**10**							+2
Digit Span					16	**12**			-4
Picture Concepts			10	**11**					+1
Coding							10	**14**	+4
Vocabulary	7	**9**							+2
Letter-Number Sequencing					10	**10**			0
Matrix Reasoning			7	**5**					-2
Comprehension	5	**9**							+4
Symbol Search							7	**11**	+4
Sum of Scaled Scores	20	**28**	25	**24**	26	**22**	17	**25**	
VC Scale Composite Scores	90	**99**							+9
PR Scale Composite Score			90	**88**					-2
WM Scale Composite Score					116	**104**			-8
PS Scale Composite Score							91	**115**	+24
FSIQ Composite Score	90	**99**							+9

Participant: 80									
Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	Difference from Time 1 to Time 2
Block Design			5	**8**					+3
Similarities	6	**8**							+2
Digit Span					9	**11**			+2
Picture Concepts			10	**10**					0
Coding							12	**13**	+1
Vocabulary	6	**6**							+2
Letter-Number Sequencing					8	**9**			+1
Matrix Reasoning			2	**9**					+7
Comprehension	8	**7**							-1
Symbol Search							11	**14**	+3
Sum of Scaled Scores	20	**21**	17	**27**	17	**20**	23	**27**	
VC Scale Composite Scores	81	**83**							+2
PR Scale Composite Score			73	**94**					+11
WM Scale Composite Score					91	**99**			+8
PS Scale Composite Score							109	**121**	+11
FSIQ Composite Score	83	**96**							+13

Participant: 81									
Subtests	VC Index Scaled Scores at Time 1	**VC Index Scaled Scores at Time 2**	PR Index Scaled Scores at Time 1	**PR Index Scaled Scores at Time 2**	WM Index Scaled Scores at Time 1	**WM Index Scaled Scores at Time 2**	PS Index Scaled Scores at Time 1	**PS Index Scaled Scores at Time 2**	Difference from Time 1 to Time 2
Block Design			11	**11**					0
Similarities	7	**7**							0
Digit Span					10	**9**			-1
Picture Concepts			8	**9**					+1
Coding							9	**14**	+5
Vocabulary	6	**7**							+1
Letter-Number Sequencing					9	**10**			+1
Matrix Reasoning			9	**13**					+4
Comprehension	7	**8**							+1
Symbol Search							7	**11**	+4
Sum of Scaled Scores	20	**22**	28	**33**	19	**19**	16	**25**	
VC Scale Composite Scores	81	**85**							+4
PR Scale Composite Score			96	**106**					+10
WM Scale Composite Score					97	**97**			0
PS Scale Composite Score							88	**115**	+27
FSIQ Composite Score	87	**99**							+12

It can be seen from Table 36 that there were rises as well as falls for each participant from Time 1 to Time 2. Table 37 below outlines the rises and falls for each participant for each subtest measure. These changes from Time 1 to Time 2 are then discussed.

Table 37. Total changes for each participant from Time 1 to Time 2 for all subtest and index scores.

Subtests	P74	P75	P76	P77	P78	P79	P80	P81	Mean change from Time 1 to Time 2
Block Design	+5	+3	+2	0	+1	0	+3	0	+1.75
Similarities	+1	+2	+2	+5	+2	+2	+2	0	+2
Digit Span	+2	0	-4	+4	+8	-4	+2	-1	+.88
Picture Concepts	+7	+1	+1	+1	+4	+1	0	+1	+2
Coding	+4	0	-1	+6	+2	+4	+1	+5	+2.63
Vocabulary	+1	+1	+1	+2	+1	+2	+2	+1	+1.38
Letter-Number Sequencing	+8	-1	-1	0	+7	0	+1	+1	+1.88
Matrix Reasoning	+2	+2	0	+1	+2	-1	+7	+4	+2.13
Comprehension	+1	+3	+3	+3	+5	+4	-1	+1	+2.38
Symbol Search	+2	+5	-1	+4	+1	+4	+3	+4	+2.75
VC Scale Composite Scores	+6	+9	+12	+17	+16	+9	+2	+4	+9.38
PR Scale Composite Score	+29	+12	+9	+6	+14	-2	+11	+10	+11.13
WM Scale Composite Score	-16	0	-14	+8	+43	-8	+8	0	+2.63
PS Scale Composite Score	+17	+15	-5	+30	+8	+24	+11	+27	+15.88
FSIQ Composite Score	+12	+13	+2	+21	+22	+9	+13	+12	+13

Note. "VC" represents "Verbal Comprehension", "PR" represents "Perceptual Reasoning", "WM" represents "Working Memory", "PS" represents "Processing Speed" and "FSIQ" represents "Full Scale IQ".
The mean change is listed in blue font.

Table 37 shows the total changes for each participant from Time 1 to Time 2 for all subtest and index scores. It can be seen from this table that there were some rises as well as falls in subtest and index scores from Time 1 to Time 2. However, when the averages are calculated, all changes in subtest and index scores were in a positive direction. More specifically, from Time 1 to Time 2, all participants showed an increase in subtest and index scores.

In considering the mean subtest changes from Time 1 to Time 2, it can be seen that the rises range from +.88 (Digit Span) to +2.75 (Symbol Search). When the scaled scores are added, it can be seen that the rises range from +2.63 points (Working Memory Scale) to +15.88

(Processing Speed Scale), with Verbal Comprehension (+9.38) and Perceptual Reasoning (+11.13) in the middle. These subtest and scaled score rises all contribute to a mean Full Scale IQ rise of +13 points. Given that full scale IQ rises were seen for all experimental participants, and not for control participants, in Chapter 5, it was not unexpected that participant's full scale IQ scores would rise following a relational intervention. It was perhaps expected that the IQ rises might be higher than were seen with the applied population in the current chapter. However, it should be noted that with any special population with varying ability levels, more variance in ability is perhaps expected than with a normally functioning population with no known or suspected learning difficulties. This is likely due to the wider range of potential deficits in such areas as attention and task completion, as well as whole host of other pre-requisite skills needed for taking part in any academic task.

It is not surprising that there were rises from Time 1 to Time 2 in both the Comprehension subtest and in the Verbal Comprehension Scale given that the strongest correlations were seen between the mean scores on the Opposite relation responding task and these scores on the WISC-IV^{UK}. However, it should also be noted that rises were seen for mean scores on all subtests and all index scores. Thus, it appears that the relational intervention has improved not only the targeted relational skills, but also has had a positive impact on all subtest and scale scores on an IQ test. The Working Memory Scale was shown to rise by the least average amount with a rise of 2.63 points. The other scales showed much larger rises, with the Verbal Comprehension Scale showing an average rise of 9.38 points, the Perceptual Reasoning Scale showing an average rise of 11.13 points, and the Processing Speed Scale showing an average rise of 15.88 points. While all scales showed some improvement, it is perhaps surprising that the Processing Speed Scale was affected consistently and to a larger extent than other scales.

The reader is reminded that the Processing Speed Scale is composed of subtests which are thought to measure the speed of an “individual’s mental and graphomotor processing” (see Wechsler, 2004, p. 4). Skills in this domain on their own make a contribution to a child’s intellectual development. Understanding how precisely the relational interventions moved the repertoire in this experiment will require careful experimental analysis. However, it seems reasonable to suggest at this point that the large battery of relational tasks administered across a range of stimuli, would have increased participants relational flexibility. That is, their ability to respond to relational tasks and to acquire new relational repertoires more generally was improved. In effect, the intervention likely functioned as a type of “learning to learn” intervention that improved some of the very general skills of learning itself. This suggestion is in line with Harlow’s (1949) proposition that humans and other highly intelligent animals not only mastered isolated tasks but also noticed patterns and shortcuts that made them more efficient learners. They not only *learned*, they *learned to learn*, becoming faster at solving new problems as they gained experience solving similar classes of problems (see Harlow, 1949). In this work, Harlow also suggested that in “learning to learn” a subject picks out a pattern in a series of learning experiences, so that the subject can learn even faster when facing similar situations in the future. An example in humans would be learning how to study correctly for a class. One might find that they do better as they go along, because they have "learned how to learn" in that class. The current analysis contributes to our understanding of this phenomenon by adding the concept of foundational relational skills which themselves require a honing of many basic learning skills (e.g., attention).

The foregoing suggestions are also in line with O’Toole et al. (in press) who have posited that using the IRAP is advantageous for precisely the reason that it measures relational flexibility. O'Toole et al. also note that the IRAP requires two patterns of responding, one consistent and the other inconsistent, with previously established relations. The assumption is that participants will respond more quickly on consistent trials as responding inconsistently is not a well established or practiced skill. Furthermore, O'Toole and colleague go on to suggest that response speed on trials may provide a useful measure of relational or cognitive flexibility, which is widely regarded to be an important component of human cognitive abilities. In fact, O'Toole et al. found that faster responding on an IRAP, and smaller difference scores between consistent and inconsistent trials, predicted higher IQ scores on the K-BIT. In light of the foregoing, it is possible that in the current research it may not be that relational ability does not correlate strongly with IQ, but rather that we have not yet investigated the right relational ability. In addition, it may be that it is not one relation specifically that is critical, but rather it may be relational flexibility itself, and not the extent of the repertoires which already exist or are established, that is the critical component in an intervention.

Many questions remain to be addressed. Specifically, O'Hora et al. (2005) have noted significant correlations between response accuracy on a complex relational skill and verbal subtests of the WAIS. So, perhaps it is possible to suggest that an even more complex relational skill than those examined in the current context would need to be employed in an intervention format before results similar to those of O'Hora et al. (2005) could be found with an applied population. Given that the experiments conducted in the current chapter alone required an entire academic year to complete, time constraints did not permit examination of each and every question that arose during the course of one doctoral research programme. It is hoped that such questions will drive future research studies.

General Discussion

In Experiment 10a of the current chapter, it was seen that the RAI did appear to be distinguishing broadly between levels of behavioural demand across the various relational repertoires. Although it did not appear that the normally developing participants employed for Experiment 10a would require a relational intervention for these relational skills, the purpose of the pilot test was merely to see that the RAI could in principle establish a child's baseline level of relational abilities against which improvements in relational repertoires could be assessed. It was also to ensure that the experimental instructions could be clearly understood by children, and to generally beta-test the specifically developed software. At the completion of Experiment 10a, it appeared that the RAI provided enough inter-subject and inter-relation variability above zero and below 100% accuracy to be useful as a measure of assessing relational skills pre-intervention in a group of children with various educational disadvantages.

In Experiment 10b, a group of children presenting with learning difficulties were recruited and administered the WISC-IV^{UK} and four individual RAI tests for the relations of Same, More-Than, Less-Than and Opposite. Correlational analyses were then conducted and these analyses revealed that there were correlations between many of the subtests on the WISC-IV^{UK}, and mean RAI scores for each relation. The strongest correlation (and the only significant one) was seen between the mean Opposite RAI score and the Comprehension subtest on the WISC-IV^{UK}. Given that Comprehension is one of the subtests which contribute to the Verbal Comprehension Scale, this result was perhaps not surprising. Furthermore, the Verbal Comprehension Scale scores were also strongly correlated with mean Opposite RAI scores. The Verbal Comprehension Index is thought to measure verbal abilities utilising reasoning, comprehension and conceptualisation (see Wechsler, 2004). These findings, therefore, support the RFT position that a functional relationship exists between verbal skills, as measured by standardised IQ tests items, and relational skills.

Several experiments in the current thesis have pointed to the relational complexity of Opposite relational responding for both adults and children, compared to Same and More-Than/Less-Than relations. Given these earlier findings, it was not surprising that the large variances in fluency in Opposite relational responding typically observed across the population should be correlated with a measure of general relational ability such as that presented by standard IQ tests. The fact that the strongest correlations found in this domain were in an expected direction and were observed between RAI scores and *verbal task* performance provides further support for the RFT suggestion that there exists a functional relationship between verbal skills, as measured by standardised IQ tests items, and relational skills.

It must also be considered at this stage, that only relational skill levels which show a large variance across the population and which are at a relatively low level of development for most individuals can be used as relational ability indices. These very relational abilities may also be those that correlate best with various intelligence measures. Of course, these correlations may be arrived at by painstaking correlational studies across an infinite number of relational skills and an almost infinite number of intelligence measures, across various populations. However, the current research, as a basic research enterprise, first attempted to establish the functional relationships between various relations, and this in turn, could now be said to have pointed to the most likely relational skill of those studied to be used as a viable RAI index for general intelligence.

In Experiment 10b, it must be noted that many other correlations were seen, in both a positive (and expected) direction, as well as in a negative direction. However, as was seen in

Chapter 5, the strongest correlations tended to be the ones that were expected from an RFT perspective (i.e., tasks in the verbal domain). These findings were once again consistent with the research findings of O'Hora et al. (2005) which supported the supposition that correlations between some subtests (e.g., verbal subtests) and DRR ability would be stronger than correlations between other subtests (e.g., performance subtests) and DRR ability. Nevertheless, there was a relative ambiguity in the ranges of correlations observed. This ambiguity was noted for RAI scores for Same, More-Than and Less-Than relations. It was concluded that it is entirely possible that even with a population of participants who are struggling in school, that a Same relational repertoire is already well established and therefore, sufficiently well developed for most individuals that it bears little relevance to IQ. Interestingly, fewer correlations in an unexpected direction were observed between the individual subtest and scale scores with mean More-Than RAI scores. This supports the foregoing suggestion insofar as this relational repertoire is less well developed among school children, and skills levels in this domain might be expected to be roughly distributed across this population in a manner related to general intellectual ability. It is perhaps relevant also that the positive correlations between More-Than RAI scores were amongst the strongest of all the correlations observed (i.e., stronger than the negative correlations). However, in considering the pattern of correlations seen between the individual subtest and scale scores with mean Less-Than RAI scores, the pattern was once again unclear. Overall, fewer correlations were seen between the individual subtest and scale scores with mean Less-Than RAI scores than with any other mean RAI score. The correlations seen were all mild. Two were in an expected direction (i.e., positive correlations), and four were in an unexpected direction (i.e., negative correlations). It has been speculated that it is entirely possible that relations of Less-Than simply do not correlate meaningfully with IQ subtests or scales.

It should also be noted at this point that although the participants recruited for Experiment 10b were all presenting with learning difficulties, the learning difficulties they presented with differed. For example, some participants presented with diagnosed general learning disabilities (i.e., measured IQ lower than 80) which might affect them in all learning settings. Others presented with early literacy difficulties which would not necessarily have any impact on measured IQ, but would certainly impact learning tasks which required them to read, and perhaps tasks which required them to sequence events. Others presented with difficulties with their mathematics curriculum for which they attended learning support teaching. Other participants presented with attention difficulties (some diagnosed and some merely suspected). Still others presented with reported difficulties in motivation. Thus, there existed a wide range of individual differences in the population of participants who took part in Experiment 10b (and 11). These individual variances in pre-existing IQ and learning ability likely accounted for some of the lack of correlations between relational ability and full scale IQ and its subtests. These individual variances are of course recognised as being a potential limitation of this current research. Future research would likely benefit from including a population sample of children with more similar learning difficulties at the start of such an intervention programme. It may emerge that RFT-based interventions are more helpful in the context of specific deficits. Indeed, this is only to be expected.

Experiment 11 retained the participant pool from Experiment 10b and administered the intervention which had been developed across Chapters 3 and 4. As described earlier, amendments included that the intervention was broken up into three parts. The first part targeted

Same relational responding. The second portion targeted More-Than/Less-Than relational responding, and the third part targeted Opposite relational responding. An equivalence intervention was not employed for three reasons. The first reason was that an equivalence intervention was deemed the functional equivalent of an intervention for Same relations and thus was unlikely to be necessary. The second reason was that the equivalence intervention employed in Chapter 2 was administered in a different format and thus, to exclude the equivalence intervention would make the intervention more fluent. The third reason was that it was decided that administering the intervention in the foregoing way would act as a time saving measure. This was important given the intensity and length of the intervention.

The results of Experiment 11 showed that all participants' mean RAI scores improved from Time 1 to Time 2. Experiment 11 also showed that all participants showed an increase in full scale IQ from Time 1 to Time 2. Aside from the general functional dependence of IQ on relational ability, there are more specific possible reasons for the rises seen in RAI scores and in IQ scores observed. For instance, the sequence of relational training employed may have been particularly effective for improving relational skills. In Experiment 11, the sequence involved a relational training intervention for Same relations first, followed by a More-Than/Less-Than relational intervention and finally, an intervention for Opposite relational responding. This sequence arose from suggestions in the discussion of Experiments 6 and 7 in Chapter 4 where it was decided to settle upon what appeared to be a "best fit" intervention for the current purposes. Given the time constraints in the current work, this practical measure was employed in order to at least make some positive impact on the lives of the participants as soon as possible. This practical measure also aimed to train relations in an order of increasing complexity so as to reduce any possible "punishing" effects of persistent failure early in an intervention. It was speculated that this effort might ensure that other factors involved in fluent relational responding (e.g., attention, control by social reinforcers) could be established first across less demanding relational tasks. It was hoped that this measure might also decrease the potential for participant's "dropping out" if the participants, their parents or teachers felt that the intervention was too difficult for them. Of course, RFT would place little emphasis on identifying the "correct" sequence in any case, but would emphasise the importance of finding a workable and maximally effective sequence, which may in turn vary across situations, and participant groups, depending on the particular deficits and relevant biological variables. However, at the conclusion of the experiments in the current chapter, it appears that the "best-fit" sequence was indeed maximally effective. More specifically, participants in Experiment 11 demonstrated rapid acquisition of Opposite relational responding comparable to their performance for More-Than/Less-Than relational responding. This strongly suggested that the acquisition of Opposite relational responding may have been facilitated by the More-Than/Less-Than relational responding intervention. In earlier chapters, normally functioning participants all showed greater differences between their More-Than/Less-Than relational responding skill and Opposite relational responding skill levels. The finding that so many children functioned almost equally well on Opposite and More-Than/Less-Than relational tasks suggests that this strategy may have been worthwhile.

The more frequent use of remedial levels employed in Experiment 11 may also have contributed to the dramatic rise in RAI and IQ scores. Berens and Hayes (2007) suggested the use of such a strategy as a means of facilitating the derivation of arbitrary relational responding when it is absent or weak. Berens et al. proposed that the use of non-arbitrary stimuli, as well as arbitrary stimuli, in training and testing protocols might be effective in such an endeavour (see also Y. Barnes-Holmes, D. Barnes-Holmes, Roche, & Smeets, 2004; Steele & Hayes, 1991). In Experiment 11, it appears that employing such a strategy increased the generalisation of control by the contextual cues by being systematically applied across both arbitrary and non arbitrary stimulus relations.

One obvious factor that is relevant to the relational intervention that may have lead to rises in RAI and IQ scores must be considered at this point. Specifically, Experiment 11 did not employ a control group for ethical reasons. That is, the experimenter did not wish to deprive any participants of an intervention that she believed would make an impact on the IQ scores of children in need of educational intervention. Thus, it is difficult to know what the effect of mere practice at relational tasks in the absence of MET may have been. Moreover, it is impossible to know what the effect of attention from, and interaction with, the experimenter in the absence of any relational training whatsoever would have been. It should be remembered, however that previous experiments did employ control groups, albeit in the context of different procedures, which established rather convincingly that the MET interventions reliably lead to improvement in relational abilities over a control procedure. Experiments 2 and 9 also showed that a MET intervention can raise IQ above those rises seen typically across time. Nevertheless, the possibility remains that a significant portion of the improvements observed in Experiment 11 can be explained by practice effects, and the improvements in general attending or other skills that may have resulted from any similar interaction with the participants.

