

**Can Teaching Flexible Relational Responding Result in Improved
Ability Scores in Children with Diagnosed Autism?**



NUI MAYNOOTH

Ollscoil na hÉireann Má Nuad

Submitted by: Keith Lyons B.A. (Hons.)

Supervisor: Dr Carol Murphy BCBA-D

Thesis Presented in Part Fulfilment of the Requirements for the Doctorate in
Psychological Science (Behaviour Analysis and Therapy), Department of
Psychology, Faculty of Science and Engineering, National University of Ireland,
Maynooth

March 31st 2013

Table of Contents

LIST OF FIGURES	VI
ACKNOWLEDGMENTS	X
ABSTRACT	XI
-CHAPTER 1-.....	1
INTRODUCTION.....	2
<i>Relational Flexibility and Intelligent Behaviour.....</i>	<i>12</i>
<i>The Current Study.....</i>	<i>17</i>
-CHAPTER 2-.....	20
METHOD	21
<i>Participants and recruitment</i>	<i>21</i>
<i>Ethical Considerations.....</i>	<i>23</i>
<i>Settings and Materials.....</i>	<i>27</i>
<i>Design</i>	<i>30</i>
<i>Inter Observer Agreement.....</i>	<i>31</i>
<i>Procedure.....</i>	<i>31</i>
<i>Decision protocol.....</i>	<i>40</i>
<i>Prompt Levels and Specific Accommodations</i>	<i>40</i>
RESULTS	43
<i>IQ Assessment Data</i>	<i>43</i>
<i>Non-arbitrary Relations (Like/Different).....</i>	<i>44</i>
<i>Mean accuracies and response times for T-IRAP and Table Top</i>	<i>57</i>
<i>Summary.....</i>	<i>59</i>

-CHAPTER 3-	62
INTRODUCTION.....	63
METHOD	66
<i>Participants</i>	66
<i>Settings and Materials</i>	67
<i>Design</i>	67
<i>Inter Observer Agreement</i>	67
<i>Procedure</i>	68
<i>Specific Accommodations</i>	71
RESULTS	72
<i>Accuracy and Speed (Like-Different Arbitrary Relations)</i>	72
SUMMARY	74
-CHAPTER 4-	76
INTRODUCTION.....	77
METHOD	80
<i>Participants and Recruitment</i>	80
<i>Settings and Materials</i>	80
<i>Design</i>	81
<i>Inter Observer Agreement</i>	81
<i>Procedure</i>	81
<i>Prompt Levels and Specific Alterations</i>	88
RESULTS	89
<i>Accuracy and Speed (Like-Different Non-arbitrary Double Reversals)</i>	89
<i>Mean Accuracies and Response Times for T-IRAP and Table-top</i>	99
<i>Post-test Verbal Ability Measures</i>	101

<i>Supplementary Analysis</i>	103
<i>Summary</i>	104
-CHAPTER 5-	106
GENERAL DISCUSSION	107
<i>Summary of Findings</i>	108
<i>Table-Top and T-IRAP Procedures</i>	113
<i>Flexibility in Relational Responding</i>	115
<i>Effectiveness of Prompting Procedures</i>	117
<i>Accounting for Failure to Generalise a Previously Learned Response</i>	118
<i>Limitations and Future Research</i>	120
<i>Conclusion</i>	122
REFERENCES	124
TABLE OF APPENDICES	135
APPENDIX 1: INFORMED CONSENT FORM FOR PARTICIPANTION IN DOCTORAL RESEARCH.....	136
APPENDIX 2: PARENT INFORMATION SHEET.....	138
APPENDIX 3: CONTINUED CONSENT FORM.....	141
APPENDIX 4: CORRESPONDENCE WITH SIGMA ASSESSMENT SYSTEMS	142
APPENDIX 5: APA GUIDELINES FOR STUDENT USE OF PSYCHOMETRIC TEST.....	143
APPENDIX 6: SAMPLES OF STIMULI USED	145

List of Tables

Table 1. <i>Standard (S) and Raw (R) Scores for Each Participant on the K-BIT and PPVT-IV</i>	44
Table 2. <i>Mean Durations for all Sessions to Date by Participant and Stimulus Set</i>	58
Table 3. <i>Standard (S) and Raw(R) Scores for Each Participant on the K-BIT and PPVT-IV at Pre and Post Test</i>	100
Table 4. <i>Mean Durations for all Sessions to Date by Participant and Stimulus Set in Seconds</i>	101

List of Figures

- FIGURE 1.* EXAMPLES OF THE FOUR TRIAL TYPES USED IN THE SQUARE/TRIANGLE STIMULUS SET T-IRAP TEACHING PROCEDURE DURING “LIKE/DIFFERENT” RELATIONAL FLEXIBILITY TRAINING. ARROWS SIGNAL THE CORRECT RESPONSE; THESE RESPONSES WERE CONSIDERED CONSISTENT WITH PRE-ESTABLISHED PATTERNS OF RESPONDING IN THE GENERAL POPULATION. THE ARROWS DID NOT APPEAR ON THE SCREEN DURING INSTRUCTION.38
- FIGURE 2.* EXAMPLES OF THE FOUR TRIAL TYPES USED IN THE ANIMAL/FOOD STIMULUS SET T-IRAP TEACHING PROCEDURE DURING “LIKE/DIFFERENT” RELATIONAL FLEXIBILITY TRAINING. ARROWS SIGNAL THE CORRECT RESPONSE; THESE RESPONSES WERE CONSIDERED CONSISTENT WITH PRE-ESTABLISHED PATTERNS OF RESPONDING IN THE GENERAL POPULATION. THE ARROWS DID NOT APPEAR ON THE SCREEN DURING INSTRUCTION.....39
- FIGURE 3.* TABLE-TOP AND T-IRAP DATA FOR ANDREW (TOP) AND EVAN (BOTTOM) SHOWING ACCURACY AND SPEED OF RESPONDING (DURATION) DATA FOR LIKE/DIFFERENT NON-ARBITRARY RELATIONAL RESPONDING WITH STIMULI THAT WERE VARIED ALONG ONE PHYSICAL DIMENSION (I.E., SHAPE: SQUARE/TRIANGLE)46
- FIGURE 4.* TABLE-TOP AND T-IRAP DATA FOR STEPHEN (TOP) AND ANN (BOTTOM) SHOWING ACCURACY AND SPEED OF RESPONDING (DURATION) DATA FOR LIKE/DIFFERENT NON-ARBITRARY RELATIONAL RESPONDING WITH STIMULI THAT WERE VARIED ALONG ONE PHYSICAL DIMENSION (I.E., SHAPE: SQUARE/TRIANGLE)50
- FIGURE 5.* TABLE-TOP AND T-IRAP DATA FOR KEVIN SHOWING ACCURACY AND SPEED OF RESPONDING (DURATION) DATA FOR LIKE/DIFFERENT NON-ARBITRARY

RELATIONAL RESPONDING WITH STIMULI THAT WERE VARIED ALONG ONE
 PHYSICAL DIMENSION (I.E., SHAPE: SQUARE/TRIANGLE).....52

FIGURE 6. TABLE-TOP AND T-IRAP DATA FOR STEPHEN (TOP) AND ANN (BOTTOM)
 SHOWING ACCURACY AND SPEED OF RESPONDING (DURATION) DATA FOR
 LIKE/DIFFERENT NON-ARBITRARY RELATIONAL RESPONDING WITH STIMULI THAT
 WERE VARIED IN RELATION TO CLASS MEMBERSHIP (I.E., CLASS: ANIMAL/FOOD)
55

FIGURE 7. TABLE-TOP AND T-IRAP DATA FOR KEVIN SHOWING ACCURACY AND SPEED
 OF RESPONDING (DURATION) DATA FOR LIKE/DIFFERENT NON-ARBITRARY
 RELATIONAL RESPONDING WITH STIMULI THAT VARIED IN RELATION TO CLASS
 MEMBERSHIP (I.E., CLASS: ANIMAL/FOOD).....57

FIGURE 8. EXAMPLES OF THE FOUR TRIAL TYPES USED IN THE ARBITRARY STIMULUS
 SET T-IRAP TEACHING PROCEDURE DURING “LIKE/DIFFERENT” RELATIONAL
 FLEXIBILITY TRAINING. ARROWS SIGNAL THE CORRECT RESPONSE; THESE
 RESPONSES WERE CONSIDERED CONSISTENT WITH PRE-ESTABLISHED PATTERNS OF
 RESPONDING IN THE GENERAL POPULATION. THE ARROWS DID NOT APPEAR ON
 THE SCREEN DURING INSTRUCTION.....70

FIGURE 9. TABLE-TOP AND T-IRAP DATA FOR KEVIN SHOWING ACCURACY AND SPEED
 OF RESPONDING (DURATION) DATA FOR LIKE/DIFFERENT ARBITRARY
 RELATIONAL RESPONDING WITH STIMULI THAT DID NOT VARY IN RELATION TO
 ANY FORMAL PROPERTY BUT ARBITRARILY.....73

FIGURE 10. EXAMPLES OF THE FOUR TRIAL TYPES USED IN THE SQUARE/TRIANGLE
 STIMULUS SET T-IRAP TEACHING PROCEDURE DURING “LIKE/DIFFERENT”
 INCONSISTENT COMPLEX RELATIONAL FLEXIBILITY TRAINING. ARROWS SIGNAL
 THE CORRECT RESPONSE; THESE RESPONSES WERE CONSIDERED INCONSISTENT

WITH PRE-ESTABLISHED PATTERNS OF RESPONDING IN THE GENERAL POPULATION.

THE ARROWS DID NOT APPEAR ON THE SCREEN DURING INSTRUCTION.....85

FIGURE 11. EXAMPLES OF THE FOUR TRIAL TYPES USED IN THE ANIMAL/FOOD
STIMULUS SET T-IRAP TEACHING PROCEDURE DURING “LIKE/DIFFERENT”
INCONSISTENT COMPLEX RELATIONAL FLEXIBILITY TRAINING. ARROWS SIGNAL
THE CORRECT RESPONSE; THESE RESPONSES WERE CONSIDERED INCONSISTENT
WITH PRE-ESTABLISHED PATTERNS OF RESPONDING IN THE GENERAL POPULATION.
THE ARROWS DID NOT APPEAR ON THE SCREEN DURING INSTRUCTION.....86

FIGURE 12. EXAMPLES OF THE FOUR TRIAL TYPES USED IN THE ARBITRARY STIMULUS
SET T-IRAP TEACHING PROCEDURE DURING “LIKE/DIFFERENT” INCONSISTENT
COMPLEX RELATIONAL FLEXIBILITY TRAINING. ARROWS SIGNAL THE CORRECT
RESPONSE; THESE RESPONSES WERE CONSIDERED INCONSISTENT WITH PRE-
ESTABLISHED PATTERNS OF RESPONDING IN THE GENERAL POPULATION. THE
ARROWS DID NOT APPEAR ON THE SCREEN DURING INSTRUCTION.87

FIGURE 13. TABLE-TOP AND T-IRAP DATA FOR ANDREW (TOP) AND EVAN (BOTTOM)
SHOWING ACCURACY AND SPEED OF RESPONDING (DURATION) DATA FOR A
DOUBLE REVERSAL LIKE/DIFFERENT NON-ARBITRARY INCONSISTENT AND
CONSISTENT RELATIONS (I.E., SHAPE: SQUARE/TRIANGLE), INCLUDING
INCONSISTENT (INCON.) AND CONSISTENT (CON.) PROBES91

FIGURE 14. TABLE-TOP AND T-IRAP DATA FOR STEVEN (TOP) AND ANN (BOTTOM)
SHOWING ACCURACY AND SPEED OF RESPONDING (DURATION) DATA FOR A
DOUBLE REVERSAL LIKE/DIFFERENT NON-ARBITRARY INCONSISTENT AND
CONSISTENT RELATIONS (I.E., CLASS: ANIMAL/FOOD), INCLUDING INCONSISTENT
(INCON.) AND CONSISTENT (CON.) PROBES.....94

FIGURE 15. TABLE-TOP AND T-IRAP DATA FOR KEVIN SHOWING ACCURACY AND
SPEED OF RESPONDING (DURATION) DATA FOR A DOUBLE REVERSAL
LIKE/DIFFERENT ARBITRARY INCONSISTENT AND CONSISTENT RELATIONS,
INCLUDING INCONSISTENT (INCON.) AND CONSISTENT (CON.) PROBES.....98

Acknowledgments

There are several people to whom I owe a debt of gratitude, and cannot thank enough for their patients, advice and time

All those who took part in this study, my participants, who gave me their time a tolerance through many long days and difficult tasks.

My wonderful partner Linda, who supported me in every way possible over the course of the last three years, offering well need distractions and words of encouragement as they were needed. I owe my completion of this thesis to her.

My fantastic mother who's words of advice and gentle encouragement helped me over hurdles, especially in the difficult moments.

My fellow doctoral students, whose support, advice and senses of humour made things all the easier.

Finally, my supervisor, Dr. Carol Murphy for her indispensable advice, unwavering dedication and all her time spent answering my countless emails at all times of the day and night. I you a huge debt of gratitude and many many thanks.

Abstract

A computerised teaching tool known as the T-IRAP has been previously shown to be an effective tool for teaching relational responding skills to children with autism, using a multiple baseline design to demonstrate greater speed and accuracy with the T-IRAP compared to Table-top procedures (Kilroe, Murphy, & Barnes-Holmes, 2011). The current study extends this by using an alternating treatments design which is not subject to problematic sequence effects that may result with multiple baseline design. Cassidy, Roche, & Hayes, (2011) found that a teaching complex forms of relational responding resulted in significant gains in cognitive ability as measured by IQ tests. The current study builds upon this by examining the IQ and verbal abilities of children with autism at pre and post test to determine if relational training will result in similar increases in IQ.

Pre and post intervention measures of IQ and verbal ability were taken to assess if relational training generates any gains on these measures. An alternating treatments design compared a Table Top (TT) procedure with a computerised procedure (T-IRAP) to teach the same sets of relational stimuli. Study 1 successfully taught five children flexible relational responding with simple non-arbitrary stimuli. Study 2 successfully taught flexible relational responding with arbitrary stimuli to one participant. Finally, Study 3 taught complex relational flexibility with double contingency reversals to five participants with diagnosed autism.

All participants demonstrated some degree of relational flexibility, with a range of competencies observed. T-IRAP was found to be more efficient and effective in terms of speed and accuracy as compared to TT teaching across the research program as a whole. No statistically significant differences were found in ability scores from pre to post-test assessment, some individual gains were observed however.

-Chapter 1-

Introduction

Basic experimental research in behaviour analysis and the discovery of fundamental principles (Skinner, 1938; 1957) resulted in the development of an applied science of human behaviour. Decades of effective documented scientific interventions to promote adaptive behaviour and replace challenging behaviour with populations with intellectual disabilities resulted from the investigation of these principles (e.g. Baer, Wolf, & Risley, 1968; Lovaas, 1987; Lovaas, Berberich, Perloff, & Schaeffer, 1966; Azrin & Foxx, 1971). In addition to successfully addressing problem behaviours with individuals with typical and atypical development, applied behaviour analysis (ABA) has been very successful with a wide range of human problems such as emotional disturbance (Matson & Coe, 1992), promoting health behaviours (DeLuca & Holborn, 1992), and promoting participation in job skills training in alcoholics and the homeless (Koffarnus, et al., 2013). Most notably it has been shown to be particularly effective with the treatment of those with development disabilities (Cohen, Amerine-Dickens, & Smith, 2006; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005). It is now widely accepted that ABA language programmes using direct reinforcement have facilitated advancements for children with autism that were not forthcoming via other methods (Sundberg & Michael, 2001; see also Larsson, 2012). Applied behaviour analysis (ABA) has traditionally focused on language development as a pivotal intervention for populations with developmental disorders, particularly children with autism spectrum disorder (ASD). Behavioural programmes designed to develop or enhance language skills in children with autism use principles such as positive reinforcement, shaping, modelling, prompting, fading and other methods derived from experimental

behavioural research, frequently commencing with one-to-one tuition from a specially trained ABA instructor for the early learner.

Skinner's (1957) *Verbal Behavior* is the foundation of ABA language and communicative teaching programmes, and Skinner's language theory proposed a functional, as distinct from a structural account of language. Verbal behaviour was described as essentially operant behaviour in that it is shaped by contingencies of environmental reinforcement similarly to other behaviours. Skinner conceded that verbal behaviour required a listener with a relevant history of reinforcement to mediate contingencies (e.g., to provide access to drinking water when asked for water, which reinforces asking in the behavior), but in the main he proposed that verbal behaviour was learned by fundamentally similar contingencies as other behaviours. Skinner (1957) outlined a number of verbal operants such as mands, tacts, echoics, textuials and intraverbals. A mand is an elementary verbal operant in which a response is under the functional control of a motivating operation (i.e. satiation or deprivation) and followed by specific reinforcement (i.e., a mand for water is reinforced by the provision of water); a mand is typically a demand or request. Manding allows the speaker indirect access to environmental reinforcers and is thought to play an important role in language acquisition (Skinner, 1957; Sundberg, 2004). It is thus usually the first verbal operant targeted by behavioural teaching programmes. The tact is also considered a basic verbal operant in Skinnerian terms. Tacts involve a response that is under the functional control of a nonverbal discriminative stimulus and followed by generalized social conditioned reinforcement. The nonverbal discriminative stimulus may be any object in the environment (e.g., the presence of a tree in the environment may evoke the verbal

response "tree") or the entire environment itself. A tact may be termed a naming or labelling response (Sundberg, 2004) but this is a simplified description.

An echoic is a response under the functional control of a verbal discriminative stimulus that has point-to-point correspondence and formal similarity with the response (e.g., hear *tree* say *tree*). Intraverbals involve a response that is under the functional control of a verbal discriminative stimulus that does not have point-to-point correspondence with the verbal stimulus (e.g., *how are you/ I'm fine*). Intraverbal behaviour is basis for social interaction and conversation, as well as academic and intellectual behaviour. Textuals have similar functional control properties as echoics but they do not have formal similarity with the response. That is, a textual is under the functional control of a verbal discriminative stimulus that has point-to-point correspondence but not formal similarity with the response. Textuals typically involve reading printed text (Sundberg, 2004). There are also a number of other minor verbal operants, many of which are extensions of the operants described above or specific cases of the operants above.

Verbal Behavior (Skinner, 1957) explained the development of language as occurring through the learner's history of reinforcement, and ABA language programmes thus focus on reinforcement of Skinner's verbal operants in educational systems (Lovaas, 1977; Greer, 2002). Frequently, learning trials arrange contingencies with a constant antecedent (e.g., "What is your name?" as antecedents to a target response such as "My name is Susie") which is immediately followed by reinforcing consequences. This type of temporal and contingent relationship between the target response and reinforcement has been termed direct reinforcement (Hayes, Barnes-Holmes, & Roche, 2001; Kilroe, Murphy, Barnes-Holmes, 2011). A limitation of this approach is that reinforcing every specific utterance may be costly

in terms of time and resources, whereas targeting the generativity that is characteristic of language to promote exponential gains may be additionally beneficial, resulting in novel speech sequences that emerge untaught for a given individual. The early capacity to generate novel speech is ubiquitous in typically-developing children and may be considered a rather amazing phenomenon that requires a precise understanding. Emergent untaught verbal behaviour was explained by Skinner (1957) largely in terms of analogy and recombination, however, a contemporary behavioural account has suggested a more detailed and comprehensive explanation for generativity and other complex aspects of language (Hayes et al.). Hayes et al. developed Relational Frame Theory building on the early stimulus equivalence research by Sidman (1971).