Many questions remain regarding the various score changes across time observed in the current intervention. For instance, given that scores on the Comprehension subtest were correlated with mean RAI scores for the relation of Opposite, it would be reasonable to assume that a relational intervention which included the relation of Opposite would increase a participants score on the Comprehension subtest. While rises in Comprehension subtest scores certainly were seen, rises were also seen for all other subtests. Perhaps this suggests that all of the relations are functionally related in some as yet undiscovered way. The possible complexity of these inter-relations is infinite, and speculation at this point would necessarily be interpretive and theoretically deductive to the point of being almost useless. Indeed, RFT researchers have consistently emphasised the inductive approach over interpretative approaches. In fact, it is precisely because Skinner's book, *Verbal Behavior* (1957), was so interpretative that it was difficult to turn into an empirical or inductive research programme. In that work, Skinner's definition of verbal behaviour was not functional and was too broad (see Hayes, Blackledge, & Barnes-Holmes, 2001), and thus failed to yield a vibrant body of research. In order to discover the ways in which the various functional relations between relational repertoires are organised, the relational interventions for Same, More-Than/Less-Than and Opposite would have to be run separately using separate groups of participants. This was not appropriate to the current research which had as its aim to pilot test a basic relational MET intervention on a population in need.

The data in the current chapter provide further support that responding relationally is indeed an operant phenomenon. DRR skills have been shown to improve across the course of a targeted intervention. Improving these DRR skills with such an intervention has also led to rises in measured IQ for both a normally developing population of children (Chapter 5), as well as a population of children with learning difficulties (Chapter 6). It is worth briefly reiterating these rises for the reader.

For the normally developing participants in Chapter 5, the rises in full scale IQ were so large that they could "move" a participant from one qualitative diagnostic range to another. In fact, some of the rises seen in Chapter 5 were so large that they constituted a change of more than one standard deviation (i.e., 15 points on index scaled scores). For example, from Time 1 (in Chapter 2) to Time 2 (in Chapter 5) experimental participant, P1, presented with a full scale IQ score rise of 33 points. This was a rise of more than two standard deviations. Experimental participant, P2, presented with a full scale IQ score rise of 18 points from Time 1 to Time 3, which constitutes a rise of more than one standard deviation. Experimental participant, P3, presented with a full scale IQ rise of 26 points from Time 1 to Time 3 which constitutes a rise of nearly two standard deviations. Finally, experimental participant, P8, presented with a full scale IQ rise of 32 points (from Time 1 to Time 3), which was a rise of more than two standard deviations.

The rises seen in full scale IQ for children presenting with learning difficulties in the current chapter were perhaps not so dramatic. Nevertheless, rises were seen for all participants. Full scale IQ rises in the current chapter included a two point rise (P76), a nine point rise (P79), two 12 point rises (P74, P81), two 13 point rises (P75, P80), a 21 point rise (P77) and a 22 point rise (P78). While only two of these rises were larger than a standard deviation (i.e., 15 points), six of the rises were large enough to yield a “move” from one qualitative range to another (i.e., more than ten points), and one approached such a move (i.e., P79, nine point rise). In fact, only one participant (P76) yielded no real changes from pre to post intervention. Generating significant improvements in full scale IQ scores for seven out of eight participants experiencing learning difficulties in the current experiment should not be underestimated. In addition, even though one participant did not appear to show any major improvement in full scale IQ, all participants showed improvements in relational skill levels as measured by the RAI from Time 1 to Time 2. Thus, it can be said that the relational intervention was indeed successful in generating improvements in relational repertoires. There are, of course, several questions remaining regarding the precise source of the various rises in subtests scores seen, and the precise process according to which the relational intervention led to full scale IQ rises. Nevertheless, at this point the effect on IQ of relational interventions of the kind used here can be said to be clearly established.

Chapter 7

General Discussion

Chapter 1 reviewed how behaviour analysts working within the field of Relational Frame Theory (RFT) have recently begun to re-examine the concept of intelligence and have suggested a theoretical framework for the analysis of those behaviours referred to as intelligent by psychometricians. Specifically, it was outlined that RFT theorists have claimed that the foundational skill for most intellectual abilities is derived relational responding (DRR). A small literature base that supports the claim that relational skills are foundational to language was also outlined (e.g., Barnes, McCullagh, & Keenan, 1990; Devany, Hayes, & Nelson, 1986; Lipkens, Hayes, & Hayes, 1993). In addition, it was shown that correlations have been found between specific relational skills (e.g., forms of rule-following involving similar/different and temporal relations and subtests on commonly used IQ tests. See O'Hora et al., 2005 and O'Toole et al., in press).

Chapter 1 also illustrated how, from a RFT perspective, many IQ test items can be conceived in relational terms that can serve as operational definitions for further research. In the field of RFT, defining relational skills, and intellectual skills, functional-analytically has led to an understanding of intelligence in which it is distinguishable from relational skill only in topographical terms. In effect, it was suggested that a functional relationship, or even considerable functional overlap, likely exists between relational skills and intellectual skills, insofar as each behavioural repertoire was perhaps only a topographical variation of the other. This, in turn, led to the crucial suggestion that an improvement in DRR repertoires may well produce a measurable change in language ability (i.e., acquisition rates, fluency, and extent of vocabulary), and even general intelligence. Chapter 1 went on to provide a rationale for the construction of interventions to raise intelligence quotients as calculated by standardised IQ tests or on specific dimensions, or subtests, of IQ. Specifically, it was proposed in Chapter 1 that training skills in derived relational responding (DRR) by utilising multiple exemplar training (MET) might accomplish this goal.

Importantly, it was also noted in Chapter 1 that other factors, such as attentional skills, the absence of sensory deficits and other diagnosed behavioural and emotional difficulties, may likely also play an important role in whether or not RFT-based interventions are effective. Thus, the effects of relational interventions may not be linear or easily predictable without understanding their relationship to a whole host of other important educational, social, biological, and psychological variables that have been studied by behaviour analysts for the past fifty years or more (see Cooper et al., 2007). Only the efforts made by researchers to address these research issues will help us to determine whether or not the RFT approach to intellectual deficit will be sufficiently useful in making a real difference to the relational repertoires, educability and lives of those who most need our assistance.

Chapter 2 attempted to investigate any possible correlations which might exist between IQ subtests and scales (verbal, performance and full) and basic relational skills (symmetry, transitivity, equivalence). This was, in part, an attempt to replicate some of the findings of O’Hora et al. (2005) with child participants. To this end, twelve participants were administered the WISC-III^{UK}, and measures of their symmetry responding and transitivity responding fluency (Experiment 1) were taken. The measured symmetry and transitivity fluency levels were combined to yield a measure of each participant’s ability to form equivalence relations. A Spearman’s Rho correlational analysis was then conducted to investigate possible correlations between these relational skill levels and full scale IQ and its subtests.

The results of Experiment 1 indicated that correlations were as expected for symmetry testing on verbal and Arithmetic subtests, rather than for tests for spatial relations or concrete manipulation of objects. The findings from Experiment 2 indicated that, as expected, there was a significant (at the 0.05 level, 2 tailed) negative correlation between the Vocabulary subtest and symmetry testing. It thus appeared that symmetry is a potentially useful index of IQ subtest scores. The strong and significant correlation found between the Vocabulary subtest and symmetry testing provided further support that there exists a functional relationship between verbal skills, as measured by standardised IQ tests items, and relational skills. This finding also supported the findings of O'Hora et al. (2005) and, more recently, of O'Toole et al. (in press). It was not surprising that, of all the correlations which emerged in Experiment 1, this was one of the strongest. Surprisingly, results also indicated that there was a significant (0.05 level, 2 tailed) negative correlation between the Symbol Search subtest and number of blocks to meet criterion on the symmetry test. This was unexpected as the Symbol Search subtest is categorised in the performance, rather than the verbal, section of the WISC-III^{UK}. However, O'Toole et al. (in press) have recently suggested that response speed (on specific trials of a complex relational responding task known as the Implicit Relational Assessment Procedure, or IRAP; see also D. Barnes-Holmes, Y. Barnes-Holmes, Power, Hayden, Milne & Stewart, 2006; McKenna, D. Barnes-Holmes, Y. Barnes-Holmes, & Stewart, 2007) may provide a useful measure of relational or cognitive flexibility. O'Toole et al. have also pointed out that assessing relational flexibility may be particularly advantageous because flexibility is widely regarded to be an important component of human cognitive abilities (e.g., Cattell, 1971; Kyllonen, Lohman, & Woltz, 1984; Premack, 2004; see also O'Toole at al., in press, for extended discussion on relational flexibility as measured by the IRAP). Thus, because Symbol Search is a timed task, it may perhaps correlate well with relational ability only when it is highly fluent. In effect, a measurement of time taken by individual participants to produce symmetry may be an even more precise index of scores on the Symbol Search subtest. Interestingly, since the current doctoral work was completed, Williams et al. (2008) have re-emphasised that speed of processing is one of the major differentiating factors across individuals' behaviour. In effect, a measure of relational skills that also reflected a speed of task completion (i.e., a fluency score) may have been more useful for the research purposes. It should be noted however, that prior to the publication of the O'Toole et al. (in press) and Williams et al. (2008) research, steps had already been taken in the current research to introduce a temporal contingency on responding to produce speed, as well as accuracy of responding, in later experiments (Experiments 10a, 10b and 11).

Two of the strongest positive correlations found in Chapter 2, Experiment 1, were between performance IQ and the number of blocks required to produce transitive relations and between Block Design and the number of blocks required to produce transitive relations. Interestingly, the results of some recent work by O'Hora, Pelaez, Barnes-Holmes, Rae, Robinson, & Chaudhary (in press) have indicated that accuracy on a temporal relational task correlated with performance on the Block Design subtest on the WAIS. However, in the work of O'Hora et al. (in press), this correlation was unexpected insofar as the Block Design subtest is in the performance, rather than verbal domain in the WAIS-III (as well as in the WISC-III^UK^). Nevertheless, the results of the current research in Chapter 2, as well as the recent work by O'Hora et al. (in press), may indicate that performance on the Block Design subtest involves some type of relational skill that has not yet been elucidated theoretically by researchers.

For transitive relations, correlations with IQ and its subtests were often weak or in an unexpected direction. For instance, a strong and significant negative correlation was found between Symbol Search (a performance subtest of the WISC-III^UK^), and the number of blocks needed to produce transitive relations. Some factors which may have contributed to the emergence of unexpected correlations and large variations in the relation between IQ and transitive relation formation across participants was suggested in Chapter 2. Specifically, it was suggested that it is possible that SE tasks simply do not co-vary reliably with IQ tests and their subtest scores, as well as a more complex relational task or a measure of relational response fluency might. In addition, it may be that at a relatively early stage of intellectual development, the functional relationships between repertoires of relational skills and other intellectual skills have not yet been fully established. Thus, there may exist many differences in the skills of young children as they continue in their intellectual development. Intellectual development in adults may not present the same levels of variability. On balance, however, it should be pointed out that the strongest of the correlations observed in Chapter 2 were in the negative direction as predicted (i.e., symmetry testing). Therefore, while a range of mild and moderate unexpected positive and negative correlations were also observed, the expected and meaningful correlations (at least from a RFT point of view) tended to be the strongest recorded.

It was also noted in Chapter 2 that the number of blocks required to produce symmetry responding showed a much stronger correlation with IQ measures than transitivity. This was surprising because, at face value, symmetry appears to represent a more basic task than transitivity (see Hayes, 1991; see also O'Hora, Roche, Barnes-Holmes, & Smeets, 2002). Instead, the data showed that, typically, participants needed larger numbers of blocks to meet criterion on symmetry rather than on transitivity tests. It was speculated that this result pointed directly to the novelty of the symmetry testing context as the source of poor stimulus control over symmetrical responding (i.e., symmetry testing was the first form of testing to which participants were introduced). In addition, perhaps participants' performances improved as the context became more familiar and this set the occasion for improved abilities at derived relational responding in general, including transitivity skills. In other words, once a participant met criterion on a symmetry test, the participant may have been more likely to produce other familiar types of relational performances. As an analogy, in school level mathematics one established skill can lay the foundation for another. The way in which this may have occurred in the current research can only be speculated upon. However, it is reasonable to suggest that familiarity with the training and testing contingencies may have aided in the acquisition of the transitive relations, thereby normalising acquisition rates across participants. This observation should serve to alert future researchers to the need for highly developed baseline measures of relational ability, which control for such factors as task format novelty. The reader should note however, that partly for this very reason a more easily administered and less novel form of relational assessment was employed in subsequent experiments (i.e., the Relational Abilities Index).

The foregoing paradoxical outcome may have been unexpected given a strictly mathematical definition of equivalence as suggested by Sidman (1971, 1986). In Sidman's definition, it is suggested that the abilities to produce reflexivity, symmetry and transitivity occur together. However, Pilgrim and Galizio (1995, 2000), and Smeets, Y. Barnes-Holmes, Apkinar and D. Barnes-Holmes (2003), have reported on the independent emergence of symmetry and transitivity. The foregoing pieces of research (as described in Chapter 2) provided evidence that reflexivity, symmetry and transitivity may not be functionally dependent, and may not be best considered as occurring together as a single functional unit called stimulus equivalence. In effect, it may well be possible that there are lines of fracture between symmetry and transitivity that manifest themselves in the independent acquisition of each under specific conditions If this is so, it may not be so surprising that symmetry and transitivity performances did not correlate equally with IQ in the current research.

Experiment 2 went on to further examine the nature of the functional relationship between IQ and DRR. The findings from Experiment 2, Chapter 2, provided evidence that DRR is indeed an operant phenomenon and that modest rises in IQ scores were generated using the DRR intervention. Specifically, all four experimental participants reached the pre-determined fluency criterion given the DRR intervention and this skill was shown to generalise to novel stimuli. However, relatively weak differences across experimental groups were observed due to rapid acquisition of relational skills by all participants. Chapter 2 concluded with the researcher pointing out the necessity of circumventing the problem of rapid acquisition of relational skills for control participants by involving more complex relations. One suggestion proposed the inclusion of training and testing protocols for increasing relation skills involving the more complex relations (i.e., relations of Same, Opposite, More/Than, Less/Than, Before/After, et cetera.). Another approach to circumventing the problem of rapid acquisition of relational skills suggested included using a multi-component task, such as the protocols employed in O'Hora et al. (2005). A further suggestion included requiring an increased level of fluency by participants. Given the complex methodological and conceptual issues that arose in merely attempting to improve derived relational responding abilities, it was decided that subsequent experiments should focus on developing an effective procedure to this end before returning to the issue of raising IQ using relational MET interventions.

It was speculated that for the control participants, exposure to the training and testing protocol served as a form of practice that may have been somewhat effective in facilitating the acquisition of relational skills in the absence of a MET intervention. While not immediately obvious, this in itself may provide support for the idea that relational skills can be acquired through operant conditioning. More specifically, by the strict definition of MET, practice, in the absence of corrective feedback, does not qualify as a relational intervention. However, by definition, if exposure to training and testing with multiple stimulus sets leads (as it did) to an improvement in relational fluency, this demonstrates that relational skills must, in principle, be subject to practice effects (i.e., can be learned). While we have no precise idea of the learning contingencies in operation during practice, at least some reinforcement must have been occurring on a trial-by-trial or at least a test-by-test basis (i.e., operant conditioning). In the absence of any obvious source of reinforcement being delivered by the computer or experimenter, it is likely that some form of conditioned reinforcer for relational consistency must have been operating during the test tasks.

Interestingly, RFT, places a heavy emphasis on the role of conditioned reinforcers in precisely this situation. RFT would suggest that perhaps verbal/relational consistency (i.e., proving consistent responses that are self-discriminated as correct) could have acted here to establish relational skills for the control participants in the absence of explicit reinforcement. Indeed, in the context of Experiments 1 and 2 of the current research, it is entirely possible that consistency in responses across trials was actually reinforced in the laboratory context during training phases. In other words, a social history in which consistency was established as a reinforcer may not have been necessary. More specifically, the cessation of training and testing phases when participants produced consistent response patterns may have served as a reinforcer for appropriate levels of relational consistency across the numerous training and testing cycles. In effect, consistency was established as a conditioned reinforcer within the research programme itself. In addition, inconsistency in relational responding was punished by the delivery of repeated training/testing phases. Thus, it is speculated that in using the concept of consistency as a conditioned reinforcer, RFT can help to understand how exposure to the training and testing format alone may have some of the effects of explicit MET for the non-MET (i.e., control) group. (See Vitale, Y. Barnes-Holmes, D. Barnes-Holmes & Campbell, 2008 for further evidence; see also Baer, Deitrich, & Weninger, 1988; Catania, Shimoff, & Mathews, 1987; Lloyd, 1980; Paniagua, 1985; Ribero, 1989; Riegler & Baer, 1989 in Roche, Y. Barnes-Holmes, D. Barnes-Holmes, Stewart, & O'Hora, 2002, for evidence of verbal coherence as a reinforcer; see Guerin, 1994, for a discussion of consistency in early social environments; see Schauss, Chase, & Hawkins, 1997, for a discussion of the maintenance of verbal consistency).

It was noted in the Discussion of Experiment 2, Chapter 2, that there may have been a fatal design flaw in Experiment 2, insofar as groups were not only administered different training and testing protocols (i.e., presence or absence of feedback during testing), but groups also differed in terms of the overall number of blocks of training and testing delivered. This difference, rather than the MET intervention per se, may have at least partly determined the differences in relational skills at the follow-up test as well as the modest rises in IQ observed following the intervention. It is important to understand, however, that this criticism potentially challenges only the extent of the effect of relational MET in Experiment 2. It does not speak to the fact that relational skills are clearly operant in nature insofar as they did improve for all participants across time.

The foregoing methodological issues notwithstanding, perhaps the most exciting finding in Chapter 2 related to the modest IQ rises that were observed for most participants in the experimental group in Experiment 2. The rises in IQ appeared to suggest a clear pattern of influence by the DRR intervention. Regardless of the precise process involved in these rises (e.g., “practice” alone versus the effects of explicit MET), these recorded rises in what is otherwise considered a psychological invariant were noted as potentially of great interest to educational psychologists, psychometricians, and specialist educators working in a wide variety of domains.

The IQ rises observed in Experiment 2 exceeded the rises typically expected across time under normal conditions, even given repeated exposure to the IQ test. The reader is reminded that the stability of scores on the WISC-III^{UK} has been assessed in a study of 353 children who were tested twice (Wechsler, 1992). In the standardisation studies, the intervals between testings ranged from 12 to 63 days with a median re-test interval of 23 days. These data showed an increase of approximately 7-8 points in the FSIQ score over a short re-test period. Discrepancies in score, due to practice effects, are smaller for the verbal IQ (VIQ) than for the performance IQ (PIQ) score. However, practice effects on the WISC-R are smaller over longer test-re-test intervals (see WISC-III^{UK} manual; see also Juliano, Haddad, & Carroll, 1988, for further discussion). Thus, the rises observed in Experiment 2, Chapter 2 seemed to be larger than the standard rises associated with practice effects. In addition, it was noted that practice effects on IQ scores are typically short-lived when measured in test-retest studies (i.e., a median time of 23 days in the WISC-III^{UK} standardisation research). However, the length of time between baseline and follow-up administrations of the IQ test in Experiment 2 was between 90 and 120 days. In effect, the rises in IQ observed in Experiment 2 are unlikely to be accounted for by test-retest effects.

It was also exciting to note that one IQ rise from baseline to follow-up was so large that it could potentially move a child from one qualitative IQ range to another (e.g., from sub-normal to normal). At this point, it is worth re-considering some definitional issues related to the IQ metric. Specifically, in general terms, a clinically relevant rise in full scale IQ score is a rise that moves a participant out of his or her “qualitative range” (see Wechsler, 1991, 1992). Using Wechsler’s qualitative system of classification, IQ scores of 130 and above are classified in the exceptionally high range. IQ scores of 120-129 are classified in the high range. IQ scores in the 110-119 range are classified in the high average range. IQ scores in the 90-109 range are classified in the average range. IQ scores in the 80-89 range are classified in the low average range. IQ scores in the 70-79 range are classified in the low range. IQ scores of 69 and below are classified in the exceptionally low range. In the current doctoral research, P8 in the experimental condition in Experiment 2, presented with an eleven-point rise in full scale IQ score from baseline to follow-up. Although, this particular participant's scores at baseline and follow-up were both within normal range, the implications of such a rise with a learning disabled child are obvious. Such an IQ score increase could "move" a child from presenting within the Borderline Learning Disabled (or "low") range to presenting within normal range. Alternatively, such a rise could "move" a child from the Mild General Learning Disability range to the Borderline range. It has been pointed out that these types of classifications may seem superfluous to the casual eye, however, they would have great impact on the types of services available to individuals and on the way these individuals might be treated by those around them.

It was also argued that participants' IQ rises seen in Experiment 2, Chapter 2, might have been due to subject expectancies. However, the reader is reminded that MET group participants were given no more details on the purpose of the study than the non-MET group. No participant or parent/guardian was privy to information regarding research condition membership. In addition, participants had no access to intervention software outside of normal session hours. Thus, the participants were unable to practice the prescribed tasks. Parents and guardians were also naïve regarding the nature of the training and testing protocols. Therefore, there was no opportunity to assist participants' learning between sessions. Also, IQ was roughly normally distributed across groups (bearing in mind the small sample size) before the beginning of the intervention. That is, at the beginning of the research study, all except one participant (P2) presented in the average IQ range (90-109). The remaining participant (P2) presented in the high average range (110-119). Thus, Chapter 2, while preliminary in nature, succeeded in making some in-roads to the analysis of the functional relationship between relational skills and intelligence.

Chapter 3 set out to establish more reliable procedures for establishing increases in DRR skill by employing an improved experimental design, and by introducing more relationally complex tasks than were employed in previous experiments (i.e., Same and Opposite relations). The aim of Chapter 3 was to circumvent the problem of rapid acquisition of DRR, and the resulting ceiling effect in performance, as well as a floor effect (i.e., low numbers of training and testing blocks required) in the data which was encountered in Chapter 2. In other words, this more complex relational training and testing procedure was expected to produce a greater data range across participants, and thus, facilitate the differentiation of participants in experimental conditions to a greater degree. An alternative training procedure to MTS was also employed in an attempt to allow a very large amount of relational training to take place in as short a period of time as possible**.** In that manner, a new procedure, which can be described as a combination of the Relational Evaluation Procedure (see Cullinan, Barnes-Holmes, & Smeets, 2001), and the Yes-No (see Fields, Adams, Verhave, & Newman, 1990) procedure was used. The new procedure employed seemed to have been an effective training strategy in that no participant in Chapter 3 failed to meet criterion for training. However, without making systematic experimental comparisons to other procedures, it was impossible to demonstrate that the new procedure was more effective than either REP or Yes-No methodologies or even other more traditional training techniques based on conditional discrimination.

It was found in Experiment 3 of Chapter 3 that relational responding in accordance with the frame of Opposite required a larger number of novel stimulus sets across training and testing than was required for participants during relational responding in accordance with the frame of Same. Several conclusions were drawn from the findings of Chapter 3. The first and most obvious conclusion was that, regardless of age or exposure to DRR in either a natural or laboratory setting, all participants in Experiment 3 demonstrated DRR in accordance with Opposition across higher numbers of blocks of training, testing, re-training and MET testing than were required in order to demonstrate DRR in accordance with Same relations. Thus, it was argued that DRR in accordance with Opposite relations is a more complex relational skill than is DRR in accordance with Same relations. Consequently, it was suggested that it may be beneficial if DRR in accordance with Same is trained before DRR in accordance with Opposite in any intervention designed to establish a comprehensive repertoire of derived relational responding. In such an intervention, each form of relational responding should ideally act as a building block for the next. In addition, exposure to interventions presented in a similar format would likely increase the generic skills (e.g., attentional skills) involved in learning relational tasks more generally. Participants would also become increasingly familiar with the stimulus types and task format before being exposed to more challenging relational tasks (i.e., Opposite). In effect, their behaviour would already have come under the control of the conditioned reinforcement involved in feedback, and the social reinforcement provided by the experimenter for participation. It was pointed out in the discussion to Chapter 3 that these factors should, at least partially, improve the effectiveness on the MET intervention for Opposite relational responding.

Despite the seemingly promising results of Experiment 3, it was unclear precisely which feature of the MET training was efficacious in establishing DRR in accordance with Same and Opposite. That is, it was not yet known if it is necessary to train and test participants to criterion on novel sets or whether mere exposure to novel stimulus sets is the crucial factor in enhancing relational skills. This issue was also raised in Chapter 2 when it was discovered that "practice" alone in the absence of corrective feedback (i.e., exposure to training and testing without feedback) on novel sets appeared to enhance relational skills. In addition, some potential flaws in the experimental design were highlighted in Chapter 3. Chapter 4 addressed these potential flaws, while also continuing to pursue the questions outlined at the beginning of Chapter 3.

Chapter 4 set out to examine a new MET procedure with various groups of adults and children, some of whom had prior exposure to multiple exemplar training and some of whom did not. Including groups without previous exposure to MET allowed the experimenter to examine the issue of a possibly optimal sequence of relational training for a preliminary applied intervention. The new procedure incorporated the availability of a remedial level which participants could have access to if they were unable to meet criterion on a probe test after a designated amount of exposures to cycles of relational training, relational testing, further relational training and multiple exemplar testing. The new procedure also included extra stimulus control tasks, and it did not include the former practice of testing to criterion on multiple exemplar tests (i.e., the new procedure involved one exposure only per test).

The results of Experiment 4 suggested that the multiple exemplar intervention employed was successful in establishing a repertoire of derived relational responding that was previously absent or weak for all of the adult participants. The introduction of the extra control tasks to preclude extraneous sources of control ensured that appropriate contextual control was established during that experiment. The addition of the remedial level possibly assisted one participant in producing DRR for Opposite relations. However, it was also noted that the adult participants in Experiment 4, on the whole, did not require a great deal of intervention in order to establish DRR with derived relations, including the relation of Opposite.

Experiment 5 employed identical procedures with ten normally functioning children. The results indicated that child participants in Experiment 5 required exposure to training and testing across a larger number of novel stimulus sets than did the adult participants in Experiment 4, for both Same and Opposite relational responding. However, the differences between Same and Opposite DRR acquisition were not quite as noticeable for adults in Experiment 4 as they were for child participants in Experiment 5. More specifically, while Same relational responding emerged rapidly for both adults and children across Experiments 4 and 5, adults required little, if any, intervention to demonstrate DRR in accordance with Same relational responding. In addition, all members of the non-MET group produced DRR in accordance with Same relations in fewer relational training and testing cycles than were required for the relation of Opposite. Two of the participants did not meet criterion for Opposite relational responding at all. However, there were no noticeable differences in the acquisition rates of Same DRR for MET and non-MET groups noted. Thus, it appeared that the multiple exemplar intervention employed was indeed more successful in establishing a repertoire of derived relational responding for Opposite relations than was training and unreinforced testing alone. For these reasons, it was suggested that exposure to multiple exemplar testing and the availability of remedial levels of intervention helped to facilitate the acquisition of DRR in accordance with Opposite relational responding within the designated number of novel stimulus set blocks (i.e., 23).

Experiment 6 attempted to establish a possibly optimal sequence of relational training that might produce a rapid emergence of a battery of relational repertoires. Specifically, this experiment examined the role of a history of DRR with one relation or set of relations (i.e., Same and Opposite) in acquiring DRR with other relations (e.g., More-Than/Less-Than). Acquisition rate differences across child and adult participants were also examined. The final issue examined was whether or not exposure to training and unreinforced testing for More-Than/Less-Than relations alone could produce similar rates of acquisition for both child and adult participants as could be produced with exposure to a multiple exemplar intervention.

The results of Experiment 6 showed that adult participants who had no prior history of DRR in accordance with Same and Opposite relational responding and who were not exposed to the multiple exemplar intervention demonstrated slower rates of acquisition than the adult participants who had previous exposure to interventions for other forms of DRR. Adults with no prior history of DRR in accordance with Same and Opposite relations who were not exposed to the multiple exemplar intervention also demonstrated slower rates of acquisition than the adult participants who had no previous exposure to DRR, but who were exposed to the multiple exemplar intervention.

Therefore, it was suggested that having exposure to a history of Same and Opposite relational responding may not determine acquisition rates of More-Than/Less-Than relational responding for adult participants.

Experiment 6 also showed that child participants with a previous history of DRR with Same and Opposite relations required more cycles of training (including re-training), relational testing and multiple exemplar testing with More-Than/Less-Than relations than did adults with the same history. Child participants with no previous history of DRR in accordance with Same and Opposite relational responding attained similar rates of acquisition of More-Than/Less-Than relational responding to their counterparts who had a laboratory history of DRR in accordance with Same and Opposite relations. Thus, it was once again suggested that having a laboratory history of DRR for Same and Opposite relational responding may not be necessary for the establishment of More-Than/Less-Than relational responding. Of course, it might also be the case that having a minimal and low-rate level of Same and Opposite relational skills below that of the current criteria levels, may play some role in the acquisition of all other relations. It was not feasible for the current research to gain access to children with no Same and Opposite relational skills whatsoever (i.e., pre-language) in order to test this idea. Nevertheless, it is likely that the pre-existing sub-criteria level Same and Opposite relational skills of these participants did in some way facilitate the emergence of More-Than/Less-Than relational skills. Indeed, child participants who had no prior laboratory history of DRR in accordance with Same and Opposite relational responding, and who were not exposed to training and unreinforced testing for More-Than/Less-Than relations, demonstrated slower rates of acquisition for More-Than/Less-Than relational responding than all adult and child participants. Nevertheless, the differences between the child participant groups were small. Thus, it appears that once again, having access to a multiple exemplar training intervention was helpful in generating DRR for More-Than/Less-Than relations. However, even in the absence of any MET intervention, it appeared that DRR for More-Than/Less-Than relations can emerge, albeit at a slower rate.

Given the foregoing conclusions, the experimenter asked whether or not the relations of Same and Opposite needed to necessarily be trained first in an intervention aimed at increasing relational skills. It had been previously suggested that training in Same relational responding should come first in an intervention with an educationally challenged population. However, it was now apparent that More-Than/Less-Than relations could emerge without intervention for adults, and even for children, in the absence of a history of multiple exemplar training with other relations in the laboratory setting. Thus, More-Than/Less-Than relations now appear to be relatively fundamental relational skills that are at a high level of proficiency for most adults and children entering the laboratory. Such relations require little intervention to raise the fluency of those skills to criterion levels. In contrast, in Experiment 5, two participants failed to demonstrated DRR in accordance with Opposite. This was a surprising outcome because Same and Opposite have been generally considered to be established earlier in life than More-Than/Less-Than relations are established. These data suggested the reverse. Thus, it was suggested that an effective intervention might train participants to first acquire DRR in accordance with Same relations, then to acquire DRR in accordance with More-Than/Less-Than relations and then finally to acquire DRR in accordance with Opposite.

The naïve child participants from Experiment 6 who were only exposed to More-Than/Less-Than relational responding were re-recruited to form the subject pool of Experiment 7. Results of Experiment 7 indicated that all participants required larger amounts of training across more novel stimulus sets to meet criterion for Opposite than those participants required to meet criterion in accordance with More-Than/Less-Than relations in Experiment 5. Three of the non-MET participants did not meet criterion for Opposite relational responding at all. However, all MET participants demonstrated DRR in accordance with Opposite within less than the maximum number of exposures to novel stimulus sets (i.e., 23).

Despite the relatively slow acquisition rates for Opposite relational responding in Experiment 7, the multiple exemplar intervention seemed to have been of benefit to the acquisition of DRR in accordance with Opposite. This is evidenced by the lower number of blocks of relational testing required to meet criterion for participants in the MET group, and by the fact that three of the participants in the control (non-MET) group failed to demonstrate the generalisation of DRR in accordance with the relation of Opposite within the pre-determined maximum number of exposures to novel stimulus sets (i.e., 23).

Experiment 7 failed to provide empirical support for the earlier suggestion that More-Than/Less-Than relations must be trained before Opposite relations in a DRR intervention. More specifically, no functional dependence was shown between these two relational responding repertoires. Indeed, Experiment 7 suggested that in terms of functional relationships, it may not matter which of these relations is trained first in a derived relational responding intervention. In effect, Experiments 6 and 7, taken together, have demonstrated that relations of More-Than/Less-Than and Opposition may be relatively functionally independent. Alternatively, they may both be functionally related to a common, less complex or more complex relational responding repertoire, not yet identified. However, across the experiments in the Chapters 3 and 4, Opposite relational responding was consistently relatively difficult to establish. This supports the emerging view that the relation of Opposite is more complex than other relations examined in this research.

The findings from Experiment 7 did not yield conclusive answers to the question of whether or not relations of Opposition functionally depend on the other trained relations which have been shown to emerge more readily. In fact, the data set from Experiment 7 suggested that the relations of More-Than/Less-Than and Opposition may be functionally independent. However, this very complex empirical question likely warrants a thorough investigation in its own right. As the current research had an applied goal as an end point, it was necessary to at least test the current intervention protocol first, and allow the results of an intervention to inform further research questions. In any case, it is important to remember that the relationships between the various relations are not fixed in stone. RFT does not see intellectual development as unfolding in a pre-determined direction for all humans, as in the Piagetian scheme. Rather RFT proposes that development acquires a unique trajectory for each human, depending on the entire gamut of biological and personal historical variables, including the effect of current context. Indeed, from an RFT perspective, finding the “correct” sequence is not important. Rather, what is important is finding a relational training sequence that is workable and maximally effective for as many individuals as possible. This training sequence may vary across situations, and participant groups, depending on the particular deficits and relevant biological variables.

As an illustration of the foregoing, it was outlined in Chapter 4 that Opposite relations might be viewed as a subset of difference relations (i.e., all opposites are different to any given samples, but not all stimuli different to a sample are opposite to each other). Roche and Barnes (1996a) noted that while difference and opposite relations might be viewed as forms of non-equivalence responding, this is not the case. Indeed, Roche and Barnes (1996a) demonstrated empirically that difference relations often require more complex forms of relational contextual control than non-equivalence (S-). As argued in Roche and Barnes (1996a) and Barnes and Roche (1996), opposite relational responding in turn requires an even finer form of contextual control than difference relational responding, and thus, likely follows the emergence of difference relational responding in young children. What was not known at that time was precisely how much more training and testing might be required to establish opposite relational responding, as opposed to difference responding. The current data suggest anecdotally that while many children might display fluent difference relational repertoires, it may take considerably more multiple exemplar training in the natural environment for opposite relational responding to emerge.

One final issue which was considered in Chapter 4 related to the seemingly paradoxical militating effect of histories of More-Than/Less-Than responding on the acquisition of Opposite relational repertoires. It appeared that, in many cases, histories of More-Than/Less-Than responding were associated with slower rather than faster rates of acquisition of Opposite relational responding. One possible explanation for this outcome pointed to the increasing complexity of the contextual control required to parse the various forms of relational responding into tight functional classes. Put simply, as experimenters attempt to establish additional forms of contextual control with a given participant pool, the opportunity for confounded forms of stimulus control increases. For example, it is likely that when first presented with relational pre-training and training for Opposite relations, many participants with a laboratory history of More-Than/Less-Than responding may have responded in early trials in terms of More-Than/Less-Than relations, rather than demonstrating control by the Opposite contextual cue. It was suggested that over several trials, the new form of control would be established. Nevertheless, an important part of the process of establishing conditionally discriminative control over responding is to “wash out” other forms of control across several trials (Sidman, 1994). This process takes time and becomes more challenging with each form of contextual control being established. This is because more and more established forms of control must be “washed out” upon each attempt to establish an additional form of contextual control. From this perspective, it may not be so surprising that prior histories of relational training sometimes appeared to militate against the acquisition of contextual controls over newer forms of DRR. Of course, this issue requires further examination in future research.

Before the emerging intervention could be delivered in a real educational setting, one final issue warranted examination. Specifically, it was not yet known whether or not a relational training intervention, as it had now been devised, would have any benefit to either a normally functioning or to an educationally challenged population. Thus, it was determined that it was important to test the general effectiveness of the intervention using participants selected from the general population. To achieve this, in Chapter 5, participants from Chapter 2 who had already taken part in training and testing for DRR in accordance with symmetry and transitivity were re-recruited. These participants also had received access to a Same and Opposite relational responding intervention identical to the one presented in Chapter 3. The aim of the experiment was to establish whether or not the extensive history of multiple exemplar training provided to these participants thus far would lead to a measurable increase in IQ scores. First however, it was necessary to provide these participants with the More-Than/Less-Than MET intervention in order that they would have been exposed to a full battery of Same, Opposite, More-Than/Less-Than relations, albeit with less effective training and testing procedures (i.e., based on conditional discriminations) than had been selected for use in the final MET battery for use in an educational setting (i.e., based on the Yes-No and Relational Evaluation procedures). Thus, Chapter 5 aimed to provide a group of experienced participants with a More-Than/Less-Than multiple exemplar relational intervention. Following the delivery of the More-Than/Less-Than relational intervention, this group would have had access to a full battery of relational interventions (symmetry, transitivity, Same and Opposite). The effect of this battery on full scale IQ could then be determined. The findings from Experiment 8 showed that the procedure for More-Than/Less-Than relational responding was successful in improving a repertoire of More-Than/Less-Than relations for all child participants in that group. At that point in the research, the novel More-Than/Less-Than training procedure had been successfully employed with both adults (Experiment 6) and children (Experiment 6) with varying levels of exposure to DRR and varying histories of exposure to relational interventions (Chapters 4 and 5).

In Experiment 9, the participants from Experiment 8 were re-recruited. These experimental participants now had exposure to a relational intervention which consisted of an entire battery of relations (i.e., symmetry, transitivity, Same, Opposite and More-Than/Less-Than). In Experiment 9, the control group from Experiment 2, Chapter 2 were also re-recruited to serve as longitudinal comparisons to the intervention group. The administration of the WISC-III^{UK} to both participant groups revealed that participants in the experimental group showed large FSIQ rises from Time 1 (Experiment 2, Chapter 2) to Time 3 (Experiment 9, Chapter 5), and that similar rises were not seen for the control group from Time 1 (Experiment 2, Chapter 2) to Time 3 (Experiment 9, Chapter 5). Thus, it appeared that the battery of MET relational training employed in Experiments 2, 3 and 12 has the potential to make real and measurable differences to the full scale IQ of normally developing children. The IQ rises observed for experimental participants were larger than those observed in Experiment 2, Chapter 2. While clinically relevant rises in IQ were neither expected nor required in this investigative research, they were nevertheless observed for experimental participants in many cases. That is, small rises were witnessed for all members of the experimental group from Time 1 (Experiment 1, Chapter 2) to Time 2 (in Experiment 2, Chapter 2) and only for one member of the control group (Experiment 2, Chapter 2). In Experiment 9, large rises in full scale IQ were observed for all members of the experimental group, but not for any member of the control group. These rises included rises of 18 full scale IQ points (P2), 26 full scale IQ points (P3), 32 full scale IQ points (P8) and 33 full scale IQ points (P1). Rises in scores on individual subtests which may have contributed to the overall rise in IQ were as follows; the Picture Arrangement subtest, the Comprehension subtest and the Information subtest. Interestingly, it was noted that correlations between precisely these subtest scores and a combined measure of stimulus equivalence ability was observed in Experiment 1, Chapter 2.

It was noted in the discussion of Chapter 5 that according to Wechsler (1992, 1991, p. 33), about 95% of all children obtain scores within the 70-130 range. This means that only 5% of children can be said to score in the exceptionally high or the exceptionally low qualitative range. Of the remaining 95% of children, scores range within 10 points of the mean of 100. Of course, it is recognised that children can score anywhere within each range in order to be classified within a qualitative diagnostic range. So a child who attained a full scale IQ score of 108 on a first IQ test and a 110 on a second IQ test would not be noteworthy, even though this child would have technically "moved" from the average to the high average range. In contrast, a rise of 10 points (e.g., from 108 to 118) would be noteworthy, as this is a larger rise than would typically be seen from test to re-test administrations (see Wechsler, 1991, 1992, p. 63; see also Juliano, et al., 1988 for discussions on practice effects).

The IQ score shifts observed in Experiment 9 are all the more impressive when one considers the variations typically seen within and across participants. More specifically, two of the experimental participants moved from the qualitative average range to the qualitative exceptionally high range. One participant moved from the qualitative high average range to the qualitative exceptionally high range. Finally, one participant moved from the qualitative average range to the qualitative high range. Importantly, the three control group participants stayed within the average range from first test administration to third test administration. The remaining control group participant moved from the high average range to the average range from first test administration to third test administration. (This change was comprised of just three full-scale IQ points).