Stimulus equivalence (Sidman, 1971, 1994, 2009) refers to the emergence of untaught stimulus relations between stimuli following a history of reinforcement for initial taught relations. For example, Sidman taught developmentally disabled children to select the relevant picture given the spoken word *doll*, and to select the printed word *doll* given the spoken word *doll*. It was subsequently found that the children could match the picture to the printed word even though this had never been explicitly trained. Put in simple terms, training the relations A-B and A-C, may result in the emergence of the untaught relations B-C and the reverse, C-B, without direct reinforcement for establishing the latter relations.

Stimulus equivalence as described by Sidman (1971), involved reflexivity, symmetry and transitivity. Reflexivity refers to the matching of a sample to itself (AA). Symmetry refers to a derived bi-directional relation between stimuli; for example, when a learner is given the spoken word *car* to identify a picture of a car (AB), then without training may say the word *car* in the presence of a picture of a car

(BA). Transitivity occurs when a learner is taught that the written word *car* (A) is equivalent to the picture (B) which is in turn equivalent to the spoken word (C), and the learner subsequently derives (untaught relation) that the written word is equivalent to the spoken word (AC) (Sidman, 1994). Sidman's research on stimulus equivalence became a critical step in the development of a behaviour analytical account of basic language that could encompass the emergent generative nature of language (Hayes et al., 2001). This pivotal and extensive work on the phenomenon of stimulus equivalence paved the way for the development of a more complete and comprehensive modern behaviour analytical account of human language known as Relational frame theory (RFT; Hayes et al., 2001).

RFT posits that equivalence is only one of many different types of derived relational responding, with equivalence being considered the most basic type of derived relational responding (DRR). RFT uses the term *relational frame* to describe particular types of relational responding. Relations frames include coordination (same), distinction (difference), opposition, comparison (more/less) sequence (before/after) and a myriad of other relations (see Hayes et al., 2001). Thus, within RFT not all relations involve equivalence or bidirectional symmetry, for example a sequential relation such as A came before B entails the derivation that B came after A which is not a symmetrical relation. RFT employs the term "mutual entailment" to encompass such relations in addition to equivalence relations. The term "combinatorial entailment" is used to describe transitivity relations, for example, when relations between A and B and between B and C combine to entail the relations between A and C and between C and A. That is, if A comes before B and B comes before C then the combination of these sequential relations may result in derivation that C comes after A and A comes before C (O'Toole & Barnes-Holmes,

2009). The term “transformation of function” is an important feature of DRR involving a change in stimulus functions based on established relations between stimuli. For example, a child who is taught via direct reinforcement that the word *sweet* refers to the item *sweet*, and subsequently learns that *candy* is the same as *sweet* (equivalence relation), may then become excited when asked “do you want some candy”, even though the word *candy* has never been directly reinforced with actual sweets. In this case some of the functions of the word *candy* are transformed due to the equivalence relation with the word *sweet*.

Relational responding that can come under to control of arbitrary contextual cues, and is not purely controlled by the formal properties of the stimuli nor direct experience with them is known as Arbitrarily Applicable Relation Responding (AARR; Hayes et al., 2001). The above example of transformation of function is also an example of AARR. The relation between the words “*sweet*” and “*candy*” is not defined by any physical properties but rather the context in which the word “*candy*” was learned. The related responses to hearing the word “*sweet*” are brought to bear on “*candy*” as a result of the social learning context; the relational responding is arbitrarily-based in the sense that the words “*sweet*” and “*candy*” are not equivalent in any physical sense, nor do they physically resemble that which they reference. These types of complex relational phenomena are thought to underpin most of, if not all, of human language and cognition (Hayes et al., 2001).

Naming objects is another prime example of AARR, a parent may point to a toy car and say the word *car* and subsequently reinforce any orienting response towards the car. In this situation the parent is training the arbitrary relation between the word *car* and the object itself. Similarly, on other occasions the parent may hold up the car and model the naming response by saying “*car*”. These types of name-

object and object-name exemplars occur regularly in early childhood, after a significant number of these exemplars the child no longer needs to be explicitly taught both the name-object and the object-name relations but abstracts out the appropriate relation in novel instances. For example, after sufficient multiple exemplar training, the child is presented with a novel object and told “this is a Coati”, the child spontaneously reverses this relation without any further explicit training and independently derives the relevant name-object relation, in other words the child generalises the derived coordination relation to novel instances (Hayes et al., 2001). It follows that when attempting to explicitly teach complex language skills to those who do not readily attain it naturally that multiple exemplar training may be necessary to achieve this level of generativity.

This ability to demonstrate emergent relational performances including AARR is readily seen in typically developing children at a very young age (Lipkens, Hayes, & Hayes, 1993). However this ability does not occur with such ease in developmentally disabled participants; for example, hearing impaired children with significant deficits in their verbal ability failed to form equivalence classes despite learning the necessary conditional discriminations just as well as their non-impaired counterparts (Barnes, McCullagh, & Keenan, 1990). Similarly, Devany, Hayes, & Nelson, (1986) compared three groups of children who were matched on a conventional measure of mental age, but differed with regard to verbal ability. They found that both typically-developing and verbally-able developmentally delayed children were able to form equivalence relations while developmentally delayed children with little to no spontaneous communication did not.

This area of behaviour analysis is catching fire with an increasing number of studies exploring relational responding. For example Nuzzolo-Gomez and Greer

(2004) examined the effects of multiple exemplar instruction on the emergence of untaught mands or tacts of adjective-object pairs across four students with autism or developmental disabilities using a multiple probe design. It was found that untaught mands or tacts emerged following multiple exemplar training for all students, findings consistent with RFT. On balance Hall and Sundberg (1987) examined the emergence of untaught mands following tact training. However, the study was unsuccessful at producing untaught mands; the participants required direct mand training providing somewhat contradictory evidence to that above and highlighting the importance of multiple exemplar instruction.

As stated previously, emergent and advanced language development is frequently deficient in the language repertoires of children with autism (Barnes et al., 1990; Devany et al., 1986). The importance of this deficit is highlighted by the fact that those children who do develop advanced language skills have a significantly better prognosis than those who do not (see Sallows & Graupner, 2005; Carr, Binkof, & Kologinsky, 1978). Language has been described as a pivotal behaviour (Koegel, Koegel, & Carter, 1998) because it can result in access to a vast number of reinforcers; one cannot, for example, discover that the study of history or mathematics is reinforcing without first accessing these topics via language. Because language may facilitate exponential gains in reinforcement via access to social and academic pursuits, language and communication skills have been traditionally considered primary learning targets for ABA programmes with children with autism and related speech deficits (Carr et al., 1978; Le Blanc, Esch, Sidener, & Firth, 2006; Sundberg & Partington, 2001; Tincani, 2004). More recently, behavioural researchers have proposed a synthesis of ABA programmes with behavioural research on derived relational responding (DRR) that encompasses generativity of

the type emergent in language (Barnes-Holmes, Barnes-Holmes, & Cullinan, 2000; Murphy, Barnes-Holmes, & Barnes-Holmes, 2005). This approach could potentially avoid the necessity of directly teaching almost every possible utterance to children with autism who do not show untaught or emergent language. Since Sidman's (1971) early work in stimulus equivalence the promise of programmes to enhance emergent or derived responding has been long awaited in the field of ABA, and a small number of recent applications have successfully demonstrated emergent responding with populations with intellectual disability (Murphy et al., 2005,2010; Berens & Hayes, 2007; Rehfeldt & Root, 2005). In addition, template applications have been set down in a handbook so that ABA practitioners can design teaching programmes to incorporate emergent or "derived" responding (see Rehfeldt, Barnes-Holmes, 2009).

Researchers aiming to synthesise Skinner's research and RFT in practical applications and interventions combine tactics such as positive (scheduled) reinforcement, prompts, fading, and such with a DRR component targeting emergent responding. For example, studies by Murphy et al. (2005) used these methods in a mand-training programme (to teach requesting), and participant children with autism demonstrated emergent untaught mands as a result. The study indicated that it is possible to establish derived mand and derived reinforcer functions following direct mand training and conditional discrimination procedures taught using positive reinforcement. This study provided evidence for the possibility of developing an effective means of teaching the verbal operants in a way that promotes the generativity that is characteristic of language to those who do not readily display it. Similarly, other research has successfully established DRR with populations with

intellectual disability (Murphy et al., 2010; Rosales & Rehfeldt, 2007; Berens et al., 2007)

Kilroe et al., (2011) built upon the work of Murphy et al., (2005) using a computerised teaching programme (T-IRAP) to teach and gradually increase the complexity of relational responding skills with four children with autism across a series of experiments. The study used positive reinforcement and started with teaching relational skills of coordination, then comparative relations, followed by oppositional relations and then analogy. The T-IRAP involves computerised stimulus presentation using the freely available Implicit Relational Assessment Procedure adapted for the purpose of teaching relational responding (<http://irapresearch.org/downloads-and-training/>; IRAP; Barnes-Holmes, et al., 2006). The study used a multiple baseline design across participants to compare participants' relational learning in terms of speed accuracy during Table-Top (TT) and T-IRAP teaching conditions to determine if the latter might be useful as a supplementary teaching tool. Initially non-arbitrary (physically-based) relations were taught, with children progressing to more complex arbitrary relations including derived relations. Results throughout showed greater speed and accuracy during T-IRAP teaching compared to Table Top teaching. However, it should be emphasised that the researchers did not suggest that the T-IRAP should replace TT teaching, but suggested that the T-IRAP might be used for student practice without the presence of an instructor, and for building speedy or 'fluent' rates of accurate responding with children with autism. Rate of responding has been seen as important to learning from the outset in behaviour analysis (Skinner, 1953), and indeed Precision Teaching. Precision Teaching was pioneered by Ogden Lindsay and is heavily based on Skinner's operant conditioning; it focuses on frequency as its primary datum (i.e.

average number of correct response per minute; White, 1986). Binder (1996) suggests that responding that is both accurate and high frequency or “*fluent*” will be retained longer and less likely to be affected by distracting conditions. It has been found that students who reach high levels of fluency outperform students taught in traditional ways (McDade, Brown & Vance, 1993). Binder (1996) also argued that fluent responding is more likely to be applied, adapted or combined in new learning situations even in the absence of instruction; in essence it is more likely to be applied arbitrarily in novel instances. For a detailed account of Precision Teaching see Lewis, 2004; Lindsley, 1990a, 1990b, 1991, 1992a, 1992b, 1995; Raybould & Solity, 1988.

Relational Flexibility and Intelligent Behaviour

The ability to respond rapidly and correctly has long been considered to be part of intelligent behaviour, and indeed practically all ability or IQ tests involve a speed component. High speeds in responding may indicate cognitive flexibility. In fact, a well-known theory of intelligence, the Cattell-Horn-Carroll theory (CHC theory; an amalgamation of Horn & Cattell’s (1966a; 1966b) Gf-Gc theory and Carroll’s (1993) three-Stratum Theory) has flexibility at its core and a key feature of fluid intelligence. Defining fluid intelligence as the ability to solve problems, the capacity to think logically, and the ability to adapt to and thrive in novel situations. Fluid intelligence is taught to be independent of acquired or learned knowledge, it is the ability to analyse, identify problems and relationships between things and logically extrapolate the solution (Cattell, 1963). The adaptability mentioned in the CHC theory relates closely to flexibility as described in this paper, responding to multiple stimulus relations is itself a form of relational flexibility (Hayes et al., 2001). As such it follows that rigidity the antithesis of flexibility may be detrimental to intelligence. This may be particularly true when treating children with diagnosed

autism. Autistic children have a tendency towards rigid behaviour with many displaying repetitive and stereotypic behaviours, aspects of rigidity are in fact a defining feature of autism (DSM-IV-TR; APA, 2000). Recently a number of authors have argued that there is a need to facilitate flexibility within programs designed to establish relational repertoires particularly in these applied settings with populations where rigidity is a problem (O'Connor, Barnes-Holmes, & Barnes-Holmes, 2011).

As stated previously, rapid, flexible responding has long been considered to be essential if not central to intelligence, for example many studies from across psychology have focused on developing measure of cognitive flexibility, for example the Stroop test (Stroop, 1935) and the Wisconsin Card Sorting Test (Grant & Berg, 1948) are two well-known measures of cognitive flexibility. The Stroop test requires the participant to attend to stimuli and rapidly discern which features are relevant and which can be ignored, the Wisconsin Card Sorting Test requires participants to demonstrate flexibility in the face of changing schedules of reinforcement. In modern intellectually complex environments, relational flexibility may confer an ability to adapt to ever changing contingencies and when necessary to respond in a manner which is contradictory to previously established norms. As a classic example, Copernicus' heliocentric universe contradicted contemporary belief in a geocentric universe in which humankind were central; inflexibility meant the point could not be conceded and Copernicus was confined under house arrest. A simpler example might be a child with autism, having learned that it is not acceptable to 'tell tales' fails to disclose that s/he (or another student) has been beaten or abused by other students. Children with autism frequently have difficulties with understanding when it is appropriate to break a rule once learnt (Grandin, 2008; von

Hahn & Murphy Bentley, 2013) in a sense this is the opposite to flexibility, and many individuals with autism are said to show 'rigid' responding.

O'Connor et al. (2011) found that children with autism appeared to have greater difficulty in establishing contextually controlled symmetry (correct) and asymmetry (incorrect). It was found that the typically developing children and a small number of those with autism learned to switch between symmetry and asymmetry responses with ease but other children with autism required specific training in this skill. Conceptually speaking those children who learned to respond to symmetry and asymmetry had greater levels of flexibility when undertaking the relational responding task than those who struggled. It has also been noted that some children with autism, who showed difficulty attaining asymmetry already had a significant history of exposure to educational programs that were tailored to their individual behavioural needs. But, since the children were being asked to effectively produce the “wrong” answer, a skill which is rarely if ever taught in educational settings, they found it difficult. This suggests that flexibility is not an implicit result of this type of teaching, specifically teaching a child the concept of “same” does not implicitly teach the concept of “not same” or “different”. It may be argued that the type of teaching procedures typically used in educational settings which are highly structured and teach only “correct” answers may foster rigidity rather than combat it. However, O'Connor et al. have shown that it is possible to teach children with autism to respond flexibly with specific training.

Relational Frame Theory posits that advanced cognitive behaviours including language, mathematics and other symbolic phenomena are all underpinned by complex relational responding, and there is a growing body of empirical evidence to support the theory. The theory predicts, unsurprisingly, that fluency in relational

responding would be essential to intelligent behaviour (Hayes, 2011), and that correlations may be seen between rapid relational responding and high scores on ability tests (e.g., IQ scores). Research testing the prediction by O'Hora, Peláez, and Barnes-Holmes (2005), examined the relationship between performances on a number of subtests of the *Wechsler Adult Intelligence Scale-Third Edition* (WAIS-III; Wechsler, 1997) with performances on complex relational tasks in which they were required to relate novel stimuli in accordance with similar/different and temporal relations. It was found that participants who passed a complex relational task produced significantly higher scores on the verbal and arithmetic subtests of the WAIS-III than participants who failed the relational test. A correlation between the number of correct responses on the relational task and the number of correct responses in the vocabulary and arithmetic subtest was also found.

Another study which aimed to further the findings of O'Hora et al., (2005) compared performances on a relational responding task to performance on the Kaufman Brief Intelligence test (K-BIT; Kaufman & Kaufman, 1990). The study measured response latency during both consistent trials, where participants responded in line with pre-established verbal relations, and inconsistent trials, where participants responded against these relations. A difference score was calculated by subtracting response latencies for the consistent trials from the inconsistent trials, difference scores were considered a measure of relational flexibility. It was found that shorter response latencies and smaller difference scores predicted higher IQ scores, providing further support for RFT's perspective on the relationship between human intelligence, verbal ability and flexible relational responding (O'Toole et al., 2009). Perhaps more surprisingly, O'Toole et al. found that speed was not the strongest predictor of high IQ scores; the strongest predictor was being able to

respond rapidly and *flexibly* when presented with relational tasks. For example, relational responding tasks involved changing contingencies so that relations initially taught were subsequently reversed and then double reversed. Participants who were able to maintain high levels of speed and accuracy throughout these procedures were those whose scores correlated with high IQ verbal ability scores.

Following on from basic research and predictions, Cassidy et al., (2011) conducted what may prove to be a ground-breaking applied study found that a training programme that employed multiple exemplar training in complex relational responding for a number of different relations (Same, opposite, More than, less than) resulted in significant gains in participant IQ. The study measured IQ at three stages; baseline, following stimulus equivalence training and following relational training in both experimental and control groups. It was found that the experimental group IQ increased marginally following stimulus equivalence training, and substantially following relational training, while no significant increase occurred in the control group. This suggests that multiple exemplar relational responding training can greatly improve IQ scores, also as only marginal increases followed stimulus equivalence training further support for RFT claims that complex relational responding underpins intelligence and higher order human cognition is garnered. However, the linear sequence of stimulus equivalence training followed by relational training along with a long delay between stimulus equivalence training and relational training make it difficult to distinguish between delayed effects of stimulus equivalence training and full relational training. It may be the case that stimulus equivalence training facilitated the subsequent acquisition of relational frames and as such the large IQ gains recorded following relational training. However, Cassidy et al. (2011) also conducted a similar procedure with a second population, this time

those with educational difficulties, and found a refined relational training intervention to be sufficient (and possibly necessary) to bring about significant increases of more than 1 SD for 7 of the 8 participants. Another interesting feature of the Cassidy et al. study was that improvements in full scale IQ at follow-up were not predicted by full scale IQ at baseline, rather a measure of fluency (summing test performances on training trials and dividing by the number of training blocks) significantly predicted changes in full scale IQ. Essentially, it was found that it was those who learned DRR efficiently and effectively improved their full scale IQ scores. This study, although preliminary, points to multiple exemplar training in relational skills as a basis of an intervention to improve general cognitive functioning. When taking the results of this study along with that provided by O'Connor et al. and the predictions made by RFT about flexibility in relational responding it suggests that providing multiple exemplar training in relational responding with a flexibility component to populations where rigidity is problematic may prove to be very beneficial.

The Current Study

The current study aims to build upon the finding of two separate studies. Firstly, Kilroe et al., (2011) successfully used the T-IRAP for teaching relational responding skills to children with autism, and used a multiple baseline design to demonstrate greater speed and accuracy with the T-IRAP compared to TT. The current study extends this research by using an alternating treatments design which is not subject to sequence effects that may result with multiple baseline design. A further extension is that the study will teach more complex relational responding with double contingency reversals as this is more likely to impact IQ (O'Toole et al. 2009). It builds upon the work of Cassidy et al. (2011) by examining the IQ and

verbal abilities of children with autism at pre and post test to determine if relational training will result in similar increases in IQ.

The current study aims to teach children with diagnosed autism fluent and flexible relational responding, while examining which of two teaching methods is more effective and simultaneously explore any potential impact on IQ and verbal ability. To do this pre and post intervention measures of IQ and verbal ability will be taken to assess if relational training generates any gains on these measures. An alternating treatments design will compare a TT procedure with a computerised procedure (T-IRAP) to teach the same sets of relational stimuli. Complexity in relational flexibility will be gradually increased over time using coordination and distinction relations. Initially stimuli will be simple and non-arbitrary, progressing in complexity to include arbitrary stimuli and double contingency reversals.

It is predicted that the T-IRAP will be more efficient and more effective at teaching relational concepts than TT procedures, partially due to the fluidity, speed and consistency of stimulus presentation T-IRAP. In addition to comparing the effectiveness of the T-IRAP to standard TT to teach complex relational responding the study will explore if training in complex relational responding will impact on IQ. Effectively, it is predicted that the participants will score a higher percentage correct in a shorter period of time when using the T-IRAP vs. TT procedures. It is also predicted that following complex relational training in flexibility a gain in IQ and verbal abilities will be seen in line with that found by Cassidy et al., (2011). In brief, the three main research questions are, 1) Can children with autism and low ability scores (e.g., IQ/verbal ability/adaptive functioning) learn to respond to relational stimuli in a flexible manner using the T-IRAP, 2) Will T-IRAP produce greater accuracy and/or faster responding than TT with this population, 3) will teaching

relational flexibility result in any gains in global or verbal cognitive functioning indicated via the children's measured scores on the Peabody Picture Vocabulary Test 4th Edition (PPVT-IV; Dunn & Dunn, 2007) and the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990).

-Chapter 2-

Method

Participants and recruitment

Six children were recruited from a local school that specialised in teaching children with intellectual and developmental disabilities where they receive 22.5 hours of schooling a week. All participants had been previously diagnosed with autism by a clinical psychologist independent of the current research in accordance with the criteria set out by the DSM-IV (APA, 2000) or the ICD-10 (WHO, 1992). All children were exposed to ABA teaching programmes during school hours. To recruit participants, consent forms with principle information were sent out to parents of all children attending the school via the children's' schoolbags, and an information sheet with expanded information about the study accompanied the consent form (Appendices 1 & 2). Parents were asked to carefully read all the information provided and if they wished for their child to participate in the study to return the signed consent form to the researcher.

A total of six parents consented to their children's' participation. The participants were five boys and one girl who ranged in age from 7-12 years old (with a mean age of 10.24 years), all with diagnosed autism spectrum disorder. Two participants were diagnosed as severe, two were diagnosed as moderate-severe and one was diagnosed as moderate. One participant failed to complete Phase 1 and his data are excluded. For the remaining participants pseudonyms are used throughout.

Evan is a 12 year old boy with moderate-severe autism who was currently awaiting a clinical assessment for suspected obsessive compulsive disorder. Evan's verbal behaviour was assessed using the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990) and the Peabody Picture Vocabulary Test, fourth

edition (PPVT-IV; Dunn & Dunn, 2007), he achieved standards scores of 40 and 20 respectively.

Steven is a twelve year old boy diagnosed with severe ASD and a moderate learning disorder. Steven initially received a diagnosis of moderate-severe autism but this was revised to severe two years later. Stevens verbal ability was also assessed using the K-BIT and PPVT-IV; he received standard scores of 40 and 20 respectively. Steven displays signs of "rigidity" around his daily routine, for example, selecting the same reinforcers in the same order each day, leaving the most preferred activity until the end regardless that he frequently fails to complete the schedule and access the most preferred reinforcer. Steven has a long history of challenging behaviour towards tutors, usually preceded by increased vocalisations and the shouting of specific terms. Setting events may be a change in the daily routine, presentation of a difficult task, a significant number of corrections during an academic task, or loud and persistent noise.

Andrew is a twelve year old boy with diagnosed ASD in the severe range. Andrew has a reasonably well developed verbal repertoire, when assessed he achieved standard scores of 40 and 23 on the K-BIT and PPVT-IV respectively. Andrew speaks about a limited number of topics, rarely initiates conversations about novel topics, and displays signs of poor comprehension around new topics or when presented with novel information. Andrew's repetitive behaviours are controlled using a DRL procedure; he demands a high level of social attention and has difficulty waiting for the attention of tutors or other adults. Andrew is very distracted from tasks by other students in class, especially when they are using the class computer.

Kevin is an eight year old boy with diagnosed ASD in the moderate range. Kevin has a limited expressive verbal repertoire scoring standard scores on 62 and 50 on the K-BIT and PPVT-IV respectively. The most significant challenge to Kevin's learning is non-compliance when presented with a non-preferred activity.

Ann is an eight year old girl with diagnosed ASD in the moderate-severe range. Ann's communicative repertoire is limited specifically her expressive vocabulary, she obtained standard scores of 40 and 34 on K-BIT and PPVT-IV tests respectively. Her learning is challenged by non-compliance when presented with a non-preferred task. Ann is very easily distracted from tasks by environmental factors such as other people moving around, noises, or the presence of any preferred items in the room. She generally displays low motivation to work on academic tasks and engages in mild stereotypy by repetitive flicking of her fingers.

Ethical Considerations

The current thesis was ethically approved by the ethics committee in the department of Psychology at the National University of Ireland, Maynooth on the 23rd of October 2012. In addition to submission to the Ethics Committee at the Department of Psychology, NUI Maynooth. The Director of services of the school was in agreement with the onsite supervisor that the current study meet the ethical requirements of the school. Both Director and Supervisor are qualified senior behaviour analysts with high level ethical training and are both accredited Board Certified Behaviour Analysts. The ethical considerations here pertain to the entire thesis.

Consent and assent. Prior to any data collection consent forms with principle information were sent out to parents of all children attending the school via the children's' schoolbags, and an information sheet with expanded information

about the study accompanied the consent form (Appendices 1 & 2). Parents were asked to carefully read all the information provided and if they wished for their child to participate in the study to return the signed consent form to the researcher. Also, approximately half way through the data collection process a continued consent form was sent home to the parents of each participant (Appendix 3). This form reminded parents of the study and offered them the chance to withdraw their consent at that point should they wish, no parent withdrew their consent at this or any other stage throughout the study. Only children whose parents returned signed consents forms were included as part of the study.

In the case of children or vulnerable individuals such as those with autism or other developmental disorder it is frequently not considered appropriate to attempt to obtain *consent* from the individual themselves as the participant may not be deemed capable of providing consent to participate if s/he has not the capacity to fully comprehend the procedures or implications of participation. In such cases it is customary to attempt to gain participant *assent* (an expression of approval or agreement) from the individual; for example, if the participant had some speech or a verbal repertoire and can answer *yes* or *no*, they were asked if s/he would like to work with the researcher. If the participant did not have an established speech/verbal repertoire an attempt was made to ascertain if that the participant was comfortable (or not) with research procedures was made by monitoring other means of participant communication, including facial expression and approach behaviours that may be indicative of assent or otherwise. For example if the child was willing to approach the researcher when asked, or was willing to sit at a table and engage in some way with the researcher and research materials, these types of behaviours would suggest that the child was at least comfortable in the presence of the research. On the other

hand, if the child was not willing to approach the researcher, and was not willing to engage in any minor way with the researcher and research material it was considered that assent was not given by the participant at that time and research sessions did not take place. All participants showed signs of comfort with the researcher and assent was assumed in all cases. The current research only used positive reinforcement techniques and provided frequent short breaks.

Assent was sought before each research session, as continued assent was monitored throughout. On one occasion one participant, Steven, showed signs of discomfort and distress, Steven began to engage in historically documented precursor behaviours, namely shouting specific phrases during a research session. Instances of physical aggression with Steven were previously noted to occur on a semi regular basis during normal school teaching and a BCBA implemented reactive strategy was already in place. Precursor behaviours occurred within a single session and that session was terminated immediately, as per the predefined protocol and no physical aggression occurred, upon consultation with the classroom tutors it was discovered that an unexpected change to Stevens schedule had occurred earlier that day which is a common setting event/precursor to potential challenging behaviour. The researcher had a well-established rapport and positive relationship with the participant and was very familiar with the reactive strategy for dealing with instances of physical aggression. Research sessions were not permanently halted as no other instances of this nature occurred.

Some potential risks were identified as a visual electronic device was being used, specifically a computer screen, there was a small risk of potential seizures, although the risk was minimal, parents were advised that children with a history or seizures should be excluded from the study. If parents of a child with a known

history of seizures wished to allow their child to participate despite the negligible risk they were to be referred to the BCBA supervisor who would reiterate that that it would be advisable not to expose their child to even this minimal risk. If a parent felt strongly, nonetheless, that their child participates, their child was not excluded from the study. This was a situation that did not arise but given the population needed to be prepared for.

Psychometric tests. Pre and post assessments were conducted with each participant to determine levels of verbal ability, the PPVT-IV (Dunn et al., 2007) and the K-BIT (Kaufman et al., 1990) were used. The researcher was deemed qualified to use these test given the test manufactures recommendations, for the K-BIT, manufactures state that the test is “easy to administer and score and may be given by individuals without any formal specialized training” (p.11; Kaufman et al., 1990). Test manufactures of the PPVT-IV, which is categorised as a level B test, state that the test may be “administered and scored by individuals with a range of educational backgrounds, provided that they are thoroughly familiar with the test materials, have received training in procedures for individual test administration” (p.iv; Dunn et al., 2007). Also the researcher received 5 hours of training in the administration and interpretation of the relevant psychometric tests by a qualified psychometrician as part of the Doctorate in Psychological Science: Behaviour Analysis and Therapy course. Furthermore having contacted Sigma Assessment Systems (A manufacturer of many psychological tests and assessments; see appendix 4) they suggest that the researcher is suitably qualified to administer level B tests, such as the PPVT-IV. Finally guidelines set out by the American Psychological Associations Task Force for test user qualifications (APA, 2000; appendix 5) stipulate that students with a level of training equivalent to that obtained by the researcher may use such tests for

research purposes provided the results from these tests do not guide or influence clinical decisions. Individual results from the assessments we're not be made available to either the school or parents as doing so may result in clinical decisions being made based upon them. Upon receipt of a formal written request from parent's access to individual results will be provided, in accordance with current freedom of information legislation, with formal written advice from the researcher and supervisor that the test scores should not be used to guide clinical decisions.

Data protection. All data gathered from both TT and T-IRAP was stored using pseudonyms on an encrypted hard drive using the Microsoft "encryption file system" which is available on all versions of windows. This method of encryption provides 265-bit encryption by default. A backup copy of all data was stored on a second encrypted drive as precaution should one drive fail. A key to identifying participant data was stored on an encrypted USB drive which was kept under lock and key by the researcher. All necessary data will be retained for the appropriate length of time in accordance with legal requirements and standard research practices, following which it will be permanently destroyed. All intended uses of data were outlined in the initial consent form, should further consent be needed additional consent forms will be sent to the participants.

Settings and Materials

IQ assessments. Each participant was assessed using two tests of verbal ability, the Kaufman Brief Intelligence test (K-BIT; Kaufman et al., 1990) and the Peabody Picture Vocabulary Test-Fourth Edition (PPVT-IV; Dunn et al., 2007) directly after school, with their respective class teacher present, acting in loco parentis. Both tests are administered using easels containing various visual stimuli and separate scoring sheets. Following initial assessments participants received 2-3

twenty minute research sessions' per-week. Sessions were conducted under standard classroom conditions during normal class hours in each child's respective classroom.

The K-BIT is a brief, individually administered test of verbal and nonverbal intelligence suitable for use with populations ranging from 4-90 years. It consists of two subtests, a vocabulary subtest (comprising of part A, Expressive vocabulary, and part B, Definitions) and a matrices subtest. The vocabulary subtest assesses crystallised or verbal intelligence, by assessing word knowledge and verbal concept formation. The matrices subtest assesses fluid or nonverbal intelligence by assessing the ability to perceive relationships and analogies, matrices involve pictorial stimuli rather than words. A correlation coefficient of .75 between K-BIT composite IQ scores and the Wechsler Adult Intelligence Scale-Revised (WAIS-R, Wechsler, 1981) full scale IQ scores is reported by the authors (Kaufman et al., 1990). Correlations between vocabulary subtests of the K-BIT and WAIS-R verbal IQ of .60 have been reported, while the matrices subtest correlates .52 with performance IQ on the WAIS-R. These provide strong support for the construct validity of the K-BIT. When examined, high levels of concurrent validity between the K-BIT and the Test of Nonverbal Intelligence (TONI; Brown, Sherbeno, & Johnsen, 1990), the Slosson Intelligence Test (SIT; Jensen & Armstrong, 1985), the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) and the Wechsler Intelligence Scale for Children- Revised (WISC-R; Wechsler 1974) have been found (Kaufman et al., 1990).

The PPVT-IV is a test of receptive vocabulary that provides a quick estimate of verbal ability. The test take approximately 30 minutes to administer, it is given verbally using pictorial stimuli and requires no reading ability. Each trail is multiple choices and can be answered by pointing to the correct stimulus or using other

gestures. As such, no vocal verbal repertoire is needed. The test can be used for anyone from 2;6-90+ years and is particularly well suited to being used with those with intellectual disabilities. The test also includes two parallel forms each containing 228 items. The PPVT-IV is a test of receptive vocabulary and as such does not directly measure fluid intelligence. However, as the PPVT-IV correlates highly with other tests of intelligence that do measure the full range of intelligences it may be a good analogue that can be administered quickly especially to students who do not have an expressive verbal repertoire.

The PPVT-IV has strong internal consistency reliability, with a split half reliability coefficient of .94 for both forms. The test also has a reliability coefficient of .89 between the alternate forms of the test and .93 for test-retest reliability. Scores obtained from the PPVT-IV correlate well with a number of other established tests, with an average correlation of $r=.82$ across ages and content areas for the Expressive Vocabulary test, Second Edition (Williams 2007), an average of $r=.61$ with the Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999) an average $r=.71$ with the Clinical Evaluation of Language Fundamentals - Fourth Edition (CELF-4; Semel, Wiig, & Wayne, 2003) and .61 with the Group Reading Assessment and Diagnostic Evaluation (GRADE; Williams, 2001).

Table Top. Table Top (TT) procedures were conducted at the participants' desks under standard classroom conditions. TT procedures were carried out using laminated cards (A4 size) each containing a sample stimulus at the top and a comparison stimulus just below. These images were sourced via the internet and were identical to those used in the T-IRAP. All images used in the TT procedure were printed to be identical, in size, shape and colour, to those used in the T-IRAP. Two additional printed cards (8"x2"), one with the word "Different" and one with

the word “Like” printed in boldface were also used. Total duration of the TT procedures was recorded using a timer. The number of correct and incorrect responses was recorded using pen and paper on a specially constructed data sheet.

T-IRAP. The T-IRAP was administered using an Intel Celeron M 450 laptop computer with a 15 inch screen, running the Microsoft Windows 7 operating system, at the participants’ desks under normal classroom conditions. The T-IRAP is a computer programme written in Visual Basic (Version 6.0) and is freely available to download and use at <http://irapresearch.org/downloads-and-training/>. It controls all aspects of stimulus presentation and data collection. The T-IRAP software records all correct and incorrect responses, as well as response latency which can be summed to give the total duration of the T-IRAP procedure. The programme was designed so that each trial presented a sample stimulus, a comparison stimulus, and two relational terms (e.g. Different/Like). Participants responded by pressing a key on the computer keyboard (e.g., D for Like and K for Different). All visual stimuli were sourced via the internet or created by the researcher using Microsoft paint.

Design

Baseline ability test scores were collected using the PPVT-IV (Dunn & Dunn, 2007) and the K-BIT (Kaufman & Kaufman, 1990). An alternating treatments design was subsequently employed to compare the relative effectiveness of TT procedures and T-IRAP at simultaneously teaching coordination and distinction relations, in this case Like/Different, to five children with diagnosed autism. The order of the intervention types was counterbalanced across participants; approximately half of the participants received the TT procedures first, and the other half received T-IRAP first. This was to control for any potential sequence effects. The order in which the K-BIT and PPVT-IV assessments were presented was also

counterbalanced across participants. Post-test ability scores were taken and compared with baseline data to determine if training in advanced relational responding skills impacted upon IQ scores which would otherwise be expected to remain stable.

Inter Observer Agreement

An independent observer who had extensive ABA experience was trained to record TT data. The observer sat at the opposite end of the table to the participant and researcher and did not interact with the participant in any way. Approximately 20% of all TT trials were observed and a trial by trial analysis was conducted. 100% agreement between researcher and the independent observer was found. The T-IRAP automatically records all responses independent of the researcher as such no IOA data was necessary for T-IRAP sessions.

Procedure

IQ assessments. Baseline K-BIT and PPVT-IV scores were calculated for each participant before they were exposed to TT and T-IRAP procedures. The K-BIT and PPVT-IV were both administered as per their respective instructions in the participants' classrooms' with their class teachers present. K-BIT trials in the vocabulary subtest involved presenting a picture, of say a "car" and asking the participant "What is this?"; matrices subtest trials consisted of pictorial stimuli, presenting a sample stimulus at this top of the page along with a number of comparison stimuli and saying "this one goes with this one, which one of these goes here", matrices also included pattern matching trials. PPVT-IV trials presented participants with four pictures on a page and required participants to "find X". Tests were conducted after school hours to avoid taking up teaching time with non-

essential assessments. The order in which these test were presented was randomly alternated across participants to prevent any possible order effects.

General experimental sequence. The relational responses targeted were coordination and distinction (Like/Different). Following initial assessments using the K-Bit and PPVT-IV all participants entered into Study 1. Participants were randomly assigned to either T-IRAP first or TT first conditions. When training the coordination and distinction relations, a number of sets of stimuli were used; these sets were designed to become progressively more complex. Stimuli set 1 consisted of stimuli that varied along simple non-arbitrary physical properties, in this case shape (squares and triangles). Stimuli set 2 used stimuli that varied in respect of which class the stimuli belong to (animals and food). Each participant progressed at his/her own rate through the experiment. In general, participants only moved to the next set of stimuli when mastery had been achieved on previous stimuli set, as such each participant progressed at his/her own rate through the stimuli sets. It was planned that each participant would be exposed to all stimuli sets but time constraints and the considerable difficulty experienced by some participants with early stimulus sets precluded exposure to complex stimuli sets for some participants.

Table Top. In the TT condition laminated cards with the word “Like” and “Different” printed on them were placed on the desk in front of the participants. With “Like” on the left and “Different” on the right, as is the positioning in the T-IRAP. Cards containing two stimuli, a sample stimulus at the top and a comparison stimulus just below the sample were held up in front of the participant. When both the target and the sample stimuli were “Like”, two squares or two triangles, participants were required to touch the card with the word “Like” printed on it. When the target and sample stimuli were different, for example one triangle and one

square, the participants were required to touch the card with “Different” printed on it. All correct responses were reinforced with social praise, edibles and/or tokens, which could be exchanged for a preferred item at the end of the task. Reinforcement contingencies were tailored for each participant and schedules closely mimicked those used on a daily basis in normal school teaching. All incorrect responses were corrected using least to most guidance, and reinforcement was withheld, in accordance with standard procedure within the school.