The relational intervention employed in Experiment 9 had a larger impact on performance IQ than on verbal IQ. This is in line with the stability data for the WISC-III^{UK} (Wechsler, 1992, 1991), which shows that discrepancies in scores due to practice effects are higher for performance IQ scores than for verbal IQ scores. Thus, it was speculated that perhaps part of the explanation for the observed rises in performance IQ lies in practice with the training and testing format and context to which participants were repeatedly exposed. Put simply, this procedure may have enhanced attentional skills as well as bringing behaviour under increasing control by the social and conditioned reinforcers delivered within and across training and testing trials. It is also entirely possible that the social engagement produced by both the experimenter, and the experimenter's voice on the computer, may have made engaging more reinforcing than would typically be expected for these types of experiments. This speculated outcome may have been related to the fact that the social engagement was socially realistic because it was randomised. The effects of social reinforcement during experimentation, however, likely warrant further examination in isolation before conclusions about their efficacy can be made with any confidence.

In addition, because the subtests which contribute to the performance IQ are all timed and are thought to measure an individual's "perceptual reasoning and organisation" and, if indeed repeated relational interventions have the effect of increasing attentional skills, we might expect to observe rises in overall performance IQ as a result of such interventions (see also Williams et al., 2008, for a recent call to behaviour analysts to pay more attention to speed of processing as a behavioural measure).

Verbal IQ scores also rose across all four experimental participants from Time 1 to Time 3. The observation of this mean rise of 27.25 VIQ points supports the RFT position that relational skills and verbal skills are functionally related and possibly even part of the same repertoire.

It was also noteworthy that both the performance and verbal domains were impacted to a large degree for all experimental participants. At the inter-subtest level, large positive changes were seen for ten of the twelve subtests administered. All four experimental participants presented with large positive change in the Comprehension subtest (verbal domain), and the Picture Arrangement subtest (performance domain). Three out of four experimental participants presented with positive change in the Information subtest (verbal domain). Thus, it would appear that in the context of this more thorough intervention spanning several relational responding repertoires, and across a large set of stimuli, most IQ subtest scores can be shifted in a positive direction.

The foregoing theoretical functional analysis of the relationship between the intervention employed across Chapters 3, 4 and 5 and IQ may point us in the direction of important relationships that can be examined more carefully in experimentation. However, the main focus of the current research remained to test such an intervention at a macro level with a population in need, rather than to tease these relationships apart purely for conceptual reasons. Thus, while many important questions remained, it was noted that they would have to be pursued in future research endeavours. At the close of Chapter 5 in the current body of research, it was clear to see that the relational intervention developed across Chapters 3 and 4 had the potential to impact on full scale IQ (in Chapter 5), as well as the many subtests which contribute to a full scale IQ score.

Chapter 6 started out by devising and pilot testing the relational abilities index (RAI) to measure children's relational abilities across four relational repertoires. These relations were Same, More-Than, Less-Than and Opposite. The RAI was first piloted with three children in Experiment 10a. In that experiment, it was shown that the RAI was, in principle, an effective measure of a child's relational abilities insofar as it produced non-maximum data and appeared to differentiate children in terms of relational abilities. In Chapter 10b, a group of children presenting with learning difficulties were recruited and administered the WISC-IV^{UK} and four individual RAI tests for relations of Same, More-Than, Less-Than and Opposite. Correlational analyses revealed modest correlations between many of the subtests on the WISC-IV^{UK} and the mean RAI scores for each relation. However, the strongest correlation (and the only significant one) was seen between the mean RAI score for Opposite relational responding and the Comprehension subtest on the WISC-IV^{UK}. This was an interesting finding because of all the four relational repertoires, generating DRR in accordance with Opposite relations has generally required the greatest amount of training and testing. In other words, Opposite relational responding ability may best discriminate participants in terms of their comprehension as measured by the WISC-IV^{UK}.

The next strongest correlation, which approached significance, was seen between the mean RAI score for Opposite relational responding and the Verbal Comprehension Index on the WISC-IV^{UK}. While not a significant correlation ($r = .614$), the observation of this relationship between Opposite relational responding ability and verbal ability, as measured by the WISC-IV^{UK}, is wholly unsurprising from an RFT perspective. Many other correlations were observed, in both a positive (and expected) direction, as well in a negative (and unexpected) direction. Once again the lack of clear relationships between DRR, in this case as measured by the RAI, and IQ measures is apparent.

It should be remembered that it is unlikely that researchers will find significant correlations with such a small participant pool. Indeed, it is speculated that more statistically significant and perhaps even stronger correlations would be found if a larger participant pool were employed (e.g., O'Hora et al., 2005). A large sample size may serve to normalise data and stabilise any emerging correlations. Of course, it should be considered that some IQ subtests may simply not correlate well with DRR ability, even if there is a broad functional overlap between the DRR and intellectual skills repertoires. Furthermore, it is important to understand that correlations between various IQ subtests and DRR ability do not need to be found in order for a *functional relationship* between DRR and IQ to exist. In other words, statistical correlation and functional dependence are not synonymous. For instance, IQ and DRR may be poorly correlated at baseline, but shifts in one repertoire may nevertheless lead directly to shifts in the other.

In Experiment 11, the participant pool from Experiment 10b was re-employed. These participants were administered the intervention which had been developed across Chapters 3 and 4. However, the intervention was divided into three parts. The first part targeted Same relational responding. The second portion targeted More-Than/Less-Than relational responding, and the third part targeted Opposite relational responding. The results showed that all participants' mean RAI scores improved from Time 1 to Time 2.

Experiment 11 also showed that all participants showed an increase in full scale IQ from Time 1 to Time 2.

Several potential reasons for the rises seen in RAI scores and in IQ scores can be offered. One reason may be that the current MET format employing a REP and Yes/No hybrid, with a large number of stimulus sets, is in fact an efficacious procedure for generalising DRR. However, the sequence in which the relations were trained during intervention may also have played some role (i.e., it was optimal or near-optimal). The use of remedial levels of intervention relying on non-arbitrary, rather than arbitrary forms of relational responding may also have assisted in the improvements in relational abilities crucial to IQ test performance. It was beyond the scope of this thesis to ascertain which specific change in the intervention format employed is most responsible for the resulting changes in RAI scores and IQ scores. However, it is likely that the various modifications together constitute a somewhat effective intervention for relational ability and IQ.

It is not possible to rule out the effects of practice on rises in RAI scores. Specifically, repeated exposure to training and testing formats across multiple relational responding repertoires, rather than effects of MET, may partly explain improvements in RAI. This can not be easily ascertained at this point because, for ethical reasons, no control group was employed. More specifically, while participants' guardians were assured that the current research was experimental and not therapeutic in nature, the experimenter expected rises in IQ to be observed for the current expanded intervention under controlled experimental conditions. Thus, it was deemed unethical to deprive any group with educational deficit of the full intervention battery in order to assess the effects of practice alone on IQ. While practice surely had some effect on RAI scores, such large and dramatic rises in full scale IQ observed for participants are difficult to explain without reference to the MET intervention itself. In addition, it has already been clarified in Experiment 9 that IQ rises have been observed only where the MET component has been included in relational training and testing interventions.

At that point in the foregoing doctoral research programme, many questions remained regarding the relational processes at work in the current intervention. For instance, given that scores on the Comprehension subtest were correlated with mean RAI scores for the relation of Opposite, it would be reasonable to assume that a relational intervention which included the relation of Opposite would increase a participant's score on the Comprehension subtest. While rises in Comprehension subtest scores certainly were seen, rises were also seen for all other subtests. Not surprisingly, the IQ subtests are functionally related, and IQ tests have been designed with this in mind. Interestingly, however, it would appear that the relatively small MET battery delivered here was sufficient to shift the entire range of IQ subtest performances in a positive direction. This is suggestive of the current battery being relatively comprehensive in targeting the full range of skills required to perform well on IQ tests. Of course, in order to discover more about the relative functional relationships between the various relations established, the relational interventions for Same, More-Than/Less-Than and Opposite would have to be run separately in dedicated basic research studies.

Interestingly, the range of educational needs of the participants recruited for Experiments 10b and 11 was large. Some participants presented with diagnosed general learning disabilities (i.e., measured IQ lower than 80). Still others presented with early literacy difficulties, which would not necessarily have any impact on measured IQ.

Others presented with difficulties with their school's mathematics curriculum, while other participants presented with attention difficulties (some formally diagnosed and some merely suspected). Finally, some participants presented with reported difficulties in motivation. These individual differences may partly explain the large variations in the relationships between IQ and DRR abilities. The important point, however, is that there was a reliable increase in IQ across participants, suggesting a strong common and non-specific functional relationship between DRR ability levels and IQ.

General Methodological Issues

While several methodological issues have been raised throughout the thesis, it is important to reiterate the most significant of these here. As a general note, however, it is worth bearing in mind that in order to ultimately administer the developed intervention to an educationally challenged sample, it was sometimes more efficient to move past interesting and important basic research questions and to be driven by practical considerations. This undoubtedly renders the current research less integrated methodologically than a similar piece of entirely laboratory-based research.

One recurring criticism related to the necessarily small size of the participant pool. There were three major reasons why the participant sample was necessarily small. The first and most important reason was that the time required to conduct many of the experiments in the current thesis was protracted. Some participants (e.g., the longitudinal groups from Chapters 2 and 5) took part in this research over the course of three years. This extended involvement required a huge time commitment on the part of the participants, their families and the researcher. Some participants in Chapters 3 and 4 took part across several weeks. Participants who took part in Experiments 10b and 11 of

Chapter 6 were employed for an entire academic year. These participants were seen approximately twice weekly. In addition, as a service to the children and their families, a comprehensive psycho-educational assessment, adhering to all standard protocols of the National Educational Psychological Services, was also provided for each child. These additional services included hours of consultation with parents and teachers, at the beginning and end of Experiments 10b and 11, as well as time spent administering additional assessment tools required by the terms of the agreement. The psycho-educational assessment involved the administration of various attainment tests (literacy, numeracy, written and oral language) if required by each individual child's set of circumstances. Based on the results of those standard assessments, reports were written for the school and various referrals were made to other appropriate professionals (i.e., speech and language therapy, occupational therapy, special educational needs organisers, resource and learning support teachers, et cetera). Thus, participation in the research by the children with educational needs required extensive involvement and time commitment by teachers, parents, other appropriate professionals, as well as by the researcher.

The second and related reason for necessarily utilising small participant groups was that many of the experiments involved several sessions of off-campus participation. Thus, the work was not only labour intensive in terms of the amount of time required for intervention groups to get from the start to finish of an intervention, but other groups (both MET and non-MET) across experiments frequently needed many sessions to complete an individual experiment. As many participants, particularly in Chapters 2 and 3, were seen in their own homes, travelling to and from private homes over the course of several sessions for each individual participant, required a large time commitment by the researcher. In addition, taking part across multiple sessions required a great time commitment by the individual participants.

The age distributions across participants and across experiments were not perfectly controlled in the current research. At the beginning of the research it was decided to employ child participants who ranged from the age of eight up to the age of 12. This chosen age range was not arbitrary. More specifically, eight was chosen as the lower age limit as children commence their education in Ireland at approximately five years of age. For the first two years of school, children are learning letters, the sounds these letters make and how to blend these letters/sounds to make words. As such, at the ages of approximately five and six years of age, Irish children are being taught what might be considered "pre-literacy" tasks. Thus, Irish children are only formally "reading", as it is commonly thought of, at the age of approximately seven years. As the researcher needed participants to be able to read the nonsense syllables which were used as stimuli throughout the thesis, it was decided to employ a group of students who had, at least in principle, mastered the basics of reading. (Children in all normally functioning groups were also designated by teachers as not presenting with any known or suspected learning difficulties). It was also speculated that by the age of eight, a child participant would be more familiar with school based learning requirements such as attending to tasks, following directions, staying in one's seat, et cetera. The upper age range of 12 was chosen as most Irish school children are enrolled in primary education up until that age. Thus, for pragmatic reasons, it was more time efficient to approach only Primary schools and to not attempt to engage with Secondary schools because this might have served to further increase travel times and involvement by additional teachers, as well as increasing other practical considerations such as classroom availability, movement of equipment from one school to another, and so on. It was also speculated that an age range encompassing any more than four years might yield too many developmental differences to make any comparisons among participants meaningful. For these reasons, the age range of eight to 12 years was utilised in early experiments involving child participants. However, by the end of Chapter 3, it became clear that, at least anecdotally, some of the younger child participants were less likely to stay on task than their older counterparts. Thus, it was decided to engage children at the older end of the chosen age range for the remaining experiments wherein child participants were employed. Thus, from Chapter 4 onwards, child participants were selected only from 6th class at a local primary school. These child participants were between the ages of 11 and 12 years old. However, the comparisons across experiments, particularly the earlier experiments, may sometimes have involved comparisons across children of slightly different ages.

The experimental work reported in Experiment 10b and 11 of Chapter 6 might also be subject to the criticism that IQ scores at baseline differed slightly across participants. It is important to remember, however, that these participants were not recruited precisely because they had low IQ scores. Rather they were recruited on the grounds that they were experiencing educational difficulties at school. In a population of children who are considered "sub-normal", there may be various different reasons why these children are not succeeding educationally, sometimes including low IQ. Some of these children may have been formally diagnosed with a general learning disability (i.e., having a diagnosis of a full scale IQ below the low average range/80 or below). In some children considered sub-normal, the difficulty presenting may not be related to presenting with lower than average IQ's at all. For instance, many children in mainstream classrooms may present with normal range IQ scores, but may present with various other diagnosed or undiagnosed behavioural, communication or even co-ordination difficulties (e.g., ADHD, Asperger's Syndrome, Autistic Spectrum Disorder, Oppositional Defiant Disorder, Dyspraxia, Specific Speech and Language Disorder, et cetera) which may also impact on their level of academic ability. Another possibility might be that participants present with a normal range IQ, but present with a specific difficulty or disability in one or more attainment areas (e.g., Dyslexia, Dyscalcula, Disorder of Oral or Written Language, et cetera). In addition, there are various other emotional and environmental factors which may impact on a child's ability to perform (e.g., family trauma, emotional or physical abuse, alcohol or drug dependence within the family setting, other mental illnesses such as anxiety or depression, low socio-economic status, et cetera.). Thus, in any such sample of children, IQ scores will vary. For ethical reasons, the sample suggested that participants who were both willing and most in need of intervention in the current research were accepted as participants. Thus, control of IQ ranges was not possible in Experiment 11. Nevertheless, it is recognised that any further research in this field would benefit from more closely matched participant samples at the outset of any intervention or experimental work.

It could also be suggested that more space could have been devoted here to disentangling the possible functional relationships between various forms of DRR and various aspects of IQ testing. However, it is important for the reader to appreciate that, as noted in Chapter 2, for behaviour analysts, IQ tests are not treated as anything more than statistically produced batteries of test items, responses to which correlate well with general educational ability. Thus, they have no status in behaviour analysis as units of behaviour, or units of measurement, and are employed here purely for the purposes of inductive research. In other words, it does not behove the behavioural researcher to account for each and every aspect of a test whose status in the experimental analysis of behaviour is questionable to say the least (see Williams et al., 2008, for criticisms of behaviour analysis for not addressing issues related to intelligence). In fact, it is research of the current kind that will help to elucidate the functional properties of IQ measures by correlating them with known behavioural processes. This, in turn, may ultimately result in these subtests being translated into behavioural terms and being subjected to rigorous analyses to determine their use in the prediction and control of behaviour.

Why an RFT Approach?

RFT provides a technical analysis of language and cognitive abilities and, more importantly, suggests procedures for making conceptual and empirical inroads to the problem of understanding the origins of, and the functional nature of intelligence. In particular, RFT claims that the foundational skill for most intellectual abilities is *Derived Relational Responding*. Behavioural research into derived relational responding performances has been greatly enhanced by the seminal work of Murray Sidman into the phenomenon now known as stimulus equivalence (see Sidman, 1994, for a review). However, RFT researchers have now provided abundant evidence that the mathematical relationships described by Sidman in his early research extend beyond mere equivalence (Y. Barnes-Holmes, D. Barnes-Holmes, Smeets, Strand, & Friman, 2004; Dymond & Barnes, 1995, 1996, 1997; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2008; Healy, Barnes-Holmes, & Smeets, 2000; O'Hora, Barnes-Holmes, Roche, & Smeets, 2004; O'Hora**,** Roche, Barnes-Holmes, & Smeets, 2002; Roche & Barnes, 1996b, 1997; Roche & Dymond, 2008; Whelan, Barnes-Holmes, & Dymond**,** 2006; Vitale, et al., 2008). In a recent paper, Sidman (2008) acknowledged this by suggesting that his original view of equivalence "does not cover other kinds of relations than equivalence, as for example, relational frame theorists attempt to do" (Sidman, 2008, p. 10).

In moving beyond mere equivalence as a metric for complex relational patterns of responding, RFT adopts a somewhat different nomenclature than Sidman. The RFT nomenclature has allowed for the analysis of a far broader range of relational skills than ever envisaged by Sidman, and those who resisted the evidence that forms of DRR other than equivalence were possible (e.g., Saunders, 1996 ; see also Galizio, 2003; Malott, 2003; McIlvane, 2003 for critiques and commentaries).

A third suggested contribution of RFT over traditional behaviour analysis is that RFT posits that the ability to derive relations is itself a learned operant skill (D. Barnes-Holmes & Y. Barnes-Holmes, 2000; Healy, Barnes-Holmes, & Smeets 2000). This sets the theory apart from Sidman's view of derived relational responding as a basic stimulus function (e.g., Sidman, 2000; see Hayes & Barnes, 1997). In simple terms, RFT suggests that the ability to derive relations is itself established by caregivers at an early stage, across multiple exemplars, often without the caregiver even being aware. RFT studies, such as these, point again to the immediate possibility of intervention when a functional relationship between reinforcement contingencies and increases in DRR skill is posited. In the Sidmanian paradigm, no such possibility is implied, although modest interventions for language training and reading programmes have been attempted by equivalence researchers (see J. T. de Rose, de Souza, Rossito, & T. M. S. de Rose, 1992; Matos, Avanzi, & McIlvane, 2006; Wilkinson & McIlvane, 2001; da Costa, Wilkinson, de Souza, & McIlvane, 2001).

One advantage of the pragmatic research approach adopted by RFT is that it leads to a research strategy in which the invariant status of constructs, such as intelligence, are arrived at through inductive research means, rather than derived or contrived statistically in advance of research efforts. This pragmatic and inductive approach is illustrated well in a commentary made by Hayes (1993) in a discussion of Lerner's (1993) epigenetic approach to human development. Specifically, Lerner's account suggests that there may exist predetermined genetic limits to human development. From a behavioural point of view, Hayes explained, there are no limits to behavioural development until they have been reached through exhaustive attempts to create exceptionally stimulating environments. In Hayes' (1993) words; "Lerner seems too quick to say how high pygmies can grow or how well a person with Down's syndrome can do. There presumably are such limits, but we cannot know when we have reached them" (p. 319). For both epistemological and ethical reasons, therefore, a behaviour analyst would never accept that the IQ of a person presenting with Down's syndrome had limit, or was invariant. Hayes goes on to explain that Lerner leaves the door open for problems by allowing for behavioural causality by genetics, and thereby limiting his analysis to only those environments that have already been studied (p. 317). The RFT account, in contrast, is situated within a short tradition of contextualistic science (Hayes, 1993b) and within a much longer tradition of behavioural pragmatism (see Barnes-Holmes, 2000).

This position has served to foster an approach within which people who might not be served by those who believe in psychological invariants, are served by the attempt to enhance core psychological abilities, precisely because limits to our ability to create extraordinary environments, and therefore extraordinary behaviour, are not anticipated.