General T-IRAP format. Before commencing, participants were provided with oral instructions by the researcher (appropriate instruction may be important to participants’ successful completion thus reducing attrition rates, so researchers may wish to view exemplar instructions available online at <http://irapresearch.files.wordpress.com/2011/11/irap-2012-experimenters-script1.pdf>. The IRAP computer software program itself can be downloaded for free at: <http://irapresearch.org/downloads-and-training/>). Instructions outlined to participants that on each trial one of two types of pictures would appear at the top of the computer screen along with another picture in the centre of the screen. Participants were also told that the response options “Like” and “Different” would appear at the bottom of the screen, and that they would be required to choose one of these options on each trial by pressing either “D” or “K” . The instructions explained that the IRAP consisted of a number of different trial types (for example two squares, two triangles, or one square and one triangle in the case of stimuli set 1). Prior starting the T-IRAP all children were given the following instructions:

“Let’s do something new today, would you like to do some work on the computer? We are going to do some matching. We will see pictures that are the same and pictures that are different. When the two pictures are the same

press the “D” button, that’s this one (researcher to point to the D key). If the two pictures are different press the “K” button, that’s this one (researcher to point to the K key). So, if a square comes up here (pointing to the top picture) and another square comes up here (pointing to bottom picture) I will point here (pointing to the onscreen prompt ‘*press D for Same*’) and you should press the ‘D’ key (pointing to the letter on the keyboard) because the pictures are the same. If a picture of a square and of a triangle comes up on the screen I will point here ‘*press K for Different*’ and you should press ‘K’ (pointing to the letter on the keyboard) because they are both different. If you get it right more pictures will come up. If you get it wrong a red X will come up, but that’s ok, we can try again to get the next one right.”

Stimuli presented on-screen consisted of pictorial images as a sample stimulus presented at the top of the computer screen and a comparison stimulus presented just below the sample stimulus close to the centre of the computer screen. Response options “Like” and “Different” were presented in the bottom left and right corners respectively as per TT procedures, sample and comparison stimuli were either “Like” or “Different” and participants were required to select the correct response option.

During all trials a correct response (selecting the key that corresponded with the correct onscreen response option) was followed by the screen clearing and presentation of the next trial. An incorrect response (selecting the key corresponding to the incorrect onscreen response option) was followed by a red “X” being presented onscreen. Participants were then required to select the correct response option in order to clear the screen and proceed to the next trial. Children who did not readily identify the red “X” as signalling an incorrect response were given verbal

corrective feedback and corrected using least to most guidance. Positive reinforcement was provided in the form of social praise, edibles and/or a token economy system to promote accurate and speedy responding. Reinforcement was specifically tailored to each participants needs in order to facilitate both speed and accuracy of responding. For example, thicker schedules of reinforcement using high potency tangible reinforcers were used with the participants who showed the lowest levels of motivation. Similarly, those who were highly motivated by the use of tokens received tokens dependent speed and accuracy of responding which could be exchanged for reinforcers of their choice at the end of the task. For example the greater the speed and accuracy of responding the greater the number of tokens awarded. The schedule on which reinforcement was delivered was determined by each participant's individual level of responding and the schedules normally used with each child during standard school teaching.

Teaching coordination and distinction relations using simple non arbitrary stimuli (Shape). The stimuli used were squares and triangles, and multiple exemplars of each were used to promote generalization (Rosales, Rehfeldt, & Lovett, 2011). Multiple exemplar training has been shown to be effective by a number of studies employing a RFT frame work (Barnes-Holmes et al., 2001a, 2001b; Berens et al., 2007; Luciano, Gómez-Becerra, & Rodríguez-Valverde, 2007). During the initial trials a sample stimulus (e.g. a square) was presented at the top of the screen while the comparison stimulus (e.g. a triangle) was presented in the middle of the screen simultaneously. Also presented were the response options of "Like" and "Different", at the bottom left and right of the screen respectively (Figure 1). The position of the response options was not counter-balanced across trials as previous studies have shown that including such counter-balancing of stimuli can

impede learning progress for children with autism (Smeets, Lancioni, & Striefel, 1987), and this was also in line with the procedure used by Kilroe et al., (2011).

Participants were required to select the appropriate term (“Like” or “Different”) to describe the relationship between the sample and comparison stimuli. For example, if the sample stimulus was a square and the comparison stimulus also a square, participants were reinforced for selecting “Like”. If the sample was a triangle and the comparison a square, participants were reinforced for selecting “Different”. All children were exposed to the same T-IRAP procedure. Those in the T-IRAP first condition received it prior to any exposure to any other experimental conditions. Those in the TT first condition received it after being exposed to the TT procedures outlined. This was done to control for any possible order effects

Teaching coordination and distinction relations using non arbitrary stimuli (Category). The TT procedure used during this phase was identical to that described above and used the same multiple exemplars as the T-IRAP, samples of the stimuli are shown below in Figure 2. The T-IRAP operated identically to the manner described above with only the stimuli set differing. The stimuli used were animals and food, with multiple exemplars of each being used to promote generalization (Rosales et al., 2011). The instructions given to participants were altered accordingly to include an example of the procedure with the current stimuli.

“Let’s do something new today, would you like to do some work on the computer? We are going to do some matching. We will see pictures that are the same and pictures that are different. When the two pictures are the same press the “D” button, that’s this one (researcher to point to the D key). If the two pictures are different press the “K” button, that’s this one (researcher to point to the K key). So, if an animal comes up here (pointing to the top

picture) and another animal comes up here (pointing to bottom picture) I will point here (pointing to the onscreen prompt '*press D for Same*') and you should press the 'D' key (pointing to the letter on the keyboard) because the pictures are the same. If a picture of an animal and of some food comes up on the screen I will point here '*press K for Different*' and you should press 'K' (pointing to the letter on the keyboard) because they are both different. If you get it right more pictures will come up. If you get it wrong a red X will come up, but that's ok, we can try again to get the next one right."

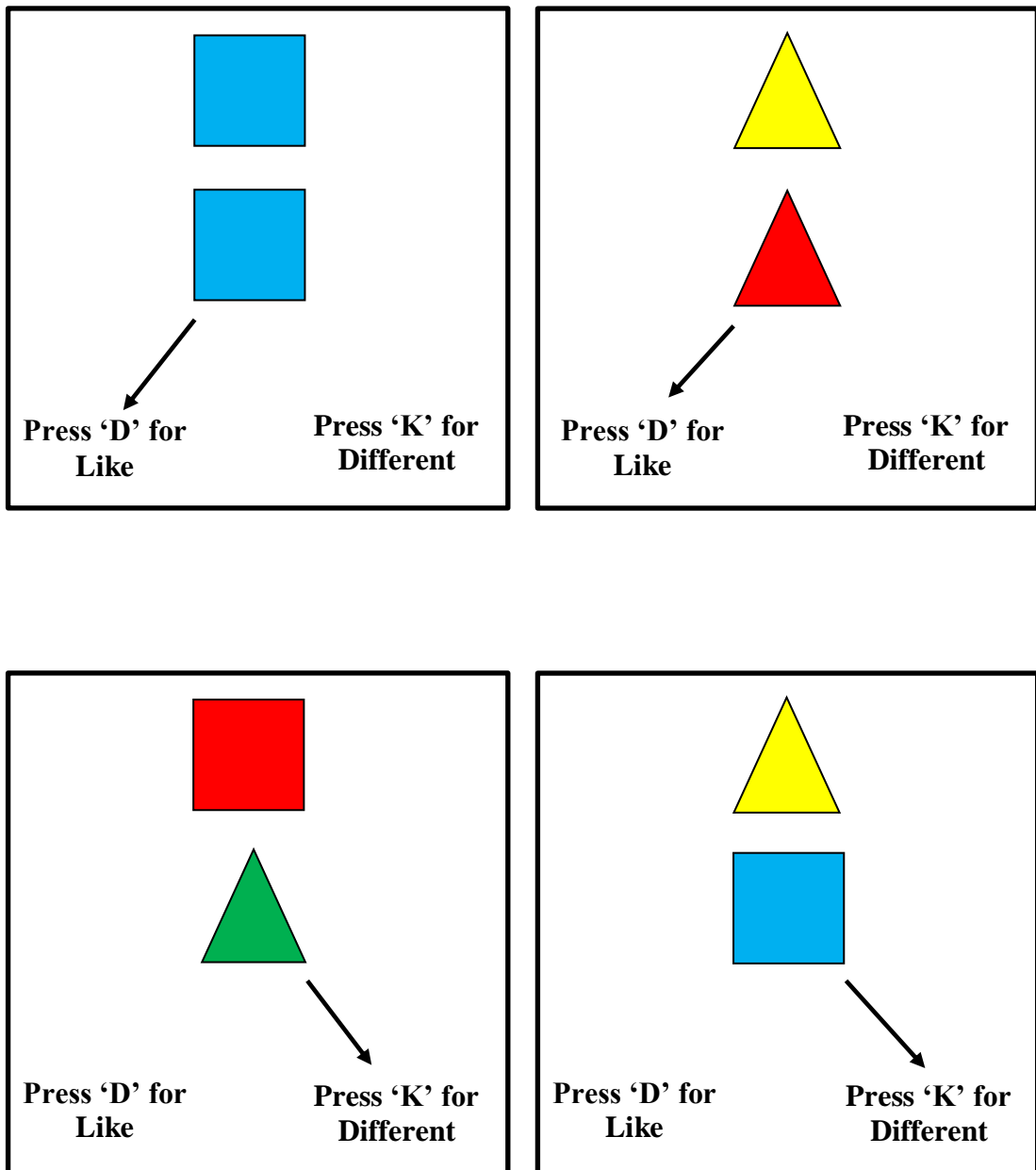


Figure 1. Examples of the four trial types used in the Square/Triangle stimulus set T-IRAP teaching procedure during “Like/Different” relational flexibility training. Arrows signal the correct response; these responses were considered consistent with pre-established patterns of responding in the general population. The arrows did not appear on the screen during instruction.

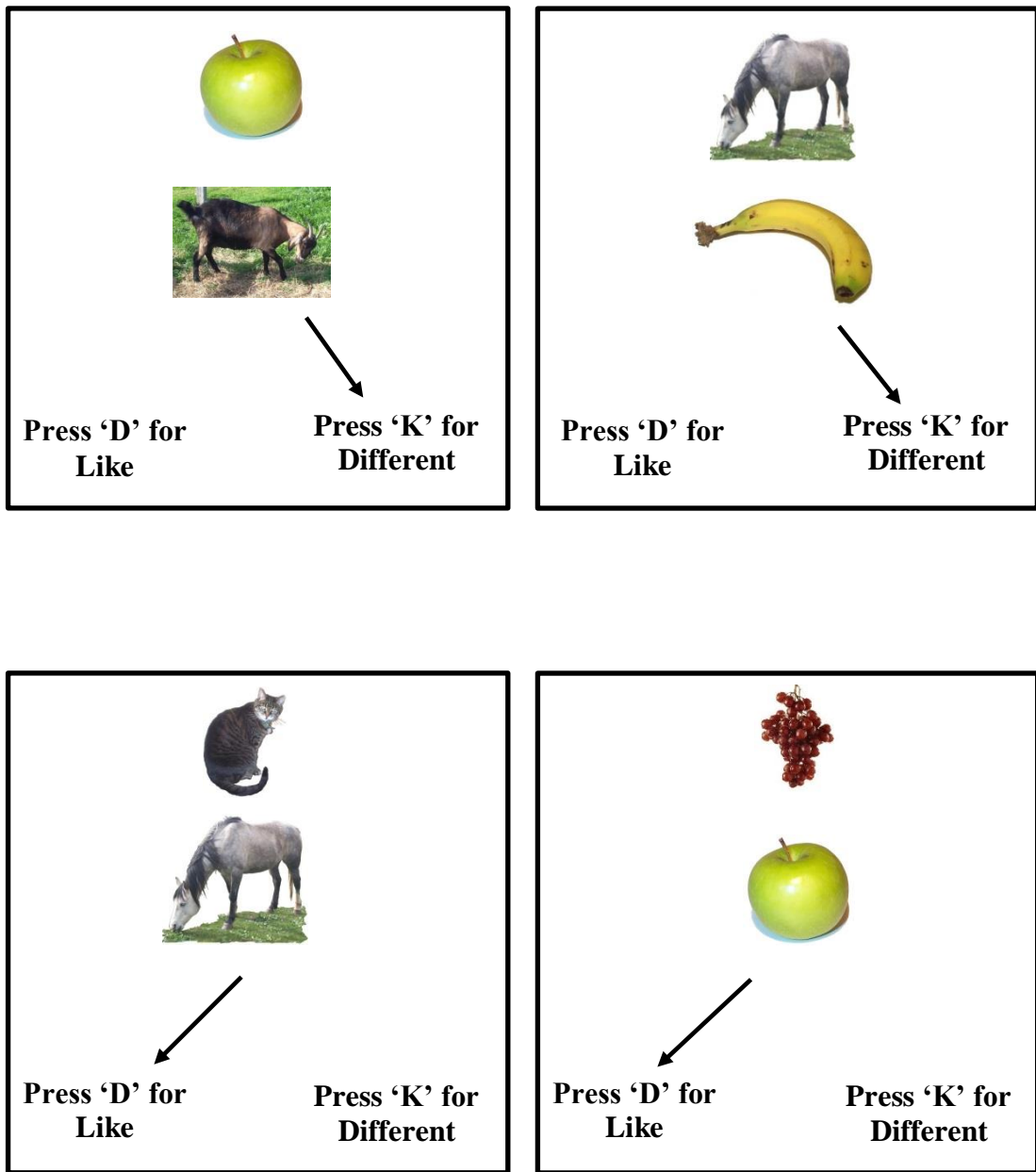


Figure 2. Examples of the four trial types used in the Animal/Food stimulus set T-IRAP teaching procedure during “Like/Different” relational flexibility training. Arrows signal the correct response; these responses were considered consistent with pre-established patterns of responding in the general population. The arrows did not appear on the screen during instruction.

Decision Protocol

A specific decision protocol based on the Comprehensive Application of Behaviour Analysis to Schooling (CABAS) schools “decision tree” was created prior to the experiment commencing (See Greer, 2002; Greer, Keohane, & Healy, 2002). The decision tree is a guide for determining where instructional programming decisions are made (Keohane & Greer, 2005). If three descending or no trend data paths are observed in the data then a decision is made to change the instructional procedures and a phase change line is drawn. If three ascending data paths occur then a decision to continue with the current instructional program is made. If there is a change in direction of the data paths within the first three data paths the decision opportunity is extended to five data paths and the same procedure applied. If there is an ascending trend continue as is, if there is a decreasing or no trend a change must be made. A mastery criterion of 80% twice or 100% on a stimuli set with a given prompt level must be reached before a prompt level could be faded. Mastery of a stimulus set was achieved when the mastery criterion of 80% twice or 100% was reached without a prompt or independently. As an alternating treatments design was used a differing trend may have been evident for each teaching condition at any one time, for example the T-IRAP may be ascending while TT shows no trend. For that reason it was decided to apply the decision protocol only to the T-IRAP data when making instructional programming decisions.

Prompt Levels and Specific Accommodations

Prompt levels were tailored to each participant to facilitate acquisition of relational concepts. For example some participants may have started the experiment with a gestural prompt while others were required to respond independently from the beginning. The decision to tailor prompt levels was made as ABA is an applied field

and all possible accommodations to facilitate the learning should be made. As such the prompt level each participant started with was based upon their previous individual learning histories and patterns. Examination of current academic programs and consultation with those who regularly work with and design programs for each individual were all taken into account when deciding what prompt levels to use in each case.

Phases with the “Independent” or “IND” label above them required the participant to respond independently to all trials in that trial block. These phases did include a least to most guidance correction procedure. Phases with the label “Gestural prompt” or “GP” label included a gestural prompt for the first 50% of trials in that block. This prompt involved the researcher silently pointing to the correct response option prior to the participant making a selection for the first 12 of each 24 trials in a given trial block. The remaining 12 trials in these blocks required the participant to respond independently or without any prompt as per the “Independent” phases.

Phases with the label “Colour Prompt” provided an extra stimulus prompt by making the relations between stimuli more salient. This was done by assigning a unique colour to each category of stimuli, for example all squares were coloured blue while all triangles were coloured red. This increased the salience of the stimuli as each time two “Like” stimuli appeared not only were they related by being the same shape they were also the same colour. Similarly, each time two “Different” stimuli were presented they differed with respect to colour as well as shape. For the animal/food stimulus sets it was not possible to change to colour of the actual picture, so the backgrounds of the stimuli was coloured in the same manner as described above.

Blocked-trials presentation. It has been demonstrated that some children with autism may struggle to develop conditional discrimination repertoires (Slocum, Miller, & Tiger, 2012). It has been shown that arranging stimuli in “blocked-trials”, where the same stimulus is presented repeatedly across trials as opposed to randomly alternating the target across trials, may aid in the development of conditional discrimination repertoires (Slocum et al., 2012; Saunders & Spradlin, 1989; 1990; 1993). The current study employed a similar strategy to aid those who did not readily acquire the relations being taught. In this case two participants, Andrew and Evan, both received blocked-trials. This procedure was TT only and involved presenting the same target relation repeatedly across trials rather than randomly presenting both. Participants received 6 trials where the target relation was “Like” followed by six “Different”, this was repeated a second time to give a total of 24 trials per block. The order in which these blocks were presented was counterbalanced across trial blocks. Correct and incorrect responses were recorded using pen and paper in the same way as all other TT procedures, duration was recorded using a timer and a least to most guidance correction procedure was used.

Results

IQ Assessment Data

Table 1 shows the pre-test standard and raw scores for each of the participants on both the K-BIT and the PPVT-IV. It should be noted that a score of 40 on the K-BIT represents the lowest possible score. The K-BIT contains two subtests, vocabulary and matrices, participants receive an individual raw score for both, these are converted to standard scores and summed. The sum of the subtest standard scores is then converted to a composite standard score, all sum of subtest scores below 90 receive a composite standard score of 40. Thus it is possible that two participants who perform vastly different receive the same standard score of 40, for example Evan and Steven, both 12 years old, received a standard score of 40, yet had raw scores of 5 and 21 respectively. A similar problem occurs at a within subjects level when comparing pre and post intervention IQ standard scores, a participant may receive a pre-test raw score of 5 and a post-test of 21, demonstrating a substantial gain yet receive a standard score of 40 on both occasions. A standard score of 20 on the PPVT-IV also represents the lowest possible score. PPVT-IV raw scores are transformed directly into standard scores, no composite scores are calculated, and as such scores of 20 represent extremely low scores, placing those who achieve it in the lowest 0.1% of the population. This is an extremely low score and may be representative of the current population having extremely low verbal ability, it may also be related to a lack of sensitivity of the ability tests used an issue discussed at length in the discussion section of the paper.

Although useful for placing the participants ability in relation to the wider population standards scores may not be sensitive enough at a within subjects levels for the purpose of the current study. Standards scores appear to be less sensitive to

change in the lower extreme of scores within which the current participants lie. As such raw scores may be a better indicator of individual gains as they may be more sensitive to subtle changes in ability. For that reason the current investigations considered raw scores in addition to standard scores when conducting data analysis, as the former may be more sensitive to subtle changes in specific repertoires rather than a general large impact which the composite and standard scores score might be expected to indicate.

Table 1

Standard (S) and Raw (R) Scores for Each Participant on the K-BIT and PPVT-IV

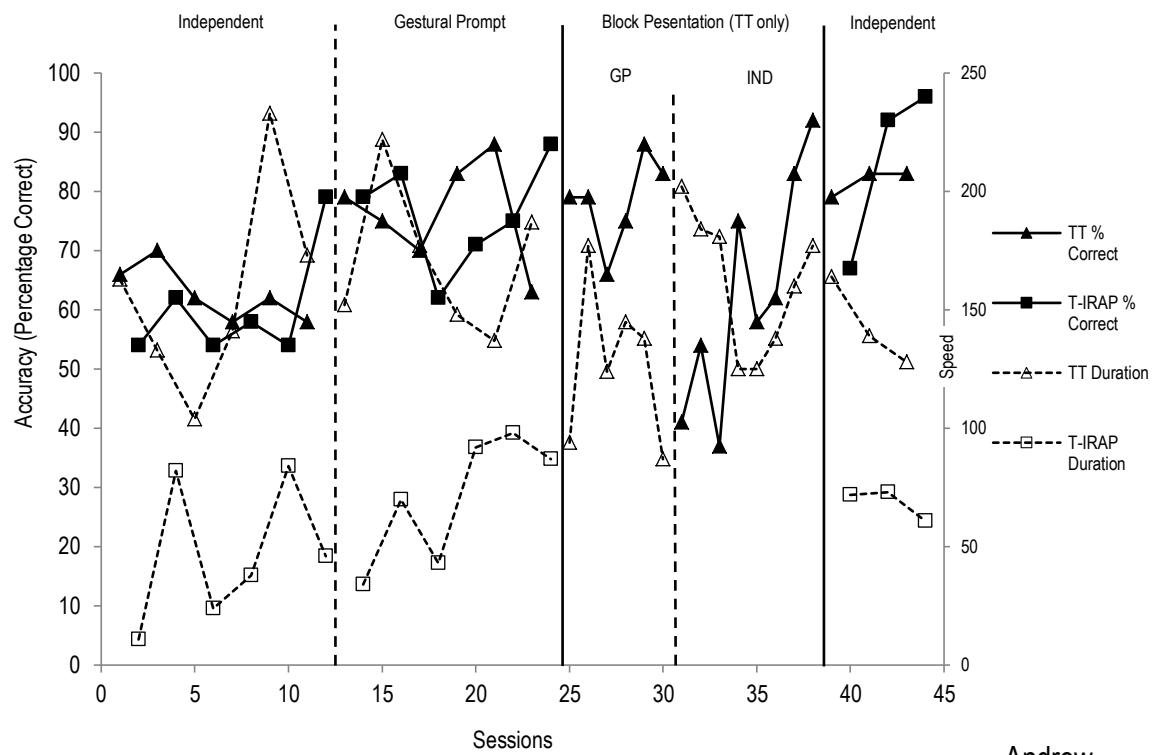
	K-BIT (S)	K-BIT(R)	PPVT-IV(S)	PPVT-IV (R)
Ann	40	10	34	34
Andrew	40	18	23	46
Steven	40	21	20	43
Kevin	62	31	50	48
Evan	40	5	20	22

Non-arbitrary Relations (Like/Different)

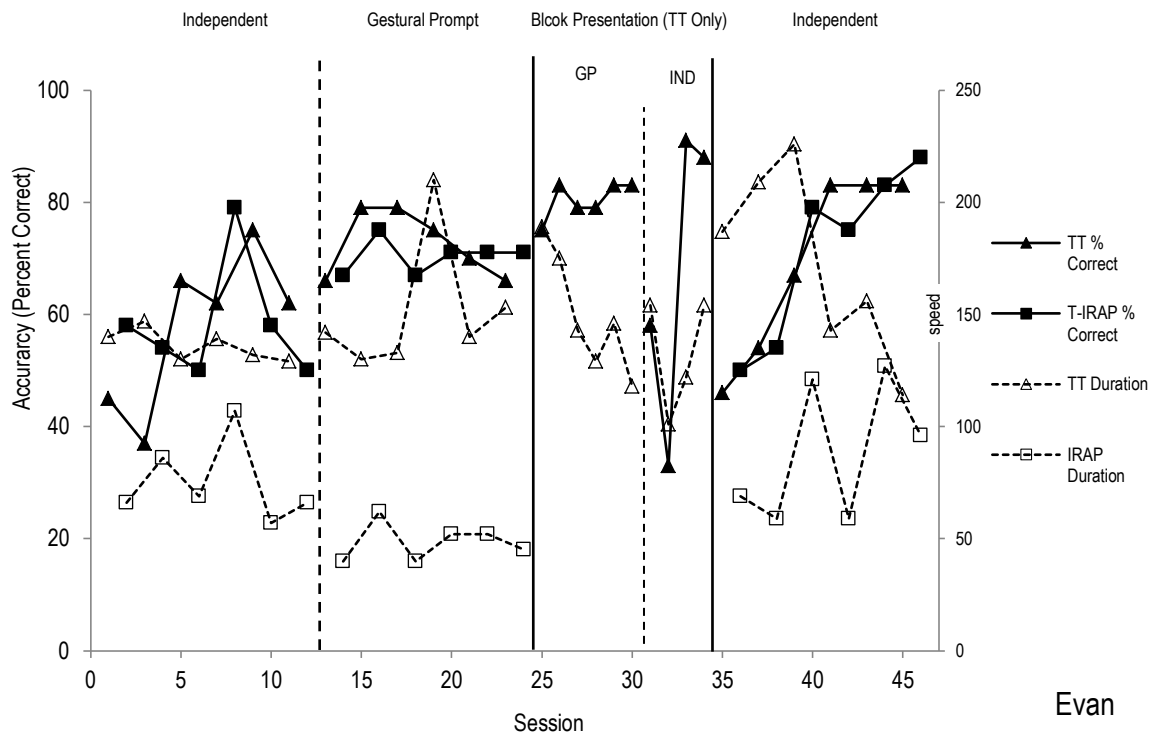
Stimulus set 1 (Square/Triangle). An alternating treatments design across five participants was used following participant completion of pre-test IQ assessments. The data for accuracy and for the speed of responding (duration of trial-blocks) using both TT and T-IRAP procedures with five participants are presented below in Figures 3 to 5. Accuracy data are represented using an unbroken line between data points and speed of responding is represented using a broken line. Accuracy was measured as percentage correct for each trial-block and data are scaled

on the primary y-axis. Duration of trial-blocks was recorded in seconds and data are scaled on the secondary y-axis (vertical axis on the right). T-IRAP data points are shown as squares while TT data points are shown as triangles. Prompts were introduced as necessary for each participant and prompt procedures were consistent across both TT and T-IRAP procedures. The introduction of each of these prompts is signalled by a vertical dashed phase-change line. The prompt introduced is indicated in the condition label above the relevant phase. More significant phase-changes or interventions are denoted by a solid unbroken vertical line. The scale on primary y-axis remains constant at 0-100% across all participants, however the scale on the secondary y-axis (speed) does not and was adjusted to accommodate variability in the duration of trial-blocks for each individual participant.

The data shown in Figure 3 for Andrew represent his performance across 5 distinct phases. Each phase represents a different prompt level or other variation within the experimental sequence. Andrew's accuracy data (percentage correct across trial-blocks) for initial independent trial-blocks during both TT and T-IRAP were a little variable at medium-high levels across five data paths with the TT accuracy data showing a downward trend and the T-IRAP data showing an upward trend that failed to demonstrate a criterion performance (80% for two consecutive data points or 100% once) . Whereas the level of accuracy data was roughly similar in TT and T-IRAP conditions, the level of speed data were markedly different for both conditions and speed of responding was greater during T-IRAP teaching compared with TT teaching in initial independent trials; the speed/duration data in both conditions were variable with no trend evident.



Andrew



Evan

Figure 3. Table-Top and T-IRAP data for Andrew (top) and Evan (bottom) showing accuracy and speed of responding (duration) data for Like/Different non-arbitrary relational responding with stimuli that were varied along one physical dimension (i.e., shape: Square/Triangle)

The accuracy data for Andrew when gestural prompt (GP) was provided were variable at medium-high levels for both teaching conditions but Andrew failed to achieve a criterion performance. The level of duration data was lower for T-IRAP teaching, indicating greater speed of responding in this condition compared to TT teaching. A subsequent TT procedure using a gestural prompt and block-trial presentation which kept location of stimuli constant (Slocum et al., 2012) resulted in Andrew achieving a criterion accuracy performance; however duration data were variable and showed no trend. The gestural prompt was faded and Andrew's data showed a criterion accuracy performance in the TT condition with independent responding and a block-trial procedure, following which Andrew achieved a criterion accuracy performance when re-exposed to both TT and T-IRAP teaching with independent trials and no modifications. The level of speed data was again notably different for T-IRAP compared to TT conditions, and responding was faster in the former rather than the latter trials. Speed of responding was variable throughout TT and T-IRAP teaching conditions. A notable difference in level between TT and T-IRAP for speed of responding is evident, with the T-IRAP producing shorter durations overall. Little difference in level of accuracy data is evident between TT and T-IRAP across all five phases.

Evan's data are shown in Figure 3 they represent his performance across 5 distinct phases. Each phase represents a different prompt level or other variation within the experimental sequence. Evan's accuracy data for initial independent trial-blocks during both TT and T-IRAP were moderately variable at medium-high levels across five data paths with the TT accuracy data showing no trend and the T-IRAP data showing an upward trend that failed to demonstrate a criterion. The level of speed data were markedly different for both conditions and speed of responding was

greater during T-IRAP teaching compared with TT teaching in initial independent trials. Whereas the speed/duration data are stable in the TT condition, they are slightly variable during T-IRAP, with no trend evident in either condition.

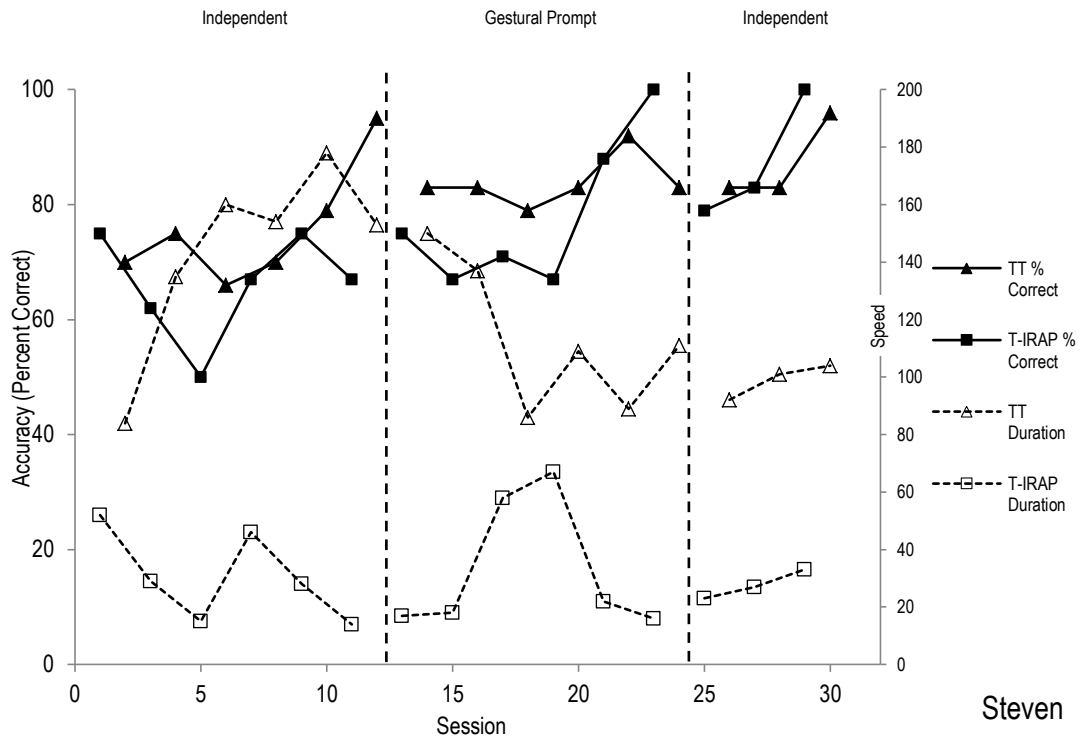
The accuracy data for Evan when gestural prompt (GP) was provided were slightly variable at medium-high levels for both teaching conditions but Evan failed to achieve a criterion performance. The level of duration data was markedly lower for T-IRAP teaching, indicating greater speed of responding in this condition compared to TT teaching. A subsequent TT procedure using a gestural prompt and block-trial presentation which kept location of stimuli constant (Slocum et al., 2012) resulted in Evan achieving a criterion accuracy performance; however duration data were variable and showed downward trend. The gestural prompt was faded and Evan's data showed a criterion accuracy performance in the TT condition with independent responding and a block-trial procedure, following which Evan achieved a criterion accuracy performance when re-exposed to both TT and T-IRAP teaching with independent trials and no modifications. The level of speed data was again notably different for T-IRAP compared to TT conditions, and responding was faster in the former rather than the latter. Variability was evident in speed data for both conditions with TT showing a downward trend, and T-IRAP an ascending trend.

Evans data as a whole show that speed of responding was variable throughout TT and T-IRAP teaching conditions, with a notable difference in level between TT and T-IRAP for speed of responding, with the T-IRAP producing shorter durations. Little difference in level of accuracy data is evident between TT and T-IRAP across all five phases.

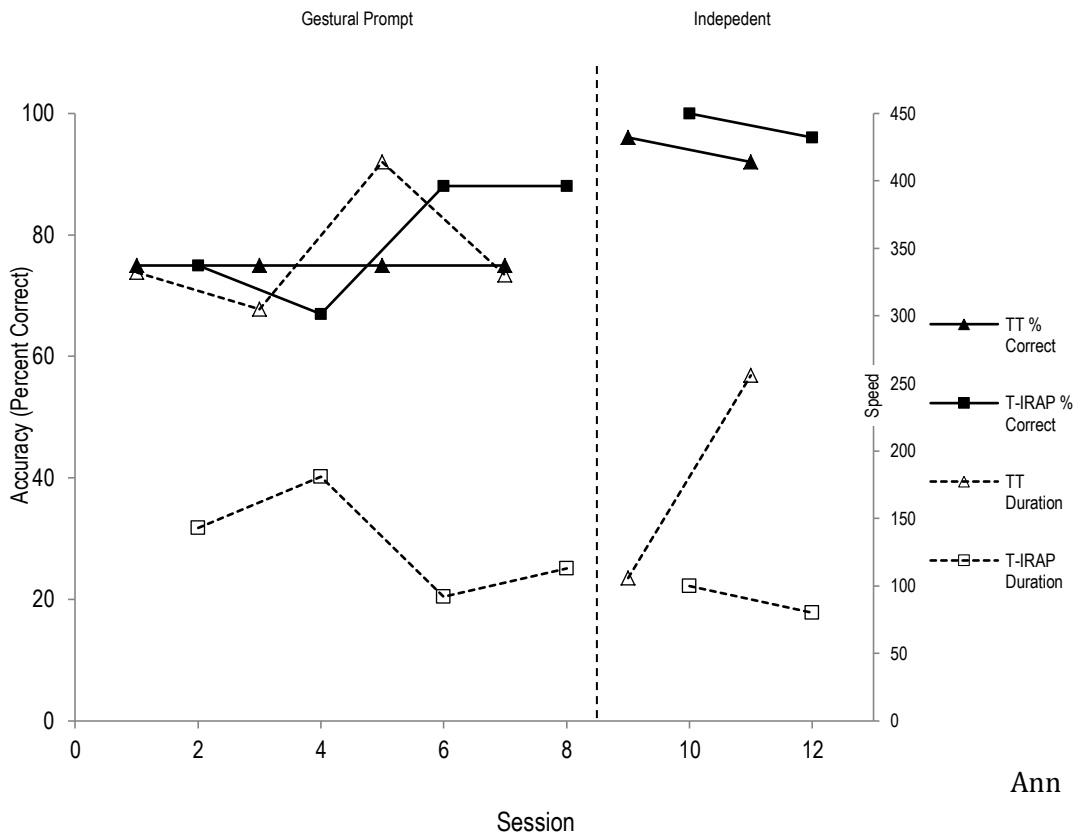
The data for Steven are shown in Figure 4, which represent his performance across 3 distinct phases. Steven's accuracy data for initial independent trial-blocks

during both TT and T-IRAP were moderately variable at medium-high levels across five data paths with the TT accuracy data showing a slight upward trend and the T-IRAP data showing no trend. Steven failed to demonstrate a criterion performance at this prompt level. While the level of accuracy data was roughly similar in TT and T-IRAP conditions, the level of speed data were markedly different for both conditions. Speed of responding was greater during T-IRAP teaching compared with TT teaching in initial independent trials; the speed/duration data in TT were variable with an upwards tend (slower responding) while T-IRAP data were more stable showing no trend.

The accuracy data for Steven when a gestural prompt (GP) was provided were slightly variable at high levels for both teaching conditions, Steven achieved a criterion performance after 5 data paths. The level of duration data was lower for T-IRAP teaching, indicating greater speed of responding in this condition compared to TT teaching, data were highly variable for both teaching conditions, with TT showing an downward trend and T-IRAP showing no trend. Following this Steven progressed to independent trials with no modifications and achieved a criterion accuracy performance after two data paths, resulting in a criterion accuracy performance for the stimulus set. The level of speed data was again notably different for T-IRAP compared to TT conditions, and again responding was faster in the former rather than the latter trials. Speed of responding was moderately variable through-out TT and T-IRAP teaching conditions. Taking Stevens data as a whole it is evident that T-IRAP produced shorter durations overall. Little difference in level of accuracy data is evident between TT and T-IRAP across all 3 phases.



Steven



Ann

Figure 4. Table-Top and T-IRAP data for Stephen (top) and Ann (bottom) showing accuracy and speed of responding (duration) data for Like/Different non-arbitrary relational responding with stimuli that were varied along one physical dimension (i.e., shape: Square/Triangle)

The data shown in Figure 4 for Ann represents her performance across 2 distinct phases. Each phase represents a different prompt level or other variation within the experimental sequence. Ann's accuracy data for initial gestural prompt trial-blocks during both TT and T-IRAP were slightly variable at high levels across three data paths with the TT accuracy data showing no trend and the T-IRAP data showing slight upward trend. Ann successfully demonstrated a criterion performance at this prompt level. The level of speed data were markedly different for both conditions yet again. Speed of responding was greater during T-IRAP teaching compared with TT teaching in initial gestural prompt trials; the speed/duration data in both conditions were slightly variable with no trend evident.

Subsequently Ann achieved a criterion accuracy performance with independent trials and no modifications after one data path meeting the mastery criterion for the stimuli set. The level of speed data show that T-IRAP was again lower compared to TT, again demonstrating faster responding in the former rather than the latter. Overall speed of responding was slightly variable for both teaching conditions, with a notable difference in level between TT and T-IRAP for speed of responding, with the T-IRAP producing shorter durations yet again.

Kevin's data are shown in Figure 5, representing his performance across a single phase. Kevin's data accuracy data for independent trial-blocks during both TT and T-IRAP were highly stable at very high levels across a single data path with no trend evident for either condition. Kevin demonstrated a criterion performance at this prompt level and achieved the mastery criterion within two data points. Both the level of accuracy and speed data were roughly similar in TT and T-IRAP conditions, with T-IRAP generating faster performances.