While the suggestion that IQ scores can be raised may be novel to psychometricians, it certainly is not new to behaviour analysts. Previous behaviour-analytic studies have shown increases in IQ test measures during interventions for severe disability (see, for example, Lovaas, 1987; see also Reed, Osbourne & Corness, 2005 for a review; see Sallows & Graupner, 2005, for a replication of Lovaas; see also Connor, 1998; Gresham & MacMillan, 1997 for criticisms; see Magiati & Howlin, 2001 for further criticisms). Thus, it is important to explicitly acknowledge the efforts of other non RFT-based research which attempts to tackle similar challenges in more traditional ways, typically employing the methods of precision teaching and applied behaviour analysis, more generally. Nevertheless, the success or failure of these approaches notwithstanding, it is important to pursue every avenue of research in order to develop the optimally effective means of helping children with intellectual deficit. RFT offers not only the same behavioural techniques available to any behaviour analysts, but also provides a fresh and innovative research paradigm and terminology with which to generate amenable research questions to address the problem of intellectual deficit.

Outside of the field of behaviour analysis, there is an abundance of research into improving various types of cognitive skills. For example, there exists a whole industry in the United States dedicated to improving scores on standardised college-entrance exams (e.g., SAT, ACT, et cetera) and to graduate level entrance exams (e.g., GRE, LSAT, et cetera). Industry can supply interested individuals with numerous practice tests, advice, short-courses, et cetera on how to score highly on these exams. However, while these programs are offering a great deal of “practice” at taking these tests, their aim is not to raise IQ, but merely to increase a person’s chances of scoring highly in the attainment areas targeted by each specific examining board. For example, the SAT and the ACT are exams which many American students are required to take (and to score within certain parameters on) to gain entrance to university in the United States. Therefore, most American high school students, who have a view to applying to colleges, will be aware that the SAT consists of a section on mathematics, a section on critical reading and a section on writing. Most American students would also be aware that the ACT consists of a section on English reading, a section on mathematics, a section on science and an optional writing section. One only has to enter the names of these standardised achievement tests into any search engine to reveal a whole host of companies (e.g., Kaplan, PrepMe, Inc, Sylvan Learning et cetera) offering preparatory courses which promise to increase scores on these tests. However, what these courses frequently offer (often at great expense) are practice at the targeted test in various formats. Thus, if a student wishes to enrol in such a preparatory course, they can be assured that they will get practice at taking old SATs and ACTs in either individual, small group setting, or in an on-line format. From a psychometrician’s perspective, these skill domains are classified as attainment areas (e.g., literacy, numeracy) separate from ability areas (e.g., verbal IQ, performance IQ, full scale IQ). In fact, if a child’s score in an attainment area was significantly lower than what would be predicted given a child’s full scale IQ, then this would warrant a diagnosis of a specific learning difficulty or disability (see

Wechsler, 2005; Wechsler, 1996a; 1996b; 1996c). In effect, the American test preparation industry interventions aim merely to enhance what would be seen here as attainment ability. In contrast, the current research aimed to achieve rises in relational ability but with a view to also testing the idea that these improvements should in turn be associated with higher full scale IQ.

Conclusion

Over the course of eleven experiments, a RFT-based multiple exemplar training intervention has been developed and tested. This intervention programme has aimed to generate repertoires of relational responding across various relational domains (i.e., symmetry, transitivity, Same, Opposite, More-Than/Less-Than). The results have shown that the developed MET interventions have been more successful at generating DRR repertoires for the examined relations than exposure to training and unreinforced testing alone. The research has also shown that, of all the relational skills studied, Opposite relational responding is the most difficult to establish. It was also found that correlations between various IQ subtests and DRR ability are both non-linear and highly variable. Much work still remains to be done to elucidate the various relations that obtain between IQ subtest performance and DRR ability.

Perhaps the most exciting and promising finding of the current research was that IQ rises were seen for participants who had access to the MET relational intervention in Chapters 2, 5 and 6. In Chapters 2 and 5, modest and significant IQ rises, respectively, were seen for the longitudinal group of participants who had access to MET across several procedures for relations of symmetry, transitivity, Same, Opposite and More-Than/Less-Than. In Chapter 6, a mean IQ rise of 13 points was seen for a group of participants presenting with learning difficulties. This exciting and heartening finding suggests that despite several outstanding methodological issues, modern behaviour analysts, adopting the RFT approach, are now in a position to make meaningful and potentially powerful additional contributions to the bodies of research of those interested in developmental disabilities, educational psychology and clinical behaviour analysis, more generally. Perhaps more importantly, and as hoped for at the outset, behaviour analysts may now well be in possession of some novel tools that will aid in making real and meaningful differences to the lives of those educationally challenged children who most need our help.

References

Arntzen, E., & Holth, P. (2000). Equivalence outcome in single subjects as a function of training structure. *The Psychological Record, 50,* 603–628.

Baer, R. A., Detrich, R., & Weninger, J. M. (1988). On the functional role of the verbalization in correspondence training procedures. *Journal of Applied Behavior Analysis, 21,* 345-356.

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

Barnes, D., Lawlor, H., Smeets, P. M., & Roche, B. (1996). Stimulus equivalence and academic self-concept among mildly mentally handicapped and nonhandicapped children. *The Psychological Record, 46,* 87-107.

Barnes, D., McCullagh, P. D., & Keenan, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *The Analysis of Verbal Behavior, 8,* 19-30.

Barnes, D., & Roche, B. (1996). Relational frame theory and stimulus equivalence are fundamentally different: A reply to Saunders, *The Psychological Record, 46*, 489-508.

Barnes, D., & Roche, B. (1997). Relational frame theory and the experimental analysis of human sexuality. *Applied and Preventive Psychology, 6,* 117-135.

Barnes, D., & Roche, B. (1997). A behavior-analytic approach to behavioral reflexivity. *The Psychological Record, 47,* 543-572.

Barnes, D., Smeets, P. M., & Leader, G. (1996). Procedures for generating emergent matching performances: Implications for stimulus equivalence. In T. R. Zentall

& P. M. Smeets (Eds.), *Stimulus class formation: Advances in psychology* (pp. 153-171). Holland: Elsevier.

Barnes-Holmes, D. (2000). Behavioral pragmatism: No place for reality and truth. *The Behavior Analyst, 23,* 191-202.

Barnes-Holmes, D., & Barnes-Holmes, Y. (2000). Explaining complex behaviour: Two perspectives on the concept of generalized operant classes. *The Psychological Record, 50,* 251-265.

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's Verbal Behavior: A possible synthesis. *The Behavior Analyst, 23,* 69-84.

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure as an indirect measure of implicit beliefs. *The Irish Psychologist, 32,* 169-177.

Barnes-Holmes, Y., Barnes-Holmes D., & Roche, B. (2001). Exemplar training and a derived transformation of function in accordance with symmetry. *The Psychological Record, 51,* 287-308.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., Healy, O., Lyddy, F., Cullinan, V., & Hayes, S. C. (2001). Psychological development. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition,* (pp. 161). New York: Plenum.

Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. (2001). Exemplar training and a derived transformation of function in accordance with symmetry II.

The Psychological Record, 51, 589-603.

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

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P., Strand, P., & Friman, P. (2004). Establishing relational responding in accordance with more-than and less-than as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4,* 531-558.

Barnes-Holmes, D., Healy, O., & Hayes, S. C. (2000). Relational frame theory and the relational evaluation procedure: Approaching human language as derived relational responding. In J. C. Leslie & D. E. Blackman (Eds.), *Experimental and applied analyses of human behavior* (pp. 149-180). Reno, NV: Context Press.

Baum, W. M. (2004). The accidental behaviourist: A review of The New Behaviorism by John Staddon. *Journal of the Experimental Analysis of Behaviour, 82,* 73-78.

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

Boelens, H. (1994). A traditional account of stimulus equivalence. *The Psychological*

Record, 44, 587-605.

Bors, D., & Forrin, B. (1995). Age, speed of information processing, recall, and fluid intelligence. *Intelligence, 20,* 229-248.

Calcagno, S., Dube, W. V., De Faria Galvao, O., & Sidman, M. (1994). Emergence of conditional discriminations after constructed-response matching-to-sample training. *The Psychological Record, 44,* 509-520.

Carroll, J. B. (1989). Factor Analysis since Spearman: Where do we stand? What do we know? In R. Kanfer, P. L. Ackerman & R. Cudek (Eds.), *Abilities, motivation and methodology* (pp.43-67). Hillsdale NJ: Lawrence Erlbaum.

Catania, A. C., Shimoff, E., & Matthews, B. A. (1987). Correspondence between definitions and procedures: A reply to Stokes, Osnes, and Guevremont. *Journal of Applied Behavior Analysis, 20,* 401-404.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40,* 153-193.

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

Cattell, R. B. (1971). *Abilities, their structure, growth and action.* Boston: Houghton Mifflin.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Second Edition. Hillsdale, NJ: Erlbaum.

Cohen, J. D., McWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments and Computers, 25,* 257-271.

Connor, M. (1998). A review of behavioural early intervention programmes for children with autism. *Educational Psychology in Practice, 14,* 109-117.

Cooper, J., Heron, T., & Heward, W. (2007). *Applied Behavior Analysis*. Second Edition. New Jersey: Pearson Education, Inc.

Cullinan, V., Barnes, D., & Smeets, P. M. (1998). A precursor to the relational evaluation procedure: Analyzing stimulus equivalence. *The Psychological Record, 48,* 121-145.

Cullinan,V., Barnes-Holmes, D., & Smeets, P. M. (2000). A precursor to the relational evaluation procedure: Analyzing stimulus equivalence II. *The Psychological Record, 50,* 467-492.

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

da Costa, A. R., Wilkinson, K. M., de Souza, D. G., & McIlvane, W. J. (2001). Emergent word-object mapping by children: Further studies using the blank comparison technique. *The Psychological Record*, *51,* 343-355.

D'Amato, M. R., & Colombo, M. (1985). Auditory matching-to-sample in monkeys (Cebus appella). *Animal Learning and Behavior*, *13,* 375-382.

D'Amato, M. R., & Worsham, R. W. (1974). Retrieval cues and short-term memory in capuchin monkeys. *Journal of Comparative and Physiological Psychology, 86,* 274-282.

Department of Education and Science. (2003). Education for Persons with Special Educational Needs Bill. Government Publications, Dublin.

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

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46,* 243-257.

Dickins, D., Singh, K., Roberts, N., Burns, P., Downes, J., Jimmieson, P., & Bentall, R. (2000). An fMRI study of stimulus equivalence, *NeuroReport, 12 (2),*1-7.

Dube, W. V., Green, G. & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 59,* 103-114.

Dube, W. V., McDonald, S. J., McIlvane, W. J., & Mackay, H. A. (1991). Constructed-response matching-to-sample and spelling instruction. *Journal of Applied Behavior Analysis, 24,* 305-317.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D.E. Blackman & H. Lejeune (Eds.), *Behaviour Analysis in Theory and Practice* (pp. 115-138). Hove: Lawrence Erlbaum.

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64,* 163-184.

Dymond, S., & Barnes, D. (1996). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness and opposition. *The Psychological Record, 46,* 271-300.

Dymond, S., & Barnes, D. (1997). Behavior analytic approaches to self-awareness. *The Psychological Record, 47*, 181-200.

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior, 88, 249-262.*

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *The Psychological Record*, *58*, 271-288.

Edwards, C. A., Jagielo, J. A., & Zentall, T. R. (1982). Acquired equivalence and distinctiveness in matching to sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes*, *8,* 244-259.

Eysenck, H. J. (1982). *A Model for intelligence.* New York: Springer.

Fields, L, Adams, B. J., & Verhave, T. (1993). The effects of equivalence class structure on test performances*.* *The Psychological Record, 43,* 697-712.

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.

Fields, L., Reeve, K. F., Varelas, A., Rosen, D., & Belanich, J. (1997). Equivalence class formation using stimulus-pairing and yes-no responding. *The Psychological*

Record, 47, 661-686.

Fredericksen, J. R. (1982). A componential theory of reading skills and their interactions. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence.* Volume 1, (pp 125-180). Hillsdale, NJ: Erlbaum.

Friman, P. C., Hayes, S. C., & Wilson, K. G. (1998). Why behavior analysts should study emotion: The example of anxiety. *Journal of Applied Behavior Analysis, 31*, 137-156.

Galizio, M. (2003). The abstracted operant: A review of Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition, edited by S. C. Hayes, D. Barnes-Holmes, and B. Roche. *The Behavior Analyst, 26,* 159–169.

Gardner, H. (1993). *Multiple Intelligences: The theory in practice.* NY: Basic Books.

Gomez, S., Lopez, F., Martin, C. B., Barnes-Holmes, Y., & Barnes-Holmes, D. (2007). Exemplar training and a derived transformation of function in accordance with symmetry and equivalence. *The Psychological Record, 57,* 273-294.

Gresham, F. M., & MacMillan, D. L. (1997). Autistic recovery? An analysis and critique of the empirical evidence on the Early Intervention Project. *Behavioral Disorders, 22,* 185-201.

Guerin, B. (1994). *Analyzing Social Behaviour.* Reno, NV: Context Press.

Harlow, H. (1949). The formation of learning sets. *Psychological Review, 56,* 51–65.

Hart, B., & Risley, T. (1995). *Meaningful Differences in the Everyday Experience of Young American Children.* Brookes: Baltimore, MD.

Hartlage, L. C., & Steele, C. T. (1977). WISC and WISC-R correlates of academic achievement. *Psychology in the Schools, 14,* 15-18.

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: Context Press.

Hayes, S. C. (1992). Verbal relations, time and suicide. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations* (pp. 109-118). Context Press: Reno, NV.

Hayes, S. C. (1993a). Discussion of Lerner, nature-nurture: two-headed arrows and wrong-headed questions. In Hayes, S. C., Hayes, L. J., Reese, H. W., & Sarbin, T. R. (Eds.), *Varieties of Scientific Contextualism* (pp. 317-319).

Hayes, S. C. (1993b). Analytic goals and the varieties of scientific contextualism. In Hayes, S. C., Hayes, L. J., Reese, H. W., & Sarbin, T. R. (Eds.), *Varieties of Scientific Contextualism* (pp. 23-27).

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

Hayes, S. C., & Barnes, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior*, *68*, 235-270.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A post-Skinnerian account of human language and cognition.* NY: Plenum.

Hayes, S. C., Blackledge, J. T., & Barnes-Holmes, D. (2001). Language and Cognition; Constructing an alternative approach within the behavioral tradition. In Hayes, Barnes-Holmes, & Roche (Eds.), *Relational Frame Theory: A post-Skinnerian* *account of human language and cognition* (pp. 9-14). NY: Plenum.

Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations, and a behaviour-analytic view of the purposes of science. *The Behavior Analyst, 9,* 175-190.

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 human language and cognition* (pp. 26-27). New York: Plenum.

Hayes, S. C., & Hayes, L. J. (1992). Verbal relations and the evolution of behaviour analysis. *American Psychologist, 47,* 1383-1395.

Hayes, S. C., & Wilson, K. G. (1993). Some applied implications of a contemporary behavior-analytic account of verbal events. *The Behavior Analyst, 16*, 283-301.

Hayes, S. C., & Wilson, K. G. (1996). Criticisms of relational frame theory: Implications for a behavior analytic account of derived stimulus relations. *The Psychological Record, 46,* 231-236.

Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of Experimental Analysis of Behavior 74 (2)*, 207-227.

Heim, A.W., Watts, K. P., & Simmonds, V. (1968, 1975). AH4 Question Book. Berkshire: NFER-Nelson Publishing Company Ltd.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence* (pp. 267-300). New York; John Wiley & Sons.

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.

Hunt, E. (1985). Verbal Ability. In R. J. Sternberg (Ed.), *Human abilities: An information processing approach* (pp.31-58). New York: Freeman.

Jensen, A. R. (1980). *Bias in mental testing.* NY: The Free Press.

Jensen, A. R. (1982). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93-132). Prenger-Verlag.

Jensen, A. R. (1998). The g factor and the design of education. In R. J. Sternberg & W. M. Williams (Eds.), *Intelligence, instruction, and assessment: Theory into practice* (pp. 111-131). Mahwah, NJ: Lawrence Erlbaum.

Juliano, J. M., Haddad, F. A., & Carroll, J. L. (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,* 317-325.

Kastak, D., & Schusterman, R. J. (1994). Transfer of visual identity matching-to-sample in two California sea lion (Zalohus Californianus). *Learning and Behavior, 22, (4),* 427-435.

Kastak, C. R., & Schusterman, R. J. (2002). Sea lions and equivalence: expanding classes by exclusion. *Journal of the Experimental Analysis of Behavior, 78,* 449-465.

Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior, 76,* 121-258.

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

Kaufman, A. S., & Kaufman, N. L. (1990). *The Kaufman Brief Intelligence Test.* Minnesota: American Guidance Service.

Kyllonen, P. C., Lohman, D. F., &, Woltz, D. J. (1984). Componential modelling of alternative strategies for performing spatial tasks. *Journal of Educational Psychology, 76,* 1325-1345.

Lattal, K. A., & Chase, P. N. (2003). *Behavior Theory and Philosophy.* London: Kluwer Academic/Plenum Publishers.

Lazar, R. M. (1977). Extending sequence-class membership with matching-to-sample. *Journal of the Experimental Analysis of Behavior, 27,* 381-392.

Lazar, R. M., & Kotlarchyk, B. J. (1986). Second-order control of sequence-class equivalences in children. *Behavioural Processes, 13,* 205-215.

Leader, G., Barnes, D., & Smeets, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record, 46,* 685-706.

Leader, G., Barnes-Holmes, D., & Smeets, P. M. (2000). Establishing equivalence relations using a respondent-type training procedure III. *The Psychological Record, 50,* 63-79.

Lipkens, G., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology, 56,* 201- 239.

Lloyd, K. E. (1980). Do as I say, not as I do. *New Zealand Psychologist, 9,* 1-8.

Lohman, D. F. (1988). Spatial abilities as traits, processes and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence 4* (pp.

181-248). Hillsdale, NJ: Erlbaum.

Lohman, D. F. (1989). Human Intelligence: An introduction to advances in theory and research. *Review of Educational Research, 59 (4),* 333-373.

Lovaas, O. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting Clinical Psychology, 55 (1),* 3-9.

Mackay, H. M., & Sidman, M. (1984). Teaching new behavior via equivalence relations. In P. H. Brooks, R. Sperber, & C. McCauley (Eds.), *Learning and cognition in the mentally retarded* (pp. 493-513). Hillsdale, NJ: Lawrence Erlbaum.

Magiati, I., & Howlin, P. A. (2001). Monitoring the progress of preschool children with autism enrolled in early intervention programmes: Problems in cognitive assessment. *Autism, 5,* 399-406.

Malott, R. (2003). Behaviour analysis and linguistic productivity. *The Analysis of Verbal Behavior, 19,* 11-18.

Markham M. R., & Dougher, M. J. (1993). Compound stimuli in emergent stimulus relations: Extending the scope of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 60,* 529-542.

Matos, M. A., Avanzi, A. L., & McIlvane, W. J. (2006). Rudimentary reading repertoires via stimulus equivalence and recombination of minimal verbal units. *Analysis of Verbal Behavior, 22,* 3-19.

McIlvane, W. (2003). A Stimulus in need of response: A Review of “Relational Frame Theory; a post Skinnerian account of human language and cognition. *The Analysis of Verbal Behavior, 19,* 29-37.

McKenna, I., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP); The first study. *International Journal of Psychology and Psychological Therapy, 7,* 253-268.

Murphy, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2005). Derived manding in children with autism: Synthesizing Skinner's Verbal Behavior with relational frame theory. *Journal of Applied Behavior Analysis, 38,* 445-462.

National Council for Special Education. (2008). Information and Guidelines for Primary Schools, Special Schools and Second Level Schools in Processing Applications for Resources for Pupils with Special Educational Needs.

O'Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record, 54,* 437-460.

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

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

O'Hora, D., Roche, B., Barnes-Holmes, D., & Smeets, P. M., (2002). Response latencies to multiple derived stimulus relations: Testing two predictions of RFT. *The Psychological Record, 52,* 51-75.

O'Toole, C., & Barnes-Holmes, D. (in press). Three chronometric Indices as predictors of relational responding on a brief intelligence test: the importance of relational flexibility. *The Psychological Record.*

Pallant, J. (2001). *SPSS Survival Manual: A Step by step guide to data analysis using SPSS for Windows (Versions 10 and 11).* Buckingham and Philadephia: Open University Press.

Paniagua, F. A. (1985). Development of self-care skills and helping behaviors in adolescents in a group home through correspondence training. *Journal of Behavior Therapy and Experimental Psychiatry*, *16,* 237-244.

Pellegrino, J. W., & Kail, R. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence*. Volume 1, (pp. 311-366). Hillsdale, NJ: Erlbaum.

Pilgrim, C., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: 1. Adults. *Journal of the Experimental Analysis of Behavior, 63,* 225-238.

Pilgrim, C. & Galizio, M. (1996). Stimulus equivalence: A class of correlations, or a correlation of classes? In Zentall T.R, Smeets P.M., (Eds.), *Stimulus class formation in humans and animals.* New York: Elsevier Science (pp. 173–195).

Pilgrim, C., & Galizio, M. (2000). Stimulus equivalence and units of analysis. In J.C. Leslie and D. Blackman (Eds.), *Experimental and Applied Analysis of Human Behavior*. Context Press: Reno, NV. (pp. 111-126).

Premack, D. (2004). Is language the key to human intelligence? *Science, 303,* 318-320.

Reed, P., Osborne, L., & Corness, M. (2005). The effectiveness of early intervention programmes for autistic spectrum disorders. *A Report for the South East Regional Special Educational Needs Partnership.* Research Partners: Bexley, Brighton & Hove, East Sussex, Kent, Mdway, Surrey, West Sussex.

Reschly, D. J. &, Reschly, J. E. (1979). Validity of WISC-R factor scores in predicting achievement and attention for four sociocultural groups. *Journal of School Psychology, 17,* 355-361.

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

Ribero, A. de F. (1989). Correspondence in children's self-report: Tacting and manding aspects. *Journal of the Experimental Analysis of Behavior, 51,* 361-367.

Riegler, H. C., & Baer, D. M. (1989). A developmental analysis of rule-following. In H. W. Reese (Ed.), *Advances in child development and behavior.* Volume 21. (pp. 191-219). New York: Academic Press.

Roche, B., & Barnes, D. (1996a). Arbitrarily applicable relational responding and human sexual categorization: a critical test of the derived difference relation. *The Psychological Record, 46,* 451-475.

Roche, B., & Barnes, D. (1996b). Sexual behaviour as an act in context. *The Irish Journal of Psychology, 16,* 409-417.

Roche, B., & Barnes, D. (1997). A transformation of respondently conditioned functions in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior, 67,* 275-301.

Roche, B., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & O'Hora, D. (2002). Relational Frame Theory: A new paradigm for the analysis of social behavior. *The Behavior Analyst, 25*, 75-91.

Roche, B., & Dymond, S. (1998). The experimental analysis of human sexual arousal: some recent developments. *The Behavior Analyst, 21,* 37-52.

Roche, B., & Dymond, S. (1999). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments and Computers, 25*, 257-271.

Roche, B., & Dymond, S. (2008). A transformation of functions in accordance with the non-arbitrary relational properties of sexual stimuli. *The Psychological Record, 58,* 71-90.

Sallows, G. O., & Graupner, T. D. (2005). *Replicating Lovaas' treatment and findings: Preliminary results.* PEACH. Putting research into Practice Conference, London.

Sattler, J. M. (1988). *Assessment of children.* Third Edition. San Diego, CA: Jerome Sattler.

Saunders, R. (1996). From Review to Commentary on Roche and Barnes: Toward a Better Understanding of Equivalence in the Context of Relational Frame Theory *The Psychological Record, 46*, 477-487.

Schauss, S. L., Chase, P. N., & Hawkins, R. P. (1997). Environment-behaviour relations, behaviour therapy and the process of persuasion and attitude change. *Behaviour Therapy of Experimental Psychiatry, 1,* 31-40.

Schenk, J. J. (1995). Complex stimuli in nonreinforced simple discrimination tasks: Emergent simple and conditional discriminations. *The Psychological Record, 45,* 477-494.

Schlinger, H. D. (2003). The Myth of intelligence. *The Psychological Record, 53,* 15-32.

Sidman, M. (1960). Tactics of scientific research. New York: Basic Books Boston.

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

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. E. Zeiler (Eds.), *Analysis and Integration of Behavioral Units* (pp 231-245). Hillside, NJ: Lawrence Erlbaum Associates.

Sidman, M. (1994). Stimulus Equivalence: A Research Story. Authors Cooperative, Boston.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74,* 127-146.

Sidman, M. (2008). Symmetry and equivalence relations in human behaviour. *Cognitive Studies, 15 (3),* 1-11.

Sigurdadottir, Z. G., Green, G., & Saunders, R. R. (1990). Equivalence classes generated by sequence training. *Journal of the Experimental Analysis of Behavior, 53,* 47-64.

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

Skinner, B. F. (1959). *Cumulative Record.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1974). *About Behaviorism.* New York: Random House Inc. and Toronto: Random House of Canada Limited.

Smeets P. M., Barnes-Holmes Y, Akpinar, D., & Barnes-Holmes, D. (2003). Reversal of equivalence relations. *The Psychological Record, 5*3, 91–119.

Smeets, P. M., Leader, G., & Barnes, D. (1997). Establishing stimulus classes with adults and children using a respondent training procedure: A follow-up study. *The Psychological Record, 47,* 285-308.

Smeets, P. M., Schenk, J. J., & Barnes, D. (1994). Establishing arbitrary stimulus classes via identity matching and nonreinforced matching with complex stimuli. *Quarterly Journal of Experimental Psychology, 48B,* 311-328.

Smeets, P. M., & Striefel, S. (1994). Matching to complex stimuli under non-reinforced conditions: Errorless transfer from identity to arbitrary matching tasks. *Quarterly Journal of Experimental Psychology, 47B,* 39-62.

Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, O. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. *American Journal on Mental Retardation, 102 (3),* 238-249.

Snow, R. E. (1981). Toward a theory of aptitude for learning: Fluid and crystallized abilities and their correlates. In M. P. Friedman, J. P. Das & N. O'Connor (Eds.), *Intelligence and learning* (pp. 345-362). New York: Plenum Press.

Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. Linn (Ed.), *Educational Measurement.* Third Edition. (pp. 263-331). New York: Macmillan.

Spearman, C. (1904) "General Intelligence". Objectively determined and measured. *American Journal of Psychology, 15,* 201-293.

Steele, D. L., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56,* 519-555.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence.* Cambridge, MA: Cambridge University Press.

Stewart, I., & Barnes-Holmes, D. (2001). Understanding metaphor: A relational frame perspective. *The Behavior Analyst, 24,* 191-199.

Stewart, I., & Barnes-Holmes, D. (2004). Relational frame theory and analogical reasoning: Empirical investigations. *International Journal of Psychology and Psychological Therapy, 4,* 241-262.

Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors, and stories. In S. C. Hayes, D. Barnes-Holmes & B. Roche (Eds.), *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (p.76). New York: Plenum.

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

Stewart, I., Barnes-Holmes, D., Roche, B., & Smeets, P. (2001). Generating derived relational networks via the abstraction of common physical properties: A possible model of analogical reasoning. *The Psychological Record, 51,* 381-408.

Stromer, R. R. (1991). Stimulus equivalence: Implications for teaching. In W. Ishaq

(Ed.), *Human behavior in today's world* (pp. 109-122). New York: Preager.

Stromer, R. R., & Mackay, H. A. (1993). Human sequential behavior: Relations among stimuli, class formation, and derived sequences. *The Psychological Record, 43,* 107-131.

Stromer, R. R., McIlvane, W. J., & Serna, R. W. (1993). Complex stimulus control and equivalence. *The Psychological Record, 43*, 584-598.

Thorndike, R., Hagen, E., & France, N. (1986). Cognitive Abilities Test (Levels A to F), Second Edition, Administration Manual. Berkshire, NFER-Nelson.

Vitale, A. (2003). *The relational frame of comparison: A systematic empirical analyses*. Unpublished doctoral dissertation. National University of Ireland, Maynooth.

Vitale, A., Barnes-Holmes, Y., Barnes-Holmes, D. & Campbell, C. (2008). Facilitating responding in accordance with the relational frame of comparison: Systematic empirical analyses. *The Psychological Record, 58,* 365-390.

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*, 33-50.

Wechsler, D. (1944). *The measurement of adult intelligence*. Third Edition. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (1981). *The Wechsler Adult Intelligence Test.* San Antonio: The Psychological Corporation.

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

Wechsler, D. (1992). *Wechsler Intelligence Scale for Children.* Third Edition. UK Manual, London: The Psychological Corporation Europe.

Wechsler, D. (1996a). *Wechsler Objective Language Dimensions.* UK Manual, London: The Psychological Corporation Europe.

Wechsler, D. (1996b). *Wechsler Objective Literacy Dimensions.* UK Manual, London: The Psychological Corporation Europe.

Wechsler, D. (1996c). *Wechsler Objective Numeracy Dimensions*. UK Manual, London: The Psychological Corporation Europe.

Wechsler. D. (2003). *Wechsler Intelligence Scale for Children.* Fourth Edition. Technical and Interpretive Manual. San Antonio, Texas, Harcourt Assessment.

Wechsler. D. (2004). *Wechsler Intelligence Scale for Children.* Fourth Edition. Administration and Scoring Manual. London, Harcourt Assessment.

Wechsler, D. (2005). *WIAT-II^{UK}, Wechsler Individual Achievement Test.* Second UK Edition. UK scoring and normative supplement. Ages 4 years to 16 years 11 months. London: Harcourt Assessment.

Whelan, R., Barnes-Holmes, D., & Dymond, S. (2006). The transformation of consequential functions in accordance with the relational frames of more-than and less-than. *Journal of the Experimental Analysis of Behavior*, *86*, 317-335.

Wilkinson, K. M., & McIlvane, W. J. (2001). Considerations in teaching graphic symbols. In J. Reichle, D. Beukelman, & J. Light (Eds.), *Exemplary strategies for beginning communicators*: *Implications for AAC*. (pp. 273-321). Baltimore, MD: Brookes.

Williams, B., Myerson, J., & Hale, S. (2008). Individual differences, intelligence, and behavior analysis. *Journal of the Experimental Analysis of Behavior, 90,* 219-231.

Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction.* New York: Columbia University Press.

Zuriff, G. E. (2005). Behaviorism makes it debut: A review of Lattal and Chase's Behavior Theory and Philosopy. *Journal of the Experimental Analysis of Behavior, 83,* 315-322.

Appendix 1

Developing an Intervention Programme to Raise IQ
Outline of the Research Being Conducted by Sarah Cassidy, NUI, Maynooth

The current research project is designed to test the utility of a psychological intervention in raising the intellectual abilities of both normal and intellectually challenged children. In order to develop and test the intervention technique, we will need a range of volunteers whose IQ we can measure both before and after our pilot intervention programme. This research is doctoral level research being conducted by Sarah Cassidy at NUI, Maynooth, under the supervision of Dr. Bryan Roche.

The volunteers we are seeking will be between the ages of 8 and 12 yrs. The children will not be made uncomfortable or endangered in any way. All that is required of them is to complete a series of puzzles on a computer. These puzzles are designed to increase the child's intellectual skills and general problem-solving ability. Thus, it is likely that each child participating in the study will benefit in some way.

Each child will be sat individually in front of a computer for 2-3 thirty minute sessions per week for several weeks. Depending on the results of this first stage of the study, we may ask to see particular children again for a small number of follow-up sessions throughout the year.

We will also be seeking a number of adult participants. Their participation will take place in an identical manner to that described for child participants.

The specific psychological theory that forms the context for the current study is known as Relational Frame Theory (RFT). RFT is a behavioural theory of human language and cognition which claims that the foundational skill for most intellectual abilities is a skill known as Derived Relational Responding (DRR). DRR is the skill of relating objects to each other in accordance with a small family of mathematical relationships (e.g., equivalence, opposition, etc.). RFT also suggests the format of an intervention for intelligence deficits, known as Multiple Exemplar Training (MET). In effect, MET involves training children in core relational skills, such as matching stimuli that are similar or opposite to each other. Once component relational skills are taught, a child should be able to solve an infinite number of other relational problems. Moreover, these improvements should have a quantifiable effect on intelligence scores.

Initially, the study will focus on equivalence relations between stimuli. That is, we will be teaching children to form categories of words that are the same as (or equivalent) to each other. The skill of forming equivalence relations lies in the important fact that each of the stimuli in an equivalence relation bears the same relationship to all other members. For example, if a television is the same as a TV and a TV is the same as a Telly, then a Television must be the same as a Telly. An intervention for intelligence based on equivalence training involves teaching children many such examples using a specifically written piece of computer software.

A Typical Matching-Game Task. The participant is taught to match the word ZID to the word PAF at the top of the computer screen. An equivalence training intervention involves teaching children to learn several such matches between words and to form classes of words that go or do not go together.

After several such intervention sessions, we expect that each participant's ability to form these relations and other relations will increase. The purpose of the study is to determine the best manner in which to conduct these interventions so that improvements in relational skills lead to changes in overall intelligence scores.

Obviously the study cannot be conducted without the help of volunteers and their parents. We very much appreciate the help of all participants and their families and believe that all will benefit from participation in this research.

Parents and teachers should feel free to contact me with any further questions at ***-*******, e-mail me at ***********@*******, or write to me at Sarah Cassidy, Department of Psychology, John Hume Building, NUI Maynooth, Maynooth, Co. Kildare.

Appendix 2

Consent Form

In agreeing to participate in the research project "Developing a Multiple Exemplar Training Intervention to Raise IQ", I understand the following:
Sarah Cassidy, Department of Psychology, NUI Maynooth and Dr. Bryan Roche, Department of Psychology, NUI Maynooth, are conducting this research. Ms. Cassidy is the principal investigator and can be contacted at (***)***-****. The principal investigator has worked extensively in Irish primary and secondary schools, as a psychologist, administering psychological assessments.

The purpose of this psychological research is to analyse different components of intelligence using a theory known as Relational Frame Theory. Each participant will be asked to engage in several different tasks that will involve problem-solving skills. These tasks are typical of tasks used in common intelligence tests. These tasks will be repeated several weeks later. Each participant will also receive some exemplar training to see if his/her problem solving skills improve. Some of these tasks may be timed, but at no time will the participant be under any undue stress. These tasks are set up as games and as such, will cause no pain or discomfort.

Specifically, each participant will be seated comfortably in front of a computer and presented with a range of tasks requiring them to learn a series of word-pairs. They will also be tested to see if they can correctly match these words together. This phase normally takes approximately 15-20 minutes, but can vary across participants.

Each participant will also be provided with a training phase designed to improve his/her performance on the word-matching test. This phase will take place across several sessions across several weeks.

Each participant will also be administered a series of cognitive abilities tests that are used in standard intelligence tests. This phase will take approximately 2 hours, but the participant can take as many breaks as he/she needs. These tests are designed to measure any improved cognitive functioning that participants may experience following the training phase.

All persons participating in this study will remain anonymous 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 and will not be noted on any school records. The data will be kept in a locked cabinet in the Experimental Psychology Laboratory in the Psychology Department at the National University of Ireland, Maynooth. This data is available to each participant or to parents/guardians of participants at their discretion and can be accessed at any time.

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

I understand that I may refuse participation for myself (or my child) or that I may withdraw myself (or my child) from the study at any stage even after giving my consent. I may withdraw my own (or my child's) data at the conclusion of participation if I still have concerns.

I understand that I may contact Dr. Bryan Roche at ***@***ie and/or the *Secretary of the National University of Ireland Maynooth Ethics Committee at ***@***.ie or 01 708 6***8* if I feel that any participant is experiencing any kind of discomfort/stress as a result of the study.

I understand that this experiment does not constitute any kind of counselling or medical treatment and that the study will not form any kind of medical diagnosis. I understand that it is experimental and not clinical in nature.

I have been informed as to the general nature of the study. I have read the research briefing. I understand that at the conclusion of my/my child's participation, any further questions or concerns I have will be fully addressed.

Signed:

______________________________ Participant (parent/guardian)

______________________________ Researcher

______________________________ Date

*If during your/your child's participation in this study you feel the information and guidelines that you were given have been neglected or disregarded in any way, or if you are unhappy about the process please contact the Secretary of the National University of Ireland Maynooth Ethics Committee at ***@***.ie of 01 708 ***8. Please be assured that you concerns will be dealt with in a sensitive manner.*

Appendix 3

Description of the WISC-III^{UK} subtests used in Experiment 1 and 2, adapted from the Manual for the WISC-III^{UK} (1992).

Picture Completion	A set of colourful pictures of common objects and scenes each of which is missing an important part which the child identifies.
Information	A series of orally presented questions that tap the child's knowledge about common events, objects, places and people.
Coding	A series of simple shapes (Coding A) or numbers (Coding B) each paired with a simple symbol. The child draws the symbol in its corresponding shape (Coding A) or under its corresponding number (Coding B), according to a key.
Similarities	A series of orally presented pairs of words for which the child explains the similarity of the everyday objects or concepts they represent.
Picture Arrangement	A set of colourful pictures, presented in a mixed-up order, which the child rearranges into a logical story sequence.
Arithmetic	A series of arithmetic problems which the child solves mentally and responds to orally.
Block Design	A set of modelled or printed two-dimensional geometric patterns which the child replicates using two-colour cubes.
Vocabulary	A series of words presented orally which the child defines
Object Assembly	A set of jig-saw puzzles of common objects, each presented in a standardised configuration, which the child assembles to form a meaningful whole.
Comprehension	A series of paired groups of symbols, each pair consisting of a target group and a search group. The child scans the two groups and indicates where or not a target symbol appears in the search group.
Digit Span	A series of orally presented number sequences which the child repeats verbatim for Digits Forward and in reverse order for Digits Backward.

Appendix 4

Arbitrary nonsense syllables employed as stimuli in Experiments 1 and 2.

A1	B1	C1	A2	B2	C2
ler	mau	cug	vek	rog	paf
wan	cil	yun	lon	cet	ril
ter	wev	hib	por	mip	jum
kon	lar	jey	wib	puh	zuj
pim	hep	luf	kib	sed	yoc
tuk	rol	vif	huv	geq	zay
kuv	jir	gep	mul	rec	hox

Appendix 5

Instructions for Baseline Conditional Discrimination Training (and MET testing) in Experiments 1 and 2.

In a Moment some words will appear on this screen. Your task is to look at the word at the top of the screen and choose one of the words at the bottom of the screen by "clicking on it" using the mouse. The computer will tell you whether or not your choice is correct. You should try to get as many answers correct as possible. This will take around half an hour to complete, but the harder you work the faster it will finish. The computer will tell you when you should take a break. If you have any questions, please ask them now. When you are ready to begin, please click the mouse button.

Appendix 6

Instructions for Testing.

In the next stage some more words will appear on this screen. Your task is to look at the word at the top of the screen and choose one of the two words at the bottom of the screen by clicking on it with the mouse. However, during this stage, the computer WILL NOT tell you if your choice was correct. This stage will be difficult but you must try to make as many correct choices as possible. Think carefully before you make your choice. When you are ready to begin, please click the mouse.

Appendix 7

IQ Subtest scores, full scale IQ scores, verbal IQ scores and performance IQ scores for each participant in Experiment 1.