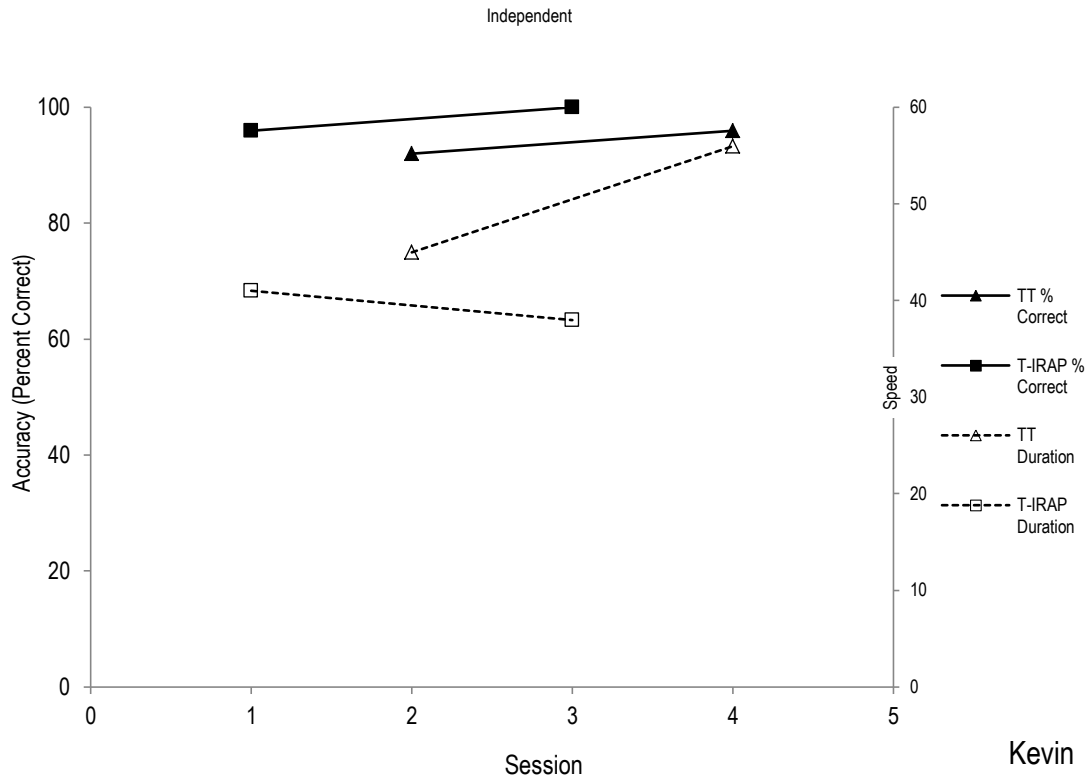


Figure 5. Table-Top and T-IRAP data for Kevin showing accuracy and speed of responding (duration) data for Like/Different non-arbitrary relational responding with stimuli that were varied along one physical dimension (i.e., shape: Square/Triangle)

Stimulus set 2 (Animal/Food). Following mastery of the triangle/square stimulus set Ann, Steven and Kevin moved onto a more complex non-arbitrary stimulus set, animals/food. Stimulus set 2 aimed to teach coordination and distinction using stimuli that varied in respect of which class of the item the stimulus belongs to, either animals or food. The experimental procedure was identical to that used in stimulus set 1.

The data for accuracy and for the speed of responding (duration of trial-blocks) using both TT and T-IRAP procedures with three participants are presented below in Figures 6 and 7. Accuracy data are represented using an unbroken line between data points and speed of responding is represented using a broken line.

Accuracy was measured as percentage correct for each trial-block and data are scaled on the primary y-axis. Duration of trial-blocks was recorded in seconds and data are scaled on the secondary y-axis (vertical axis on the right of the graph). T-IRAP data points are shown as squares while TT data points are shown as triangles on the graph. Prompts were introduced as necessary for each participant and prompt procedures were consistent across both TT and T-IRAP procedures. The introduction of each of these prompts is signalled by a vertical dashed phase-change line graph. The prompt introduced is indicated in the condition label above the relevant phase. More significant phase-changes or interventions are denoted by a solid unbroken vertical line. The scale on primary y-axis remains constant at 0-100% across all participants, however the scale on the secondary y-axis (speed) does not and was adjusted to accommodate variability in the duration of trial-blocks for each individual participant.

The data for Steven are shown in Figure 6, it represents his performance across 3 distinct phases. Each phase represents a different prompt level or other variation within the experimental sequence. Steven's accuracy data for initial independent trial-blocks during both TT and T-IRAP were moderately variable at medium-high levels across five data paths with the TT accuracy data showing a slight upward trend and the T-IRAP data showing no trend. Steven failed to demonstrate a criterion performance at this prompt level. Whereas the level of accuracy data was roughly similar in TT and T-IRAP conditions, the level of speed data were markedly different for both conditions. Speed of responding was greater during T-IRAP teaching compared with TT teaching in initial independent trials; the speed/duration data in TT were variable with an upwards trend (slower responding) while T-IRAP data were more stable showing no trend.

Figure 6 shows the accuracy data for Steven when a gestural prompt (GP) was provided, data for T-IRAP were relatively stable at high levels with an upwards trend, while data were slightly variable at high levels for TT with no trend evident. Steven achieved a criterion performance after 5 data paths. The level of speed data was markedly lower and more stable for T-IRAP teaching compared to TT teaching, data were variable for TT with no trend evident. Subsequently Steven progressed to independent trials with no modifications and achieved a criterion accuracy performance after two data points, resulting in a criterion accuracy performance for the stimulus set. The level of speed data was again notably lower for T-IRAP compared to TT conditions, speed of responding was moderately variable throughout TT, while T-IRAP was relatively stable.

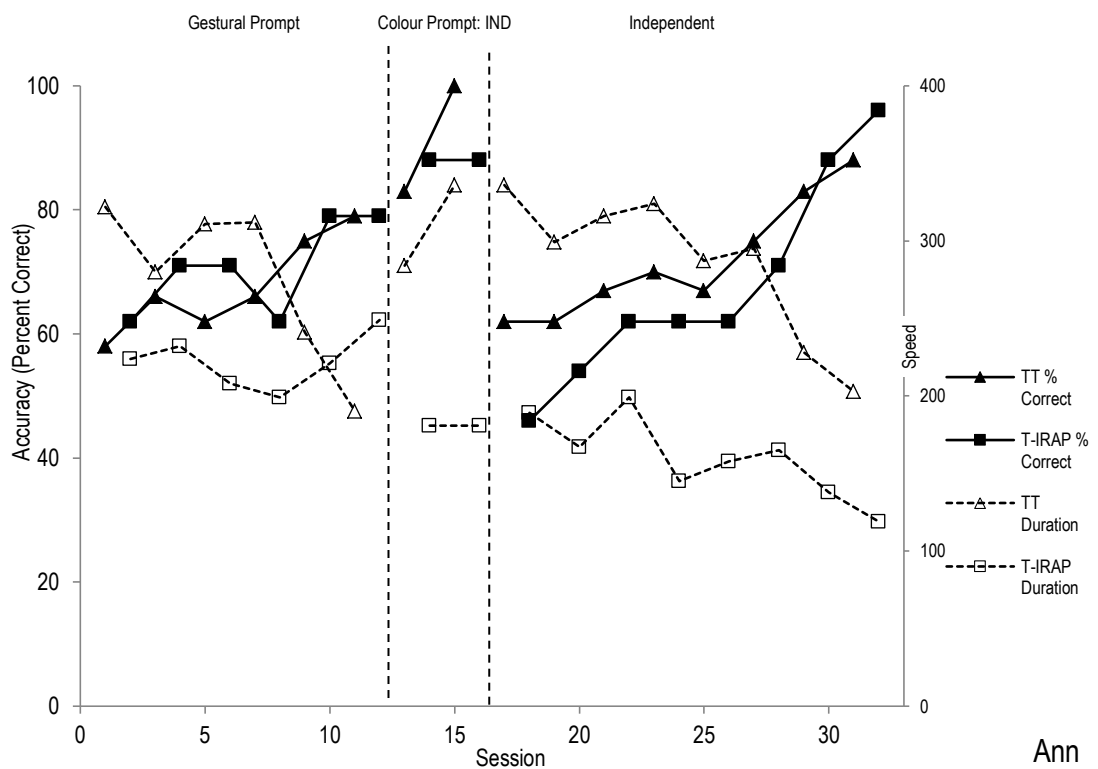
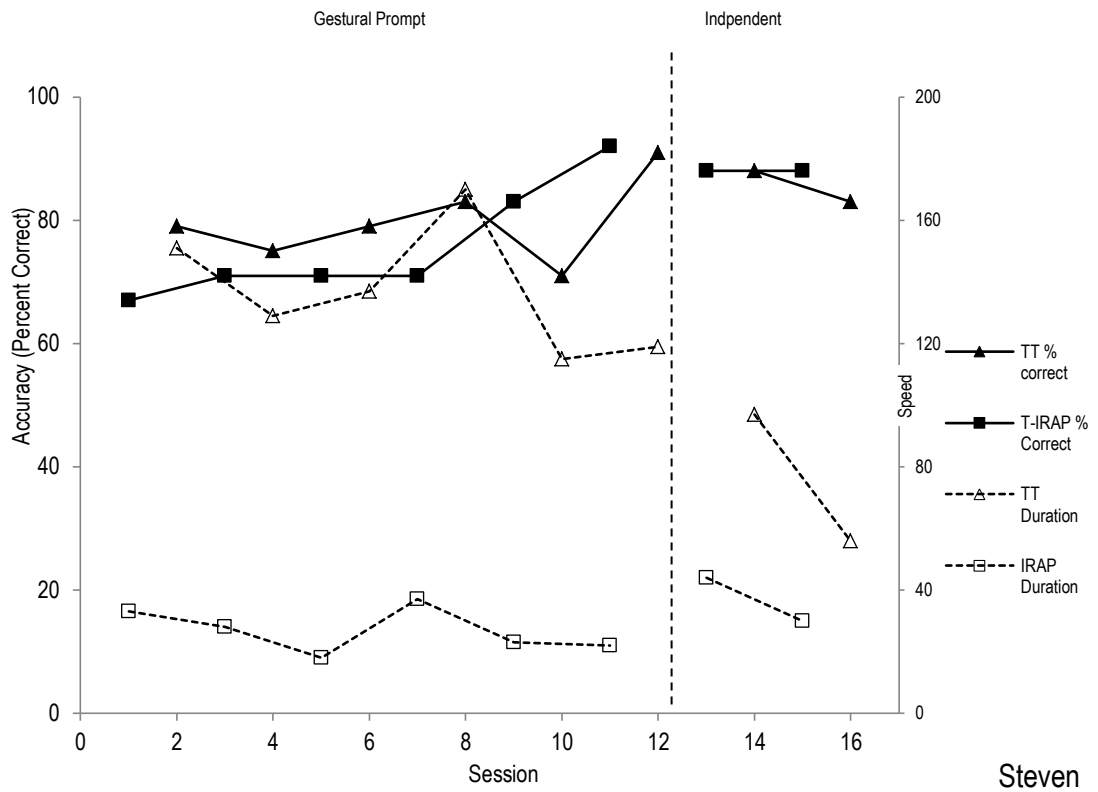


Figure 6. Table-Top and T-IRAP data for Stephen (top) and Ann (bottom) showing accuracy and speed of responding (duration) data for Like/Different non-arbitrary relational responding with stimuli that were varied in relation to class membership (i.e., class: Animal/Food)

The data shown in Figure 6 for Ann represents her performance across 3 distinct phases. Ann's accuracy data for initial gestural prompt trial-blocks during both TT and T-IRAP were moderately variable at medium-high levels across five data paths with the TT data showing a slight upward trend and the T-IRAP data showing no trend. Ann did not demonstrate a criterion performance with the gestural prompt. The level of speed data were slightly different for both conditions yet again. Speed of responding was slightly faster during T-IRAP teaching compared with TT teaching in initial gestural prompt trials; the speed data in TT were slightly variable with a downward trend evident, with T-IRAP being more stable showing no trend.

A subsequent stimulus prompt teaching procedure where the salience of the stimuli was increased by colour coding the stimuli was implemented (hereafter referred to as a colour prompt). Ann achieved a criterion accuracy performance with this colour prompt procedure after one data path. Following which she achieved a criterion accuracy performance with independent trials and no modifications, meeting the mastery criterion for the stimulus set. The level of accuracy data for both TT and T-IRAP was roughly similar and showed a steady upward trend across seven data paths. The level of speed data show that T-IRAP was again lower compared to TT, with both showing a slight downward trend some to slight-moderate variability. Overall speed of responding was moderately variable for both teaching conditions, with a notable difference in level between TT and T-IRAP for speed of responding, with the T-IRAP producing shorter durations yet again.

Kevin's data are shown in Figure 7, representing his performance across a single phase. Kevin's data accuracy data for independent trial-blocks during both TT and T-IRAP were highly stable at very high levels across a two data paths with no trend evident for TT and an upwards trend for T-IRAP. Kevin demonstrated a

criterion performance with independent trials and no modifications, achieving the mastery criterion for the stimulus set within two data paths. The level of accuracy data was roughly similar in TT and T-IRAP conditions at very high levels, with TT generating marginally higher scores. Speed data were slightly variable in T-IRAP and TT conditions, with the former showing no trend and the later showing a downward trend. T-IRAP generated faster performances compared to TT.

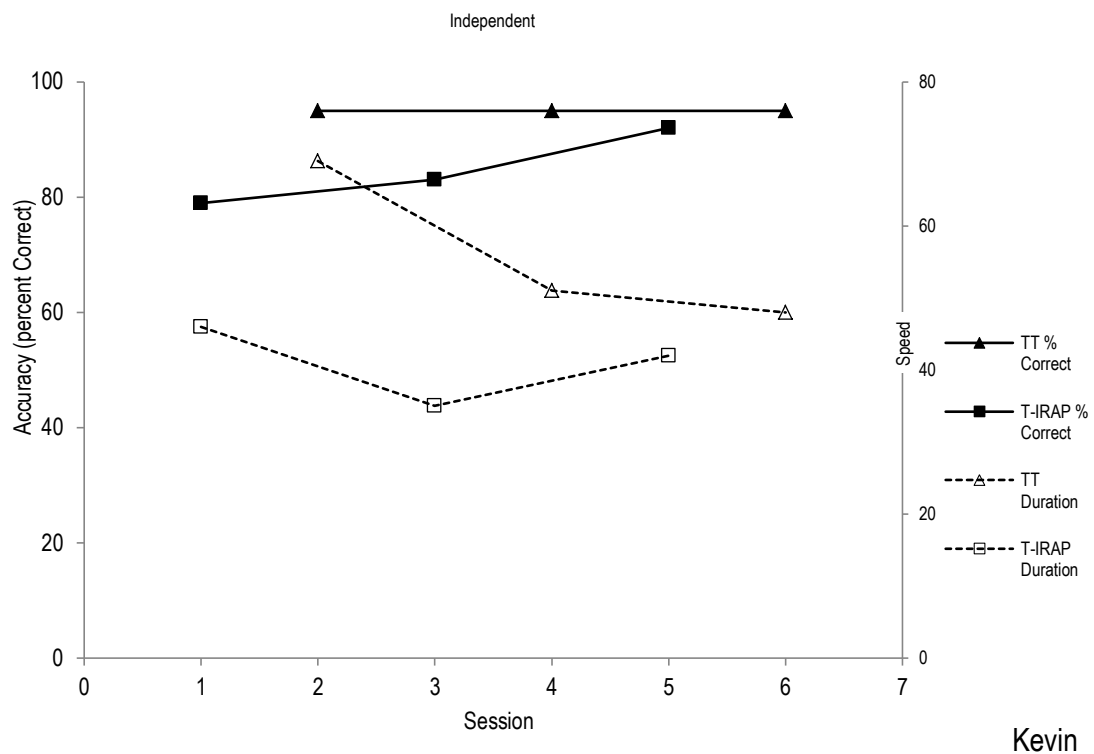


Figure 7. Table-Top and T-IRAP data for Kevin showing accuracy and speed of responding (duration) data for Like/Different non-arbitrary relational responding with stimuli that varied in relation to class membership (i.e., class: Animal/Food).

Mean Accuracies and Response Times for T-IRAP and Table Top

Mean response times and accuracies for all sessions to date by participants and stimulus set are presented in Table 2. Mean response times were calculated by summing the duration of each trial block for each participants and dividing by the total number of trial blocks. Table 2 shows that mean T-IRAP scores are

approximately 2.6 times shorter on average compared to TT. With the largest difference evident for Steven during Square/Triangle trials, with T-IRAP being 4.8 times faster, Kevin’s data show the smallest difference with T-IRAP being 1.3 times faster than TT. Mean accuracies are highly similar across both TT and T-IRAP for all participants and stimulus sets.

Table 2

Mean Durations for all Sessions to Date by Participant and Stimulus Set

	T-IRAP Mean Duration (Seconds)	Table-Top Mean Duration (Seconds)	T-IRAP Mean Accuracy (Percent)	Table-Top Mean Accuracy (Percent)
<i><u>Stimulus set 1</u></i>				
Andrew	48	166.8	70.72%	70.72%
Evan	67.8	138.1	66.67%	69.64%
Steven	30.7	148.5	75.07%	81.33%
Kevin	39.5	50.5	98%	94%
Ann	124.7	312.6	85.67%	81.33%
<i><u>Stimulus set 2</u></i>				
Steven	29.4	121.8	78.88%	81.13%
Kevin	41	56	84.6%	95%
Ann	185.9	285.3	71.31%	72.69%

A paired samples t-test was conducted to compare average duration in TT and T-IRAP condition. It was found that for all trials to date there was a significant difference in average duration for T-IRAP ($M=64.4$, $SD=28.22$) and TT ($M=155.03$, $SD=91.19$); $t(7) = 3.563$, $p=0.009$. It was also found that T-IRAP was significantly faster during both stimulus sets individually. In the Square/Triangle stimulus set average duration for T-IRAP ($M=64.07$, $SD=34.19$) was significantly different from

TT ($M=155.45$, $SD=87.09$); $t(4) = -3.56$, $p=0.024$). In the Animal/Food stimulus set average duration for T-IRAP ($M=85.43$, $SD=87.22$) was not significantly different from TT ($M=154.33$, $SD=118.04$); $t(2) = -2.55$, $p= 0.126$).

Paired samples t-test examining mean accuracy scores for all session to date between T-IRAP ($M=78.97$, $SD=10.13$) and TT ($M=88.18$, $SD=44.59$) indicated no significant difference; $t(7) = .561$, $p=0.592$. Similarly, for stimulus set 1 mean accuracy differences between T-IRAP ($M=79.4$, $SD=12.52$), and TT ($M=79.03$, $SD=10.5$) were not significant; $t(4) = .192$, $p=.857$. nor were there differences between T-IRAP ($M=78.26$, $SD=6.66$) and TT ($M=82.93$, $SD=11.27$) mean accuracy scores during stimulus set 2 significantly different; $t(2) = -1.627$, $p=.245$. These results show that accuracy data for both conditions was roughly similar throughout all stimulus sets to date.

Summary

The data show it was possible to teach five children with autism and low ability scores to use a T-IRAP teaching procedure to learn Like/Different flexible relational responding with non-arbitrary stimuli. Data also suggest that the T-IRAP may be a more efficient method of teaching rapid and fluent responding compared to TT. These findings are in line with those of Kilroe et al (2011) which found T-IRAP to produce higher levels of accuracy at greater speed as compared to TT. The current study did not find T-IRAP to produce significantly greater levels of accuracy than TT.

TT and T-IRAP were compared in an alternating treatments design, data across all participants showed T-IRAP to be more effective in terms of speed, while there was little difference in accuracy. This suggests that the T-IRAP may be useful as a supplementary teaching tool as it allows children to work independently and at a

high rate of efficiency, which would allow students build additional fluency in responding. Fluency has been noted to be beneficial in terms of acquisition, retention and resilience to distraction (Binder, 1996), also RFT suggests that fluency in relational responding may be an important part of intelligence (Hayes, 2011). Taken together, the T-IRAP represents a promising method of building fluency in relational responding which may in turn prove to be beneficial in terms of acquisition, retention and general cognitive ability.