P1

Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		10
Information	12	
Coding		11
Similarities	12	
Picture Arrangement		8
Arithmetic	12	
Block Design		9
Vocabulary	9	
Object Assembly		7
Comprehension	8	
(Symbol Search)		(7)
(Digit Span)	(10)	
Sum of Scaled Scores	53	45
Verbal IQ	103	
Performance IQ	92	
Full Scale IQ	98	

P2		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		15
Information	14	
Coding		10
Similarities	13	
Picture Arrangement		14
Arithmetic	12	
Block Design		10
Vocabulary	12	
Object Assembly		11
Comprehension	15	
(Symbol Search)		(12)
(Digit Span)	(13)	
Sum of Scaled Scores	66	60
Verbal IQ	120	
Performance IQ	115	
Full Scale IQ	119	

P3		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		13
Information	8	
Coding		9
Similarities	12	
Picture Arrangement		10
Arithmetic	13	
Block Design		9
Vocabulary	13	
Object Assembly		11
Comprehension	15	
(Symbol Search)		(8)
(Digit Span)	(10)	
Sum of Scaled Scores	61	52
Verbal IQ	113	
Performance IQ	103	
Full Scale IQ	109	

P4		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		8
Information	11	
Coding		9
Similarities	12	
Picture Arrangement		11
Arithmetic	13	
Block Design		8
Vocabulary	10	
Object Assembly		10
Comprehension	14	
(Symbol Search)		(11)
(Digit Span)	(13)	
Sum of Scaled Scores	60	46
Verbal IQ	111	
Performance IQ	94	
Full Scale IQ	104	

P5		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		11
Information	10	
Coding		12
Similarities	16	
Picture Arrangement		15
Arithmetic	9	
Block Design		8
Vocabulary	11	
Object Assembly		12
Comprehension	14	
(Symbol Search)		(14)
(Digit Span)	(9)	
Sum of Scaled Scores	60	58
Verbal IQ	111	
Performance IQ	112	
Full Scale IQ	113	

P6		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		10
Information	14	
Coding		14
Similarities	10	
Picture Arrangement		6
Arithmetic	10	
Block Design		14
Vocabulary	13	
Object Assembly		11
Comprehension	13	
(Symbol Search)		(12)
(Digit Span)	(10)	
Sum of Scaled Scores	60	55
Verbal IQ	111	
Performance IQ	107	
Full Scale IQ	111	

P7		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		11
Information	10	
Coding		13
Similarities	9	
Picture Arrangement		12
Arithmetic	10	
Block Design		9
Vocabulary	11	
Object Assembly		10
Comprehension	11	
(Symbol Search)		(10)
(Digit Span)	(10)	
Sum of Scaled Scores	51	55
Verbal IQ	101	
Performance IQ	107	
Full Scale IQ	104	

P8		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		7
Information	7	
Coding		9
Similarities	13	
Picture Arrangement		13
Arithmetic	10	
Block Design		6
Vocabulary	9	
Object Assembly		9
Comprehension	12	
(Symbol Search)		(12)
(Digit Span)	(12)	
Sum of Scaled Scores	51	54
Verbal IQ	101	
Performance IQ	91	
Full Scale IQ	96	

P9		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		10
Information	15	
Coding		14
Similarities	9	
Picture Arrangement		10
Arithmetic	13	
Block Design		9
Vocabulary	12	
Object Assembly		9
Comprehension	10	
(Symbol Search)		(15)
(Digit Span)	(7)	
Sum of Scaled Scores	59	52
Verbal IQ	110	
Performance IQ	103	
Full Scale IQ	107	

P10		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		12
Information	14	
Coding		18
Similarities	19	
Picture Arrangement		10
Arithmetic	15	
Block Design		11
Vocabulary	14	
Object Assembly		12
Comprehension	19	
(Symbol Search)		(14)
(Digit Span)	(12)	
Sum of Scaled Scores	81	63
Verbal IQ	138	
Performance IQ	119	
Full Scale IQ	133	

P11		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		8
Information	10	
Coding		16
Similarities	13	
Picture Arrangement		11
Arithmetic	12	
Block Design		11
Vocabulary	9	
Object Assembly		12
Comprehension	12	
(Symbol Search)		(11)
(Digit Span)	(9)	
Sum of Scaled Scores	56	58
Verbal IQ	107	
Performance IQ	112	
Full Scale IQ	110	

P12		
Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		10
Information	14	
Coding		7
Similarities	13	
Picture Arrangement		10
Arithmetic	11	
Block Design		18
Vocabulary	10	
Object Assembly		11
Comprehension	9	
(Symbol Search)		(6)
(Digit Span)	(14)	
Sum of Scaled Scores	57	56
Verbal IQ	108	
Performance IQ	109	
Full Scale IQ	109	

Appendix 8

Each participant's Baseline Conditional Discrimination training scores, Symmetry Testing scores, Transitivity Testing scores and Symmetry and Transitivity combined testing scores (SE measure). Experiment 2.

Participant	Baseline Condition Discrimination Training	Symmetry Testing	Transitivity Testing	Combined Symmetry and Transitivity Testing
P1	65	15	2	17
P2	55	4	35	39
P3	3	2	1	3
P8	27	2	3	5
P4	3	2	4	6
P6	10	2	2	4
P7	13	4	5	9
P9	9	1	2	3

Appendix 9

IQ Subtest scores, full scale IQ scores, verbal IQ scores and performance IQ scores for each participant at baseline and follow-up. Experiment 2.

Experimental P1				
Subtests	Verbal Scaled Scores at baseline	Verbal IQ Scaled Scores at follow-up	Performance IQ Scaled Scores at baseline	Performance IQ Scaled Scores at follow-up
Picture Completion			10	10
Information	12	12		
Coding			11	12
Similarities	12	11		
Picture Arrangement			8	12
Arithmetic	12	8		
Block Design			9	11
Vocabulary	9	10		
Object Assembly			7	10
Comprehension	8	11		
(Symbol Search)			(7)	(13)
(Digit Span)	(10)	(11)		
Sum of Scaled Scores	53	52	45	55
Verbal IQ	103	102		
Performance IQ			92	107
Full Scale IQ	98	105		

Experimental P2				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			15	12
Information	14	16		
Coding			10	12
Similarities	13	15		
Picture Arrangement			14	16
Arithmetic	12	10		
Block Design			10	8
Vocabulary	12	12		
Object Assembly			11	10
Comprehension	15	13		
(Symbol Search)			(12)	(11)
(Digit Span)	(13)	(11)		
Sum of Scaled Scores	66	66	60	58
Verbal IQ	120	120		
Performance IQ			115	112
Full Scale IQ	119	118		

Experimental P3				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			13	14
Information	8	9		
Coding			9	10
Similarities	12	12		
Picture Arrangement			10	11
Arithmetic	13	10		
Block Design			9	11
Vocabulary	13	13		
Object Assembly			11	12
Comprehension	15	14		
(Symbol Search)			(8)	(11)
(Digit Span)	(10)	(10)		
Sum of Scaled Scores	61	58	52	58
Verbal IQ	113	109		
Performance IQ			103	112
Full Scale IQ	109	111		

Experimental P8				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			7	7
Information	7	9		
Coding			9	14
Similarities	13	12		
Picture Arrangement			13	19
Arithmetic	10	8		
Block Design			6	7
Vocabulary	9	9		
Object Assembly			9	13
Comprehension	12	12		
(Symbol Search)			(12)	(18)
(Digit Span)	(12)	(12)		
Sum of Scaled Scores	51	50	54	60
Verbal IQ	101	100		
Performance IQ			91	115
Full Scale IQ	96	107		

Control P4				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			8	8
Information	11	12		
Coding			9	10
Similarities	12	9		
Picture Arrangement			11	16
Arithmetic	13	9		
Block Design			8	9
Vocabulary	10	7		
Object Assembly			10	10
Comprehension	14	11		
(Symbol Search)			(11)	(5)
(Digit Span)	(13)	(14)		
Sum of Scaled Scores	60	48	46	53
Verbal IQ	111	98		
Performance IQ	94	104	94	104
Full Scale IQ	104	101		

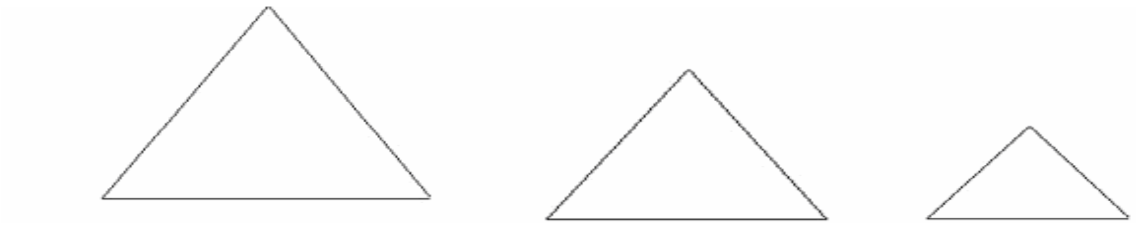
Control P6				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			10	13
Information	14	15		
Coding			14	14
Similarities	10	11		
Picture Arrangement			6	10
Arithmetic	10	11		
Block Design			14	8
Vocabulary	13	12		
Object Assembly			11	10
Comprehension	13	11		
(Symbol Search)			(12)	(13)
(Digit Span)	(10)	(10)		
Sum of Scaled Scores	60	60	55	55
Verbal IQ	111	111		
Performance IQ			107	107
Full Scale IQ	111	111		

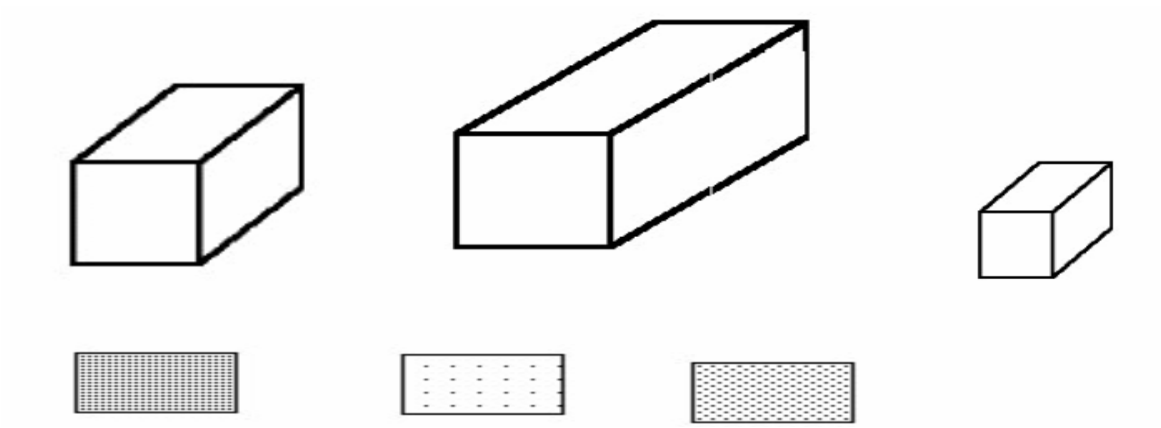
Control P7				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			11	11
Information	10	11		
Coding			13	10
Similarities	9	12		
Picture Arrangement			12	16
Arithmetic	10	13		
Block Design			9	10
Vocabulary	11	10		
Object Assembly			10	12
Comprehension	11	11		
(Symbol Search)			(10)	(11)
(Digit Span)	(10)	(10)		
Sum of Scaled Scores	51	57	55	59
Verbal IQ	101	108		
Performance IQ			107	113
Full Scale IQ	104	111		

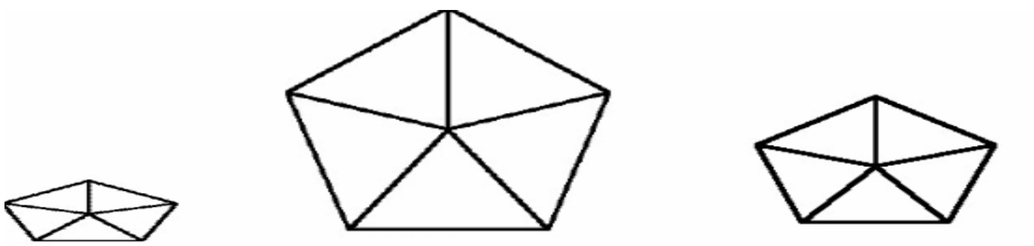
Control P9				
Subtests	Verbal Scaled Scores at baseline	Verbal Scaled Scores at follow-up	Performance Scaled Scores at baseline	Performance Scaled Scores at follow-up
Picture Completion			10	10
Information	15	15		
Coding			14	16
Similarities	9	9		
Picture Arrangement			10	7
Arithmetic	13	13		
Block Design			9	6
Vocabulary	12	13		
Object Assembly			9	9
Comprehension	10	11		
(Symbol Search)			(15)	(14)
(Digit Span)	(7)	(10)		
Sum of Scaled Scores	59	61	52	48
Verbal IQ	110	113		
Performance IQ			103	96
Full Scale IQ	107	106		

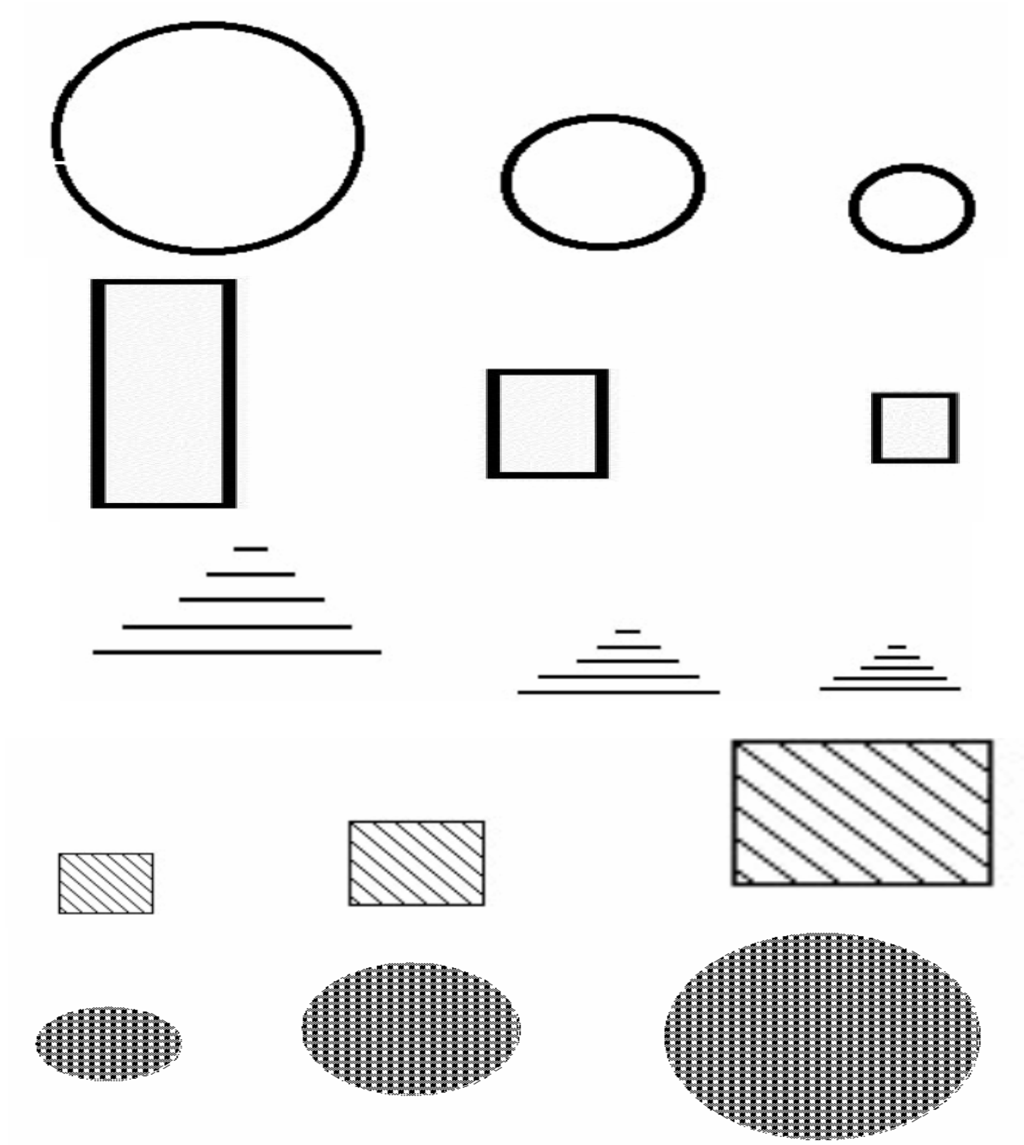
<u>Appendix 10</u>

Non-arbitrary stimulus sets employed during Same/Opposite Pre-training phases and Remedial phases (where applicable). Chapters 3-6.









<u>Appendix 11</u>

Same Arbitrary Nonsense Syllables employed in Chapter 3 as stimuli.

A1	B1	C1	D1	N1	N2
moq	niw	bup	ivr	xor	ziy
vew	zaq	qef	fik	dal	lus
xaf	ces	vup	biy	nor	maw
zog	lih	kuj	jek	haz	gox
fic	duv	seb	paq	yor	tiy
rup	wef	qag	moh	nij	buk
laz	xec	ven	buq	nie	mor
qat	wey	rui	tis	yod	pah
aej	sul	diz	fov	gan	heq
juw	kie	lor	zat	xey	cui

Appendix 12

Opposite Arbitrary Nonsense Syllables employed in Chapter 3 as stimuli.

A1	B1	C1	D1	N1	N2
guj	hik	loz	xac	veb	nim
qem	wir	tos	yud	pag	feh
sug	dif	fos	gah	hej	kua
liw	zor	xat	cey	vur	biq
niv	mab	qov	wem	eix	rop
tas	yed	puh	sif	doj	kal
vio	baf	nea	mus	qud	wif
eug	rih	toj	kaz	lex	zui
aqu	ewi	roe	orp	uts	aya
sed	gif	hoj	klu	zax	cev
vil	sup	peb	gev	pil	tod
cek	rac	hik	yeq	zeg	bov
wox	qiq	fil	sar	dow	muk
quk	faq	ruk	det	reh	xek
pey	dac	gov	yit	tad	voj
zal	zaj	huf	sog	kod	naq
yiy	fal	lim	ruh	qej	sur
haf	fez	cep	cak	fec	sed
pef	moj	jac	qoz	zep	jef
nix	pox	vak	kow	ruq	pir
jak	yif	miq	lum	cos	mur
noy	qok	pek	woq	mut	goq

Appendix 13

List of all artists and titles of dance music songs employed. Chapters 3-6.

1. Jennifer Lopez - A'int it funny
2. Stonebridge - Put 'em high
3. Moloko - Sing it back
4. Deepest Blue- Deepest blue
5. Madonna - Hung up
6. Missy Elliot - Get ur freak on
7. Jennifer Lopez- Get right
8. Madonna - Ray of light
9. Christina Aguilera - Genie in a bottle
10. Faithless - Insomnia
11. Kylie Minogue - In your eyes
12. Maria Carey - It's like that
13. Mojo - Lady
14. Justin Timberlake - Love you
15. Kylie MInogue - Can't get you out of my head
16. Sugababes - Push the button
17. Yeah - Usher

Appendix 14

Instructions for Same and Opposite Pre-Training. Chapter 3-6.

In a moment some objects will appear on this screen.
Your task is to look at the object at the top of the screen, then look at the object in the middle of the screen and then choose one of the three objects at the bottom.
The computer will tell you whether or not your choice is correct.
You should try to get as many answers correct as possible by paying close attention to whether the computer tells you your choices are right or wrong.
The harder you work to figure out how to make correct choices the faster this stage will finish.
The computer will tell you when you should take a break.
If you have any questions, please ask them now.
When you are ready, please click on the mouse button.

Appendix 15

Instructions for Same and Opposite Pre-Training testing. Chapter 3-6.

In a moment some more objects will appear on this screen.
Like before, your task is to look at the object at the top of the screen, then look at the object in the middle of the screen and then choose one of the three objects at the bottom of the screen by "clicking on it" using the mouse.
This time the computer will NOT tell you whether or not your choice is correct.
However, you still need to get as many answers correct as possible. It will help if you try to remember what you learned up to this point.
The harder you work to figure out how to make correct choices, the faster this stage will finish. The computer will tell you when to take a break.
If you have any questions, please ask them now.
When you are ready, please click on the mouse button.

Appendix 16

Instructions for Same and Opposite Training. Chapters 3-6.

In a moment some items will appear at the top of this screen.
You can treat these items as if they form a statement.
You can read this statement from left to right as you would read any other statement.
Then the words YES and NO will appear at the bottom of the screen. If you agree with the statement, click on YES with the mouse button. If you do not agree with the statement, click on NO with the mouse button.
The computer will tell you if your choice was right or wrong.
You should try to get as many correct choices as possible.
If you have any questions, please ask them now.
When you are ready to begin, please click on the mouse button.