The data show a range of competencies between participants with three of the five participants achieving mastery on two stimulus sets, while two only achieved mastery on one stimulus sets. It was also necessary to implement a number of specific alterations to the teaching procedure to remediate difficulties some participants experienced. With stimulus set 1 two participants required a block-trial presentation procedure after failing to reach criterion following the introduction of a gestural prompt, the block-trial presentation procedure proved to be effective. Two participants required a gestural prompt to acquire the response, while one achieved mastery independently without the introduction of any additional prompts. Of the three participants who progressed to stimulus set 2, one achieved mastery without any alterations, one required a gestural prompt, and the last required a colour prompt procedure after failing to meet criterion with a gestural prompt. The colour prompt procedure proved highly effective. Each of the prompts used were effective at remediating the difficulties each participants was experiencing and represent possible interventions for future use in similar tasks.

It must be noted that speed data were quite variable for a number of participants and in most cases be accounted for by environmental distractions. Taking Andrew as an illustrative example his speed data (Figure 3) are moderately-

highly variable throughout and is largely due to environmental distractions, Andrew was easily distracted when his classmates had access to one of his preferred reinforcers, the class computer. Interestingly speed data for Steven for the second stimulus set show considerably more variability in the TT condition than T-IRAP, this may be due to the fluidity of stimulus presentation in the T-IRAP which may be more resistant to environmental distractors (Binder 1996).

The current study achieved its goals of teaching with autism and low IQ scores fluent and flexible relational responding using a T-IRAP, although a range of competencies are evident, all participants did achieve a level of relational flexibility. It is also clear from the data that T-IRAP was more efficient method of teaching the target relations compared to TT; however there was some variability in data for both conditions.

-Chapter 3-

Introduction

Arbitrarily Applicable Relational Responding (AARR) is the ability to identify relations between and among stimuli and events not just by their physical properties but by arbitrary contextual cues that are unrelated to the physical properties involved (Hayes et al. 2001). It has been suggested that learning to respond to non-arbitrary stimuli may be a prerequisite for learning to respond to arbitrary stimuli (Berens et al., 2007; Dymond & Barnes, 1995). As such Study 2 builds upon Study 1 by attempting to teach arbitrary flexible relational responding to one participant from Study 1 who successfully complete the both non-arbitrary stimuli sets.

Relational frame theorists have argued that AARR is a form of operant behaviour and as such can be altered through systematic intervention. Berens et al., (2007) experimentally demonstrated this with a number of participants who initially could not perform a series of problem-solving tasks involving arbitrary more than and less than relations. Multiple-exemplar training was shown to facilitate the development of arbitrary comparative relations; it was also found that training in non-arbitrary comparative relations was a potential prerequisite for successful acquisition of arbitrary comparative relations. Dymond et al. (1995) also found that non-arbitrary relational responding may be a prerequisite for more complex forms of relational responding. It was found that participants who had been exposed to a non-arbitrary pre-training procedure performed better on a novel complex relational task involving arbitrary stimuli than participants who did not receive pre-training.

A number of studies have demonstrated that AARR is susceptible to the same forms of control and manipulation as other operants (see Barnes-Holmes et al., 2001a, 2001b; Healy, Barnes-Holmes, & Smeets, 2000). For example Roche,

Barnes-Holmes, Smeets, Barnes-Holmes, & McGeady (2000) demonstrated that different contextual cues can bring about different patterns of relational responding. Similarly O'Connor et al., (2011) brought symmetry and asymmetry performances under contextual control in typically developing children and some children with autism. These studies show that AARR is functionally similar to other operant behaviours, and like other operant behaviours is inherently flexible (O'Toole et al., 2009). That is, verbally sophisticated individuals may alter their responses depending on the current context and adapt to changing contingencies.

Such studies support RFT's claims that AARR is an integral part of advanced human language and cognition. For example, it is arbitrary relations that allow humans to understand more complex constructs like the value of money, as the worth of a given coin bears no physical relation to its perceived value. Take a one euro coin, it is smaller in physical size than a fifty cent coin but it is worth more in monetary terms. Such complex arbitrary relations require a level of verbal-sophistication to comprehend to derive that fifty cent is less than one euro (Barnes-Holmes, Barnes-Holmes, & Mchugh, 2004). Such relations are arbitrary as no physical relationship exists between the referents and are common place within day to day life. It has been argued that arbitrary applicable relational responding may be brought to bear on any set of stimuli given an appropriate context (Hayes, et al., 2001; Barnes-Holmes et al.). Consider a student asked to pretend correct is incorrect and incorrect is correct, in this example regardless of the current relations that exists between the stimuli the relational functions of correct and incorrect are now applied arbitrarily. Such examples highlight the importance of AARR in everyday life and its involvement with advanced language and cognitive abilities.

It follows then that teaching AARR to those with severely limited language abilities may potentially result in exponential gains in advanced verbal ability. Study 2 aims to build upon the literature and Study 1 by teaching flexible arbitrary applicable relational responding to one child with diagnosed autism and extremely low IQ scores using both T-IRAP and TT in an alternating treatments design.

Method

Study 2 utilised arbitrary stimuli in the sense that they bore no physical relation to each other. Previous studies have demonstrated that relations can be established with arbitrary stimuli, Kilroe et al.,(2011) showed that stimuli related on an arbitrary basis that are used in educational and real-world settings can be taught using the T-IRAP format. The stimuli chosen are commonly used in educational settings to teach both fractions and percentages to primary school children. The arbitrary stimuli for the “Like/Different” relations involved teaching children to relate the percentage symbol for half (50%) to a visual graphic representation of half (●). Similarly participants were taught to relate graphic symbols representing a quarter (⊕) to its corresponding percentage symbol (25%). This was taught using the same procedure outlined in Study 1 using both Table-top and T-IRAP procedures.

Participants

One participant from Study 1 took part in Study 2; Kevin achieved the mastery criterion of Study 1 and proceeded to the current study. Kevin had been previously diagnosed with autism by a clinical psychologist independent of the current research. Clinical diagnoses were in accordance with the DSM-IV criteria (APA, 2000). Consent was sought prior to the commencement of Study 1 for the entire research program, as such no additional consent form was sent out the participant’s parents specifically for Study 2. As outlined previously Kevin is an eight year old boy with diagnosed autism in the moderate range. Kevin achieved the highest score of all during pre-test IQ assessments carried out prior to Study 1, scoring 50 and 62 on the PPVT-IV and K-BIT respectively. For a more detailed description please see the participants section of Study 1.

Settings and Materials.

Setting. Kevin received 2-3 twenty minute research sessions a week. These sessions were conducted during normal class hours in his usual class room, under standard class room conditions.

Table Top and T-IRAP. TT and T-IRAP procedures were carried out in a manner identical to Study 1, only varying in relational to the stimuli used. The stimuli used in Study 2 were images created by the researcher using the Microsoft Paint tool that is part of Microsoft's Windows operating system and were identical in both TT and T-IRAP procedures. During TT the number of correct and incorrect responses was recorded using pen and paper on a specially constructed data sheet. While data collection was automated during T-IRAP procedures.

Design

An alternating treatments design was employed during the study to compare the relative effectiveness of TT procedures and T-IRAP at teaching flexibility in relational concepts, in this case Like/Different using arbitrary stimuli, to one child with diagnosed autism. The order of the treatments was counter balanced across participants across the entire research program so that half the participants received the table top procedures first, and the other half received T-IRAP first, as such the order of the presentation for Kevin was carried over from Study 1.

Inter Observer Agreement

An independent observer who had extensive ABA experience was trained to record TT data. The observer sat at the opposite end of the table to the participant and researcher and did not interact with the participant in any way. Approximately 33% of all TT trials were observed and a trail by trial analysis was conducted. 100% agreement between researcher and the independent observer was found.

Procedure

Experimental sequence. Following successful completion of the stimuli sets included in Study 1, Kevin progressed directly into Study 2 which used arbitrary stimuli as opposed to the non-arbitrary stimuli used in Study 1. The format of stimulus presentation in both TT and T-IRAP remains unchanged from Study 1, with only the stimuli changing. The stimuli used in Study 2 were circles and squares of different colours; these had either one or two bisecting lines drawn through them to divide the shape into either two halves or four quarters. In both cases one of the resulting sections was coloured in to produce shapes which had either one half or one quarter coloured in (\oplus , \odot). Used in conjunction with these pictorial representations were the corresponding percentage i.e., “50%” and “25%”, these were also in a multitude of colours (Figure 8). The response options “Like” and “Different” were presented in the same fashion as during Study 1 and again the position of the response options was not counter balanced across trials on either TT or T-IRAP.

Kevin was required to select the appropriate term (“Like” or “Different”) to describe the relationship between the sample and comparison stimuli. For example, if the sample stimulus was \odot and the comparison stimulus was “50%”, Kevin was required to select “Like”. Similarly, if the sample \odot and the comparison stimulus was “25%”, Kevin was required to select “Different”. Reinforcement was delivered contingent on speedy and accurate responding and again tailored to the needs of the participant. Prior to starting the T-IRAP Kevin was given the following instructions which were adapted from those used in Study 1:

“Let’s do something new today, would you like to do some work on the computer? We are going to do some matching. We will see pictures that are the same and pictures that are different. When the two pictures are the same

press the “D” button, that’s this one (researcher to point to the D key). If the two pictures are different press the “K” button, that’s this one (researcher to point to the K key). So, if a picture like this, that has this part (i.e. Half) coloured in comes up here, (pointing to the top picture) and this picture of 50% comes up here (pointing to bottom picture) I will point here (pointing to the onscreen prompt ‘*press D for Like*’) and you should press the ‘D’ key (pointing to the letter on the keyboard) because the pictures are the same. If a picture like this, that has this part (i.e. Quarter) coloured in comes up here, I will point here ‘*press K for Different*’ and you should press ‘K’ (pointing to the letter on the keyboard) because they are both different. If you get it right more pictures will come up. If you get it wrong a red X will come up, but that’s ok, we can try again to get the next one right.”

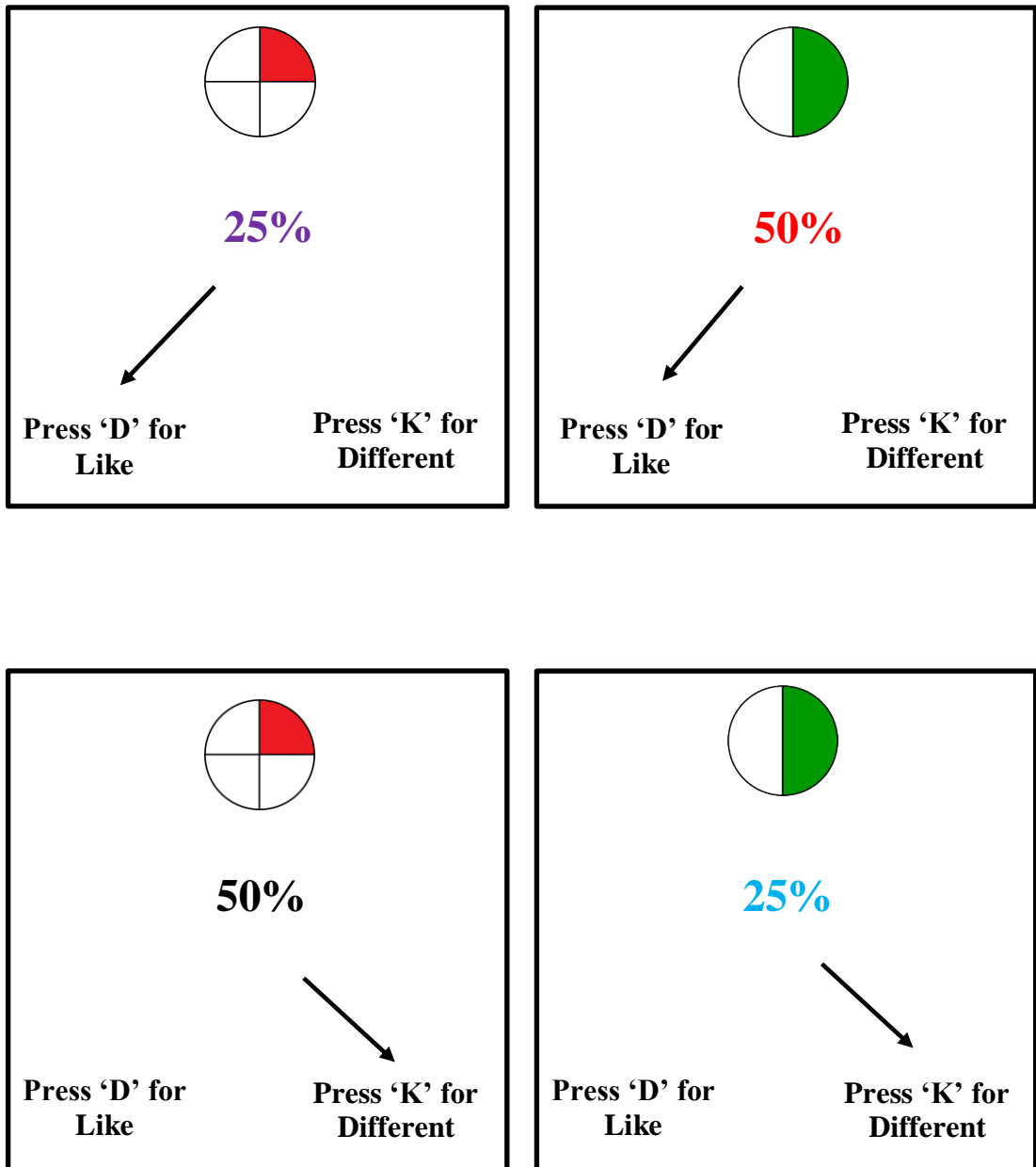


Figure 8. Examples of the four trial types used in the Arbitrary stimulus set T-IRAP teaching procedure during “Like/Different” relational flexibility training. Arrows signal the correct response; these responses were considered consistent with pre-established patterns of responding in the general population. The arrows did not appear on the screen during instruction.

Specific Accommodations

Kevin failed to meet the response criterion of 80% x 2 consecutive sessions or 100% once on the standard experimental procedure described and a program intervention was implemented in accordance with the decision protocol. It was decided to increase the salience of the stimuli by adding a colour prompt to the stimuli. The colour prompt procedure was identical to that used in Study 1. All percentage symbols for half (50%) and the graphic representations of half (◐) were green, while all percentage symbols for one quarter and graphic representations (◑) were red. This meant that when two stimuli were alike both stimuli were the same colour. Similarly if the response required was “Different” the stimuli differed in colour. This colour prompt procedure was matched across both TT and T-IRAP procedures. This essentially added a non-arbitrary component to the otherwise arbitrary stimuli to facilitate the acquisition of the arbitrary relation when the colour prompt procedure was removed.

Results

Accuracy and Speed (Like-Different Arbitrary Relations)

An alternating treatments design with one participant from Study 1 was used following participant completion of both stimulus sets in Study 1. The data for accuracy and for the speed of responding using both TT and T-IRAP procedures are presented below in Figure 9. Data are represented in the same format as Study 1.

The data shown in Figure 9 for Kevin represents his performance across 5 distinct phases. Each phase represents a different prompt level or other variation within the experimental sequence. Kevin's accuracy data for initial independent trial-blocks during both TT and T-IRAP were a quite variable at medium levels across five data paths with neither the TT nor T-IRAP showing any trend. Kevin failed to demonstrate a criterion. The level of accuracy data was roughly similar in TT and T-IRAP conditions, while the level of speed data was markedly different between both conditions. Speed of responding was faster during T-IRAP teaching compared with TT teaching in initial independent trials; the speed data in both conditions were highly variable with no trend evident.

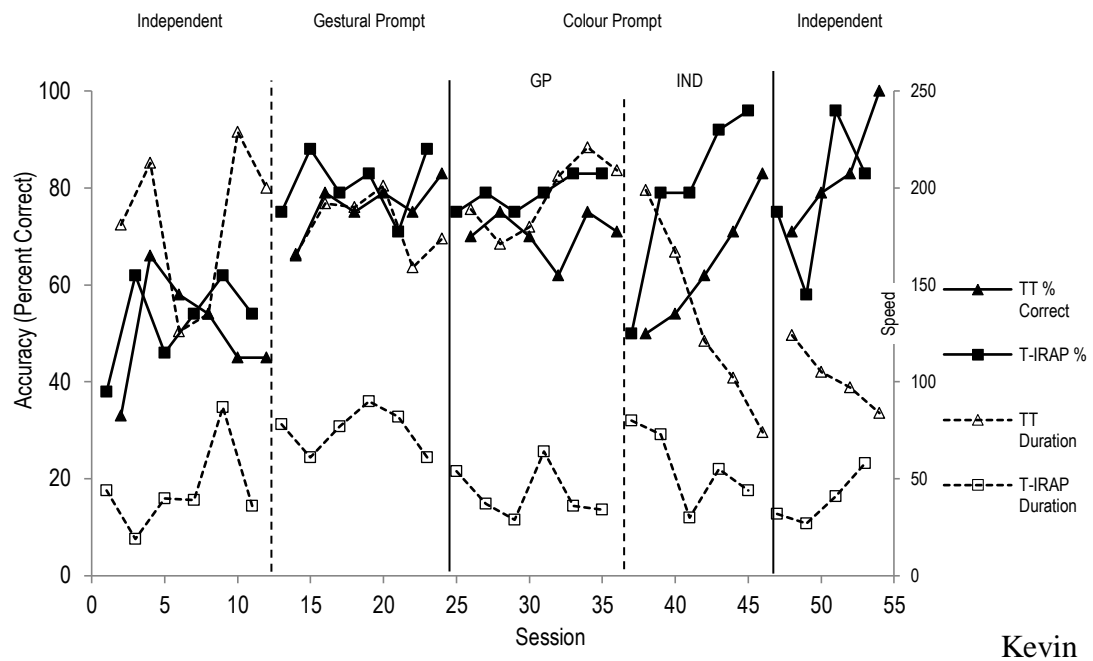


Figure 9. Table-Top and T-IRAP data for Kevin showing accuracy and speed of responding (duration) data for Like/Different arbitrary relational responding with stimuli that did not vary in relation to any formal property but arbitrarily

The accuracy data for Kevin when a gestural prompt (GP) was provided were variable at high levels for both teaching conditions but Kevin failed to achieve a criterion performance. The level of duration data was lower for T-IRAP teaching, indicating greater speed of responding in this condition compared to TT teaching. A subsequent colour prompt procedure using a gestural prompt resulted in Kevin achieving a criterion accuracy performance. Speed data for TT and T-IRAP were variable and showed no trend, a notable difference in level was again evident with T-IRAP producing faster responses. The gestural prompt was faded and Kevin's data showed a criterion accuracy performance in the colour prompt procedure with independent responding, both TT and T-IRAP data showed a definite ascending trend, with T-IRAP producing slightly higher levels. TT speed data showed a clear downward trend from high to low levels, approaching levels similar to T-IRAP data

which was slightly variable at low levels. Subsequently, Kevin achieved a criterion accuracy performance with independent trials and no modifications, meeting the mastery criterion for the stimulus set. Accuracy data for T-IRAP were quite variable, while TT showed a steady increase. The level of speed data was again notably different for T-IRAP compared to TT conditions, and responding was faster in the former rather than the latter trials. TT speed data showed a stable downward trend, while for T-IRAP an ascending trend is evident.