Appendix 17

List of all verbal reinforcement statements employed during training and MET testing phases in Chapters 3, 4, 5 and 6.

1. Good effort, but you need to get them all right in order to move on to the next part. Make sure you concentrate on whether the computer tells you you're right or wrong. Let's give it another go.
2. I know you're working hard, but I need you to try your best to focus on whether the computer tells you you're right or wrong. Are you ready to try again?
3. Ok, you're getting there, but I'm not sure you're concentrating enough on what the computer is telling you. Will we go again?
4. Nice try, but are you sure you're listening to the computer? It will tell you whether your choices are right or wrong. Try again.
5. Good effort, but you must try to remember whether the computer told you were right or wrong for each one. Let's have another go.
6. Well done for trying, but you need to pay closer attention to what the computer is telling you. Ready?? Let's go again.
7. Not quite there yet. Are you sure you're paying attention to what the computer is telling you?? Let's give this another go.
8. Ah no! Not quite there yet. Make sure you're listening to whether the computer tells you you're right or wrong.
9. Ok, good effort, but you need to make sure you're paying attention to whether the computer tells you you're right or wrong. Let's go again.
10. Nice try. Just make sure you're paying close attention to what the computer is telling you. Let's give it another shot.

Appendix 18a

Instructions for Same and Opposite (non-MET) Testing. Chapters 3-6.

In this stage, some more items will appear at the top of the screen. Again you can read these items as if they form a statement. You may read this statement from left to right as you would read any other statement. Then the words YES and NO will appear at the bottom of the screen. As before, if you agree with the statement, you should click on YES with the mouse button. If you do not agree with the statement, you should click on NO with the mouse button. This time the computer will NOT tell you whether your choice was right or wrong. This stage will be a bit harder so you need to think carefully before you make your choice. As always, you should try to get as many correct choices as possible.

If you have any questions, please ask them now.

When you are ready to begin, please click on the mouse button.

Appendix 18b

Instructions for Same and Opposite MET Testing. Chapters 3-6.

In a moment some items will appear at the top of this screen.
You can treat these items as if they form a statement.
You can read this statement from left to right as you would read any other statement.
Then the words YES and NO will appear at the bottom of the screen. If you agree with the statement, click on YES with the mouse button. If you do not agree with the statement, click on NO with the mouse button.
The computer will tell you if your choice was right or wrong.
You should try to get as many correct choices as possible.
If you have any questions, please ask them now.
When you are ready to begin, please click on the mouse button.

Appendix 19

Same arbitrary nonsense syllables employed as stimuli in Chapters 4, 5 and 6.

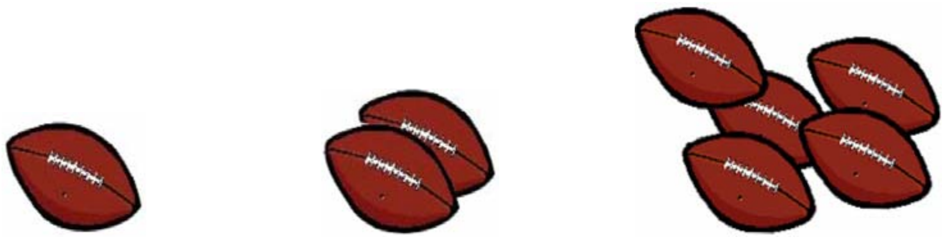
A	B	C	D	N1	N2
Mup	han	xiz	tun	jij	caq
rej	kam	fos	haq	jaf	kap
cav	req	jox	tup	yiz	juj
piv	muq	vuw	cer	wub	ney
zot	noy	riq	poq	yiq	bom
pom	xon	vur	rel	yaj	wew
rin	woc	xag	fes	taf	fuw
cez	yom	zer	yix	qog	zuv
fap	yip	bim	joz	kor	gah
zuz	quv	dah	ciw	dib	beb
tes	gub	dil	qit	cet	nof
cor	jov	yok	nin	qaf	jud
pex	kiw	taf	kol	zey	bip
kux	xuz	vuy	vev	vaz	zus
mul	rup	xos	vac	pul	tam
teq	zet	gef	huw	zuw	vaf
hap	wol	bal	qah	bah	sog
xey	gam	luh	tak	res	vut
fex	xab	tel	wem	toj	qac
qoz	zew	vix	qox	wok	ruw

<u>Appendix 20</u>

Opposite arbitrary nonsense syllables employed as stimuli in Chapters 4, 5 and 6

A	B	C	D	N1	N2
bik	yuq	tah	qir	yim	tiq
kir	mun	cas	mus	noz	rol
luc	wer	xoq	nuv	gud	gak
fev	luj	yal	xux	zop	sev
nux	yoh	vec	bof	zuc	huy
siz	loy	puq	saz	yof	ciy
lim	mir	woj	quk	tol	wud
wuq	fec	nad	hos	gug	yan
miw	hoq	noc	roj	piq	suy
kon	hek	goz	wif	gof	dex
xip	jod	rak	hih	seq	tox
gir	ruj	fov	liy	kek	maj
moh	hik	juw	nes	vim	gat
neq	xas	ler	qas	yaz	fip
yuf	qev	dow	gef	vos	kic
qeq	xuq	zur	tih	vem	tes
soq	xav	lib	zuk	vek	weh
voq	div	fim	gub	buc	vov
lor	yac	xes	pex	yel	waw
yex	nas	jub	yeg	lep	xaq
kiz	xim	jej	luz	doy	yej
kes	zup	fer	mal	xul	tev
dof	keq	hol	qor	xan	jat
yax	kif	jit	kih	wih	jij
qom	tud	mij	zan	puz	foh
fos	sov	nen	cij	qot	qaq
guw	quv	rud	von	mok	baz
bot	quy	kog	puh	hir	vap
mew	baz	cor	ver	cuk	xuj
xug	wew	buh	goc	pij	qen
zum	qop	jiw	zas	yuh	jeg
sew	pag	rah	pid	liw	koj
var	feh	qes	xik	teh	pel
haj	xun	jiv	gox	joq	niy
hud	naq	qex	vox	kol	nav
toz	yev	geh	niw	lof	wab
zav	rul	bax	Juy	giw	cin
tuj	jix	fik	Nux	paj	wud
suy	goz	dex	Xip	hih	seq
gof	gir	ruj	Fov	liy	kek

<u>Appendix 21</u>
More-Than/Less-Than pre-training and remedial non-arbitrary stimulus sets.
Experiments 6, 7, 8, 10b and 11.



Appendix 22

More-Than/Less-Than nonsense syllables, Experiments, 6, 7, 8 and 11.

A	B	C	D	N1	N2
zet	req	foz	nug	cid	pem
bem	bor	het	yoj	geg	jaf
ziv	yaf	xup	yuh	qav	tuj
tuh	cav	zer	yiz	piv	vuw
nuw	hik	fet	puy	neb	ruh
cer	wub	zot	pok	yic	xon
rek	mij	cak	lig	zur	xeq
caq	tup	vur	yaj	rin	woq
zij	fal	yig	vax	ruf	yog
xag	fes	fuw	yom	qog	zuv
hak	yib	woc	fer	lib	meq
yix	fap	bim	kor	zuz	quv
duw	raj	xur	jey	zux	fos
ciw	beb	gub	qit	naf	jov
qox	juw	yac	ceq	waf	bof
yok	qaf	jud	kiw	kol	zey
wox	gew	qos	gic	vav	zup
kux	xuz	vev	haw	zus	rup
zos	dih	tur	qov	baz	yiy
vuy	zav	xos	teq	zet	gef

Appendix 23a

Instructions for More/Less Pre-Training, Experiments 6, 7, 8 and 11.

In a moment some objects will appear on this screen.
Your task is to look at the object at the top of the screen, then look at the object in the middle of the screen and then choose one of the three objects at the bottom.
The computer will tell you whether or not your choice is correct.
You should try to get as many answers correct as possible by paying close attention to whether the computer tells you your choices are right or wrong.
The harder you work to figure out how to make correct choices the faster this stage will finish.
The computer will tell you when you should take a break.
If you have any questions, please ask them now.
When you are ready, please click on the mouse button.

Appendix 23b

More-Than/Less-Than pre-training test instructions, Experiments 6, 7, 8, 11.

In a moment some more objects will appear on this screen.
Like before, your task is to look at the object at the top of the screen, then look at the object in the middle of the screen and then choose one of the three objects at the bottom of the screen by “clicking on it” using the mouse.
This time the computer will NOT tell you whether or not your choice is correct.
However, you still need to get as many answers correct as possible. It will help if you try to remember what you learned up to this point.

Appendix 24

More-Than/Less-Than relational training instructions, Experiments 6, 7, 8, 11.

In a moment some items will appear at the top of this screen.
You can treat these items as if they form a statement.
You can read this statement from left to right as you would read any other statement.
Then the words YES and NO will appear at the bottom of the screen. If you agree with the statement, click on YES with the mouse button. If you do not agree with the statement, click on NO with the mouse button.
The computer will tell you if your choice was right or wrong.
You should try to get as many correct choices as possible.
If you have any questions, please ask them now.
When you are ready to begin, please click on the mouse button.

Appendix 25a

More-Than/Less-Than relational testing instructions (non-MET), Experiments 6, 7, 8, 11.

In this stage, some more items will appear at the top of the screen. Again you can read these items as if they form a statement. You may read this statement from left to right as you would read any other statement. Then the words YES and NO will appear at the bottom of the screen. As before, if you agree with the statement, you should click on YES with the mouse button. If you do not agree with the statement, you should click on NO with the mouse button. This time the computer will NOT tell you whether your choice was right or wrong. This stage will be a bit harder so you need to think carefully before you make your choice. As always, you should try to get as many correct choices as possible.
If you have any questions, please ask them now.
When you are ready to begin, please click on the mouse button.

Appendix 25b

More-Than/Less-Than relational testing instructions (MET), Experiments 6, 7, 8, 11.

In a moment some items will appear at the top of this screen.
You can treat these items as if they form a statement.
You can read this statement from left to right as you would read any other statement.
Then the words YES and NO will appear at the bottom of the screen. If you agree with the statement, click on YES with the mouse button. If you do not agree with the statement, click on NO with the mouse button.
The computer will tell you if your choice was right or wrong.
You should try to get as many correct choices as possible.
If you have any questions, please ask them now.
When you are ready to begin, please click on the mouse button.

Appendix 26

Experimental participant's (P1, P2, P3, P8) individual subtest scores contributing to Full Scale IQ, Verbal IQ and Performance IQ at Time 3.

Experimental Participant, P1.		
Subtests	**Verbal Scaled Scores**	**Performance Scaled Scores**
Picture Completion		**13**
Information	**16**	
Coding		**12**
Similarities	**15**	
Picture Arrangement		**19**
Arithmetic	**12**	
Block Design		**14**
Vocabulary	**11**	
Object Assembly		**13**
Comprehension	**14**	
(Symbol Search)		**(16)**
(Digit Span)	**(9)**	
Sum of Scaled Scores	**68**	**71**
Verbal IQ	**122**	
Performance IQ	**132**	
Full Scale IQ	**131**	

Experimental Participant, P2.		
Subtests	**Verbal Scaled Scores**	**Performance Scaled Scores**
Picture Completion		**17**
Information	**15**	
Coding		**12**
Similarities	**18**	
Picture Arrangement		**18**
Arithmetic	**17**	
Block Design		**12**
Vocabulary	**13**	
Object Assembly		**12**
Comprehension	**19**	
(Symbol Search)		**(14)**
(Digit Span)	**(17)**	
Sum of Scaled Scores	**82**	**71**
Verbal IQ	**139**	
Performance IQ	**132**	
Full Scale IQ	**137**	

Experimental Participant, P3.		
Subtests	**Verbal Scaled Scores**	**Performance Scaled Scores**
Picture Completion		**16**
Information	**15**	
Coding		**11**
Similarities	**15**	
Picture Arrangement		**19**
Arithmetic	**16**	
Block Design		**10**
Vocabulary	**14**	
Object Assembly		**13**
Comprehension	**19**	
(Symbol Search)		**(11)**
(Digit Span)	**(9)**	
Sum of Scaled Scores	**79**	**69**
Verbal IQ	**136**	
Performance IQ	**130**	
Full Scale IQ	**135**	

Experimental Participant, P8.		
Subtests	**Verbal Scaled Scores**	**Performance Scaled Scores**
Picture Completion		**11**
Information	**12**	
Coding		**13**
Similarities	**11**	
Picture Arrangement		**18**
Arithmetic	**12**	
Block Design		**9**
Vocabulary	**9**	
Object Assembly		**12**
Comprehension	**16**	
(Symbol Search)		**(13)**
(Digit Span)	**(11)**	
Sum of Scaled Scores	**60**	**76**
Verbal IQ	**111**	
Performance IQ	**137**	
Full Scale IQ	**128**	

Appendix 27

Control participant's (P4, P6, P7, P9) individual subtest scores contributing to Full Scale IQ, Verbal IQ and Performance IQ at Time 3.

Control Participant, P4.

Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		**6**
Information	**14**	
Coding		**14**
Similarities	**9**	
Picture Arrangement		**12**
Arithmetic	**9**	
Block Design		**8**
Vocabulary	**7**	
Object Assembly		**10**
Comprehension	**10**	
(Symbol Search)		**(11)**
(Digit Span)	**(11)**	
Sum of Scaled Scores	**49**	**50**
Verbal IQ	**99**	
Performance IQ	**99**	
Full Scale IQ	**99**	

Control Participant, P6.

Subtests	Verbal Scaled Scores	Performance Scaled Scores
Picture Completion		**8**
Information	**16**	
Coding		**15**
Similarities	**13**	
Picture Arrangement		**9**
Arithmetic	**13**	
Block Design		**8**
Vocabulary	**12**	
Object Assembly		**10**
Comprehension	**8**	
(Symbol Search)		**(9)**
(Digit Span)	**(14)**	
Sum of Scaled Scores	**62**	**50**
Verbal IQ	**115**	
Performance IQ	**99**	
Full Scale IQ	**108**	

Control Participant, P7.		
Subtests	**Verbal Scaled Scores**	**Performance Scaled Scores**
Picture Completion		**9**
Information	**13**	
Coding		**10**
Similarities	**11**	
Picture Arrangement		**15**
Arithmetic	**9**	
Block Design		**12**
Vocabulary	**12**	
Object Assembly		**10**
Comprehension	**8**	
(Symbol Search)		**(11)**
(Digit Span)	**(9)**	
Sum of Scaled Scores	**53**	**56**
Verbal IQ	**103**	
Performance IQ	**109**	
Full Scale IQ	**106**	

Control Participant, P9.		
Subtests	**Verbal Scaled Scores**	**Performance Scaled Scores**
Picture Completion		**8**
Information	**15**	
Coding		**12**
Similarities	**12**	
Picture Arrangement		**7**
Arithmetic	**14**	
Block Design		**7**
Vocabulary	**13**	
Object Assembly		**8**
Comprehension	**10**	
(Symbol Search)		**(6)**
(Digit Span)	**(8)**	
Sum of Scaled Scores	**64**	**42**
Verbal IQ	**117**	
Performance IQ	**88**	
Full Scale IQ	**104**	

Appendix 28

RAI nonsense syllables employed as stimuli across Experiments 10 and 11.

Same			More			Less			Opposite		
A	B	C	A	B	C	A	B	C	A	B	C
cav	rin	gub	dex	yuf	weh	xim	puz	xan	teq	ziw	yal
req	wom	nof	xip	zur	vov	jej	qom	kif	zet	nuv	zop
tup	taf	jov	jod	qev	fim	luz	kih	jit	gef	vut	loy
yiz	yom	zey	hih	vem	jub	zup	yej	fos	vaf	fex	saz
piv	zuv	bip	ruj	vek	yeg	fep	xul	cij	boh	qoz	yof
wub	fap	kux	fov	zuk	pex	dof	jat	qot	gam	ruw	tol
zot	bim	vev	liy	yex	kiz	keq	foh	guw	luh	kir	wuq
noy	ciw	zus	kek	fip	voq	yax	sov	rud	xab	yim	gug
vur	dib	rup	qas	ler	tih	sij	nen	mew	wem	cas	miw
yaj	beb	tam	vos	xas	gef	tud	zan	cuk	toj	tak	gof

Appendix 29a

RAI instructions-no time limit

In a moment some statements will appear at the top of this screen. A question will also be presented underneath these statements. Then the words YES and NO will appear at the bottom of the screen. You should read the two statements carefully and then answer the questions by clicking on YES or NO using the mouse.

It is important that you get as many answers correct as possible.
If you have any questions, please ask them now.
When you are ready to begin, please click on the mouse button.

Appendix 29b

RAI instructions-with time limit

In a moment some more statements will appear at the top of this screen. A question will also be presented underneath these statements. Then the words YES and NO will appear at the bottom of the screen. You should read the two statements carefully and then answer the questions by clicking on YES or NO using the mouse.

However, this time there will be a time limit for each question. You only have a few seconds to answer the question by clicking on YES or NO using the mouse. If you do not make a choice, an incorrect answer will be recorded by the computer and the next statements and question will be presented.

It is important that you get as many answers correct as possible.
If you have any questions, please ask them now.

Appendix 30

Description of the WISC-IV^{UK} subtests used in Experiment 10b and 11, adapted from the Manual for the WISC-IV^{UK} (2004).

Subtest	Subtest Description
Similarities	This subtest is designed to measure verbal reasoning and concept formation. It also involves auditory comprehension, memory, distinction between nonessential and essential features and verbal expression.
Vocabulary	This subtest is designed to measure word knowledge, learning ability, long-term memory and the degree of language development.
Comprehension	This subtest is designed to measure verbal reasoning and conceptualization, verbal comprehension and expression, the ability to evaluate and use past experience, and the ability to demonstrate practical information.
Block Design	This subtest is designed to measure the ability to analyze and synthesize abstract visual stimuli.
Picture Concepts	This subtest is designed to measure abstract, categorical reasoning ability.
Matrix Reasoning	This subtest is designed to provide a reliable measure of visual information processing and abstract reasoning skills.
Digit Span	This subtest is designed to measure auditory short-term memory, sequencing skills, attention, and concentration.
Letter-Number Sequencing	This subtest involves sequencing, mental manipulation, attention, short-term auditory memory, visuospatial imaging, and processing speed.
Coding	This subtest measures processing speed, short-term memory, learning ability, visual perception, visual-motor co-ordination, visual scanning ability, cognitive flexibility, attention and motivation.
Symbol Search	This subtest measures processing speed. It also involves the use of short-term visual memory, visual-motor coordination, cognitive flexibility, visual discrimination and concentration.

Appendix 31

Description of the WISC-IV^{UK} indices and the core subtests which contribute to the indices adapted from the Technical and Interpretive Manual for the WISC-IV^{UK} (2004).

Indices	Subtests which contribute to the indices
The Verbal Comprehension Index (VCI) is composed of subtests measuring verbal abilities utilizing reasoning, comprehension and conceptualization.	Similarities, Vocabulary, Comprehension,
The Perceptual Reasoning Index (PRI) is composed of subtests measuring perceptual reasoning and organization.	Block Design, Picture Concepts, Matrix Reasoning
The Working Memory Index (WMI) is composed of subtests measuring attention, concentration and working memory.	Digit Span, Letter-Number Sequencing
The Processing Speed Index (PSI) is composed of subtests measuring the speed of mental and graphomotor processing.	Coding, Symbol Search

Appendix 32

Mean RAI rises for each participant for each relation in Experiment 11.

Participant ID	Same mean RAI rises	More-Than mean RAI rises	Less-Than mean RAI rises	Opposite mean RAI rises
P74	4.67	0	9	5
P75	12.67	8	1	3.67
P76	8.67	10	1	5.66
P77	12.34	1.34	1.66	1.34
P78	11.33	1	1.34	4
P79	10.33	3.33	7.33	10
P80	10	9.34	5.67	8.34
P81	10.67	17	13	8.33
Average mean rises	**10.09**	**6.25**	**5**	**5.79**