Overall, moderate levels of variability are evident in accuracy data with both TT and T-IRAP showing data of roughly the same level. Speed data showed a marked difference between T-IRAP and TT, with the former showing lower levels (faster responding) through-out. On average T-IRAP was 3.1 times faster than TT with average block durations of 52.1s and 163.5s respectively, with mean accuracies of 67.9% during TT and 73.4% during T-IRAP. As there was only one participant no statistical analysis was carried out.

Summary

One participant with diagnosed autism learned flexible relational responding with arbitrary stimuli. An alternating treatments design was used to examine the data, which clearly showed the T-IRAP was an effective method of teaching the target relation. It is also clear the T-IRAP was again notable faster than TT and also produced marginally higher mean accuracy. The data support the previous findings that T-IRAP is a more efficient method of stimulus presentation and again support the finding of Kilroe et al. (2011).

Study 2 aimed to teach arbitrary flexible relational responding; however it was necessary to introduce a colour prompt procedure to facilitate acquisition of the arbitrary relations. This procedure involved colour coding the stimuli so that “Like”

stimuli were the same colour and “Different” stimuli were different colours. This essentially added a non-arbitrary component to the otherwise arbitrary stimuli. When a criterion performance was achieved with the colour prompt procedure was removed and Kevin was required to select the correct response option with arbitrary relational stimuli. Kevin had previously mastered two sets of non-arbitrary stimuli prior to being exposed to the current arbitrary stimuli and done so after relatively few exposures. As such he may not have received sufficient multiple exemplar training with non-arbitrary stimuli, which Berens et al., (2007) suggest may be a prerequisite to establishing arbitrary relations. The colour prompt procedure represents specific training in the relevant relation with related non-arbitrary stimuli which successfully facilitated the establishing of an arbitrary stimulus relation, supporting the findings of Berens et al., (2007). Future research look to examining the extent to which non-arbitrary relations are prerequisites to arbitrary relations, in terms of the level of fluency required or the amount of multiple exemplars needed.

-Chapter 4-

Introduction

Relational Frame Theory posits that advanced cognitive behaviours including language, mathematics and other symbolic phenomena are all underpinned by complex relational responding, and there is a growing body of empirical evidence to support this. For example, O'Hora et al., (2005) examined the relationship between performances on a number of subtests of the WAIS-III (Wechsler, 1997) with performances on complex relational tasks. It was found that participants who passed the complex relational task produced significantly higher scores on the verbal and arithmetic subtests of the WAIS-III than participants who failed the relational test. Similarly, O'Toole et al., (2009) found that, when participants were required to respond to both consistent trials (in line with pre-established verbal relations) and inconsistent trials, (against pre-established relations) shorter response latencies and smaller difference scores, a measure of relational flexibility calculated by subtracting response latencies on inconsistent trials from consistent trials, predicted higher IQ scores. For example, relational responding tasks involved changing contingencies so that relations initially taught were subsequently reversed. Participants who were able to maintain high levels of speed and accuracy throughout these procedures were those whose scores correlated with high ability scores.

Cassidy et al., (2011) found that a training programme which employed multiple exemplar training in stimulus equivalence and complex relational responding for a number of different relations resulted in significant gains in participant IQ. It was found that experimental group IQ increased marginally following stimulus equivalence training, and substantially, following more complex relational training, while no significant increase occurred in the control group. Similarly it was found with a second population that a refined relational training

intervention alone was sufficient to bring about significant increases of more than 1 SD for 7 of the 8 participants. Surprisingly, Cassidy et al. found that a measure of fluency (summing test performances on training trials and dividing by the number of training blocks) and not baseline IQ significantly predicted gains in full scale IQ at follow up. Essentially, it was found that it was those who learned the target relations efficiently and effectively improved their full scale IQ scores.

These studies support the RFT prediction that fluent relational responding is essential to intelligent behaviour, and that correlations may be seen between rapid relational responding and high scores on ability tests. When taken together with the predictions made by RFT about the role of flexibility in relational responding and intelligence, along with the findings of basic and applied research in the area, it is predicted that providing multiple exemplar training in complex forms of flexible relational responding will positively impact upon cognitive functioning.

The current study aims to teach complex flexible relational responding to five children with autism and low IQ scores who participated in Study 1. Complexity in the relational flexibility was increased by implementing double contingency reversals. That is, the relations initially trained in Studies 1 and 2 were reversed (inconsistent trials), and subsequently reversed back to the original relations taught in Studies 1 and 2 (consistent trials). Trials following the first reversal were termed inconsistent trials, as they required participants to respond in a manner that was inconsistent with the previously established relational. Trials following the subsequent double reversal were termed consistent trials as they were consistent with the previously established relations in Studies 1 and 2. The double reversal was seen as adding additional complexity to the level of flexibility required as previously correct responses became incorrect under changed contingencies. As task children

with autism may find particularly difficult as they frequently have difficulty understanding when it is appropriate to break a rule or go against a previously established pattern of responding (Grandin, 2008; von Hahn et al., 2013).

Method

Participants and Recruitment

Five children that completed Study 1 also completed Study 3, detailed profiles of each participant are included in Study 1. A continued consent form was sent out to parents approximately half way through the research project which outlined the parents' rights to withdraw consent and offered them the chance to do so, no parent withdrew consent at this stage.

Settings and Materials.

IQ assessments. Each participant was assessed using two tests of verbal ability, the K-BIT (Kaufman et al., 1990) and the PPVT-IV (Dunn et al., 2007) prior to commencing Study 1 and again following completion of Study 3. The order in which these test were administered was counter balanced across participants. All assessments were carried out outside of teaching hours as to not take up valuable teaching time. Both tests are administered using easels contain various visual stimuli and separate scoring sheets. A detailed description of each of the tests is included in Study 1.

Table Top and T-IRAP. TT and T-IRAP procedures were carried in a manner identical to that used in studies 1 and 2. All stimuli were sourced via Google image search or created using the Microsoft paint function. The stimuli used were consistent across TT and T-IRAP and had been previously used in studies 1 and 2.

Settings. Each participant received 2-3 twenty minute research sessions a week. These sessions were conducted during normal class hours in each child's respective class room, under standard class room conditions. Assessments were carried out directly after school in the participants own class room with their respective teachers present acting in loco parentis for the duration.

Design

An alternating treatments design with double contingencies reversals and two contingency probes was employed. During the study the relative effectiveness of TT and T-IRAP procedures at teaching relational concepts that included both inconsistent and consistent trials with a Like/Different relation to five children with diagnosed autism was assessed. The order of the procedures was counter balanced across participants for the entire study, as such the order in which each participant received TT and T-IRAP conditions was carried over from Study 1, or in Kevin's case Study 2.

Inter Observer Agreement

An independent observer who had extensive ABA experience was trained to record TT data. The observer sat at the opposite end of the table to the participant and researcher and did not interact with the participant in any way. Approximately 23% of all TT trials were observed and a trial by trial analysis was conducted. 100% agreement between researcher and the independent observer was found.

Procedure

Post-test IQ assessments. Pre-test K-BIT and PPVT-IV scores were calculated for each participant prior to Study 1 and again following Study 3. The post-test assessment procedure was identical to pre-test in all regards. The PPVT-IV has two alternate forms to account for practice effects when re-testing, pre-test assessments were conducted using form A while post-test assessments used form B. The test manufactures report a high level of reliability between the two forms, with a reliability coefficient of .89, scores from both forms are considered comparable.

Specific method for double reversal for relations taught. Study 3 employed double contingency reversals, that is, the relations initially trained in

Studies 1 and 2 were reversed (inconsistent trials), and subsequently reversed back to the original relations taught in Studies 1 and 2 (consistent trials). Essentially, three contingencies were used in the study as a whole; first the contingency used in studies 1 and 2, then a second used at the beginning of Study 3 following the first reversal and a third used following the second reversal in Study 3. Both the 1st and 3rd contingencies reinforced selecting "Like" when stimuli were in fact like and "Different" when the stimuli were indeed different, a pattern of responding that is consistent with coordination and distinction relations in the general community. The 2nd contingency used required selecting "Like" when the stimuli were in fact different and "Different" when the stimuli were in fact the same, a pattern of responding that was inconsistent with previously taught relations and relations in the general community. Put simply, following each reversal the previously correct answers became incorrect and the previously incorrect answers became correct. As an illustrative example during the inconsistent trials in contingency 2 when two squares were presented participants were required to select "Different" rather than "Like", giving the opposite response to that which was previously taught during contingency 1. Similarly when a food item and an animal were presented together a response of "Like" was required. These type trials were termed *inconsistent* trials, as participants were required to respond in a manner that was inconsistent with the previously established relations, or put simply to give the *wrong* answer. Following inconsistent trials a second reversal was implemented and *consistent* trial blocks were presented. Consistent trial blocks required responses consistent with those taught in studies 1 and 2. For example when a triangle appeared with another triangle selecting the response option "Like" was required.

Subsequently two additional probes were carried out, firstly an inconsistent trials probe and secondly a consistent trials probe. Probes assess previous learning and do not provide contingent reinforcement or correction. Probes were implemented to briefly assess if participants could readily identify the changing contingencies of reinforcement independently and without the specific intervention that had been applied in all phases up to now. Participant who achieved high scores in both inconsistent and consistent probes more readily adapted to the changing contingencies and thus demonstrated greater levels of complex relational flexibility. Essentially the sequence used in study was as follows: consistent trials (studies 1 and 2) → inconsistent trials → consistent trials → inconsistent trials (probe) → consistent trials (probe). This alternation between inconsistent and consistent trial blocks was seen to require a greater level of flexibility in relational responding than that required when just responding to two different relations (coordination and distinction) simultaneously.

The TT procedure to teach both inconsistent and consistent trials was identical to that used in studies 1 and 2. Only the response requirements changed in accordance with the inconsistent and consistent trials response requirements. Similarly the T-IRAP procedure was identical to that used in studies 1 and 2. The instructions given were altered to accommodate the new response requirements in the inconsistent trials, each participant received a version of the following instructions dependent on the stimulus set to be used:

“Let’s do something new today, would you like to do some work on the computer? We are going to do some matching. We will see pictures that are the same and pictures that are different. When the two pictures are the same we need to pretend they are different and press the “K” button, that’s this one

(researcher to point to the K key). If the two pictures are different we need to pretend they are the same and press the “D” button, that’s this one (researcher to point to the D key). So, if a square comes up here (pointing to the top picture) and another square comes up here (pointing to bottom picture) I will point here (pointing to the onscreen prompt ‘*press K for Different*’) and you should press the ‘K’ key (pointing to the letter on the keyboard). If a picture of a Square and of a triangle comes up on the screen I will point here ‘*press D for Like*’ and you should press ‘D’ (pointing to the letter on the keyboard). If you get it right more pictures will come up. If you get it wrong a red X will come up, but that’s ok, we can try again to get the next one.”

Multiple stimuli sets were used in the current study to accommodate the range of competencies found within the population. For example, Andrew and Evan used stimulus set 1 (Square/Triangle), Steven and Ann used stimulus set 2 (Animal/food), while Kevin used stimulus set 3 (arbitrary stimuli). Illustrative examples of the three different stimulus sets used and the responses required during inconsistent trials are shown in Figures 10, 11 and 12.

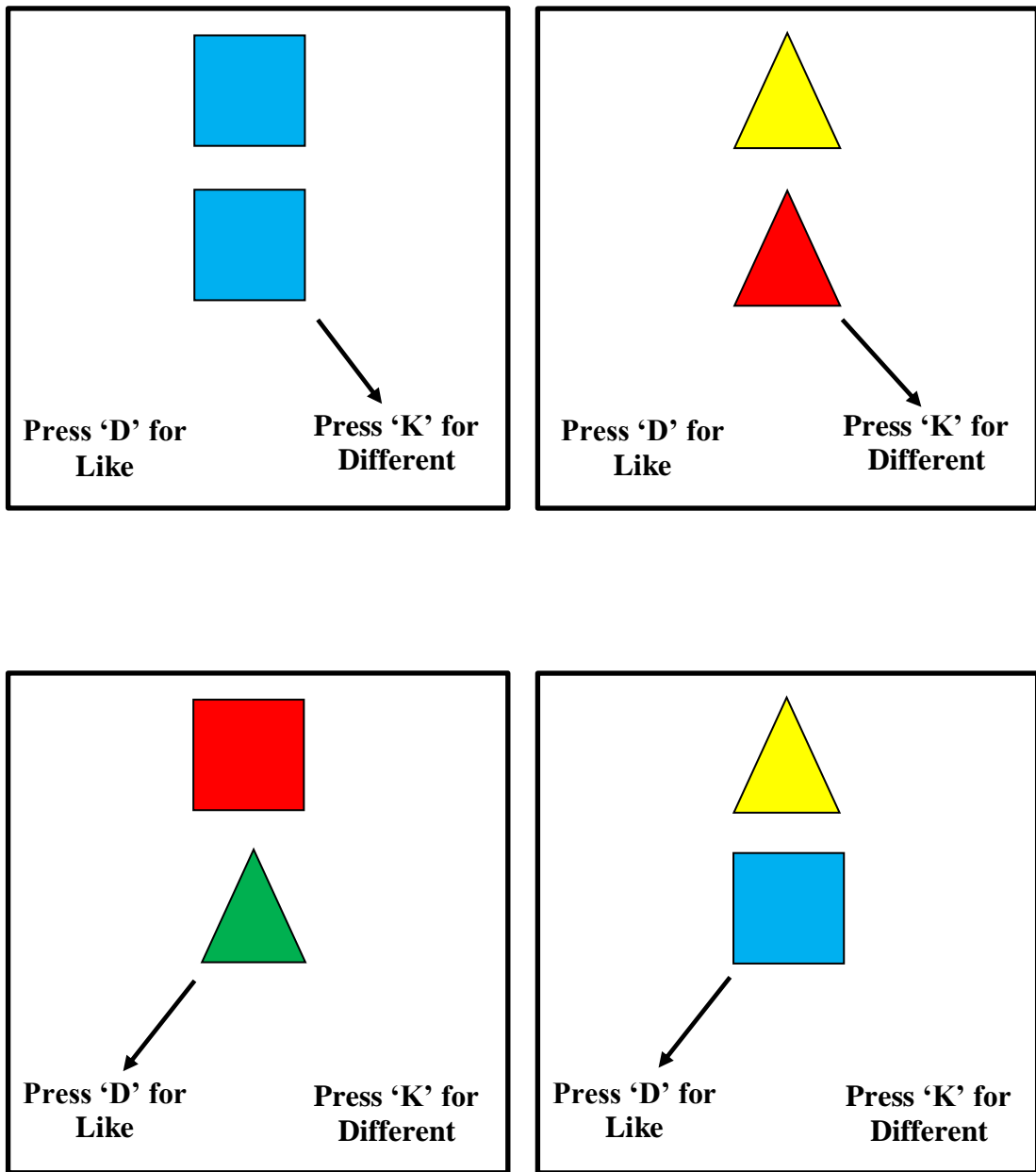


Figure 10. Examples of the four trial types used in the Square/Triangle stimulus set T-IRAP teaching procedure during “Like/Different” inconsistent complex relational flexibility training. Arrows signal the correct response; these responses were considered inconsistent with pre-established patterns of responding in the general population. The arrows did not appear on the screen during instruction

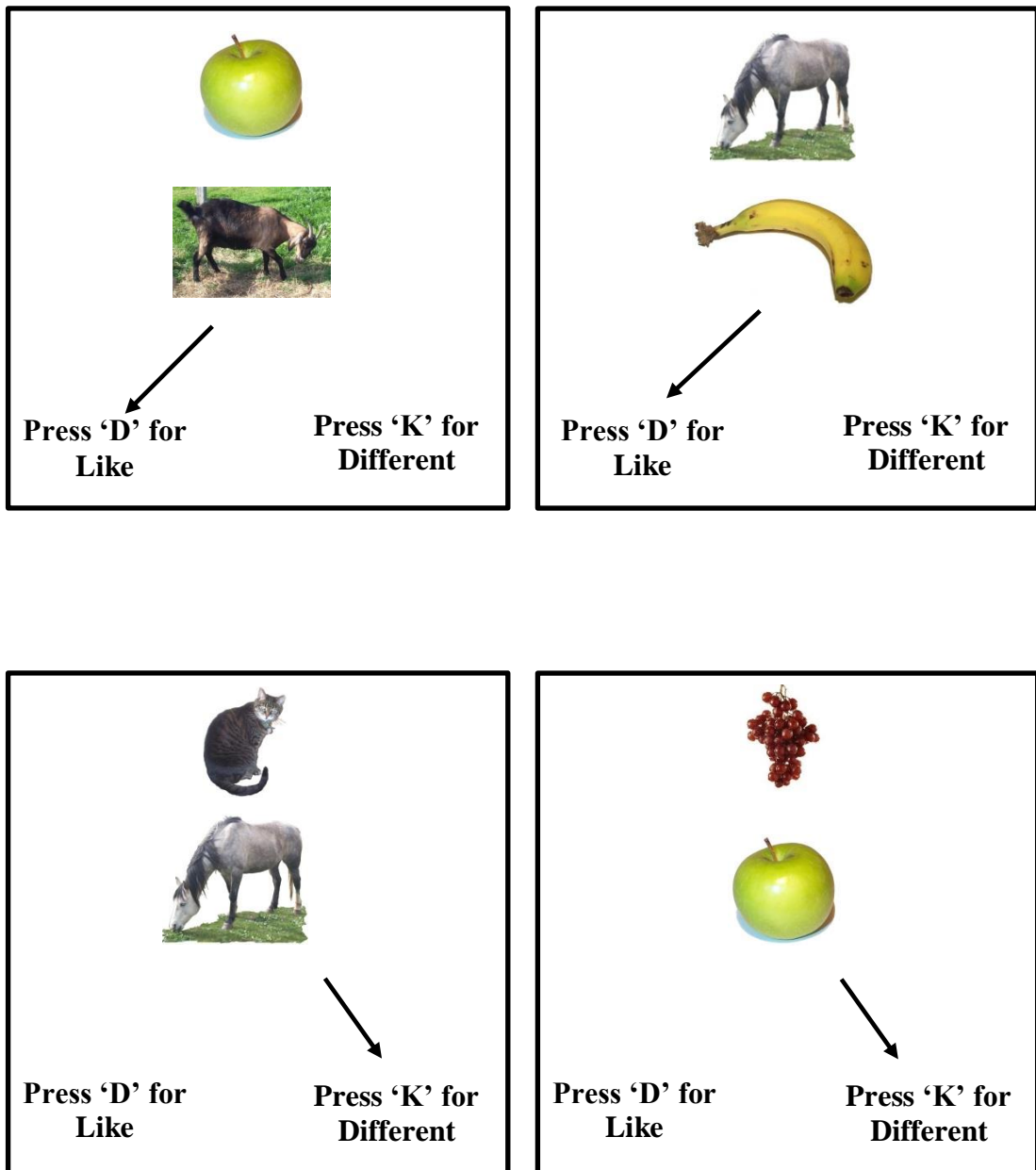


Figure 11. Examples of the four trial types used in the Animal/Food stimulus set T-IRAP teaching procedure during “Like/Different” inconsistent complex relational flexibility training. Arrows signal the correct response; these responses were considered inconsistent with pre-established patterns of responding in the general population. The arrows did not appear on the screen during instruction.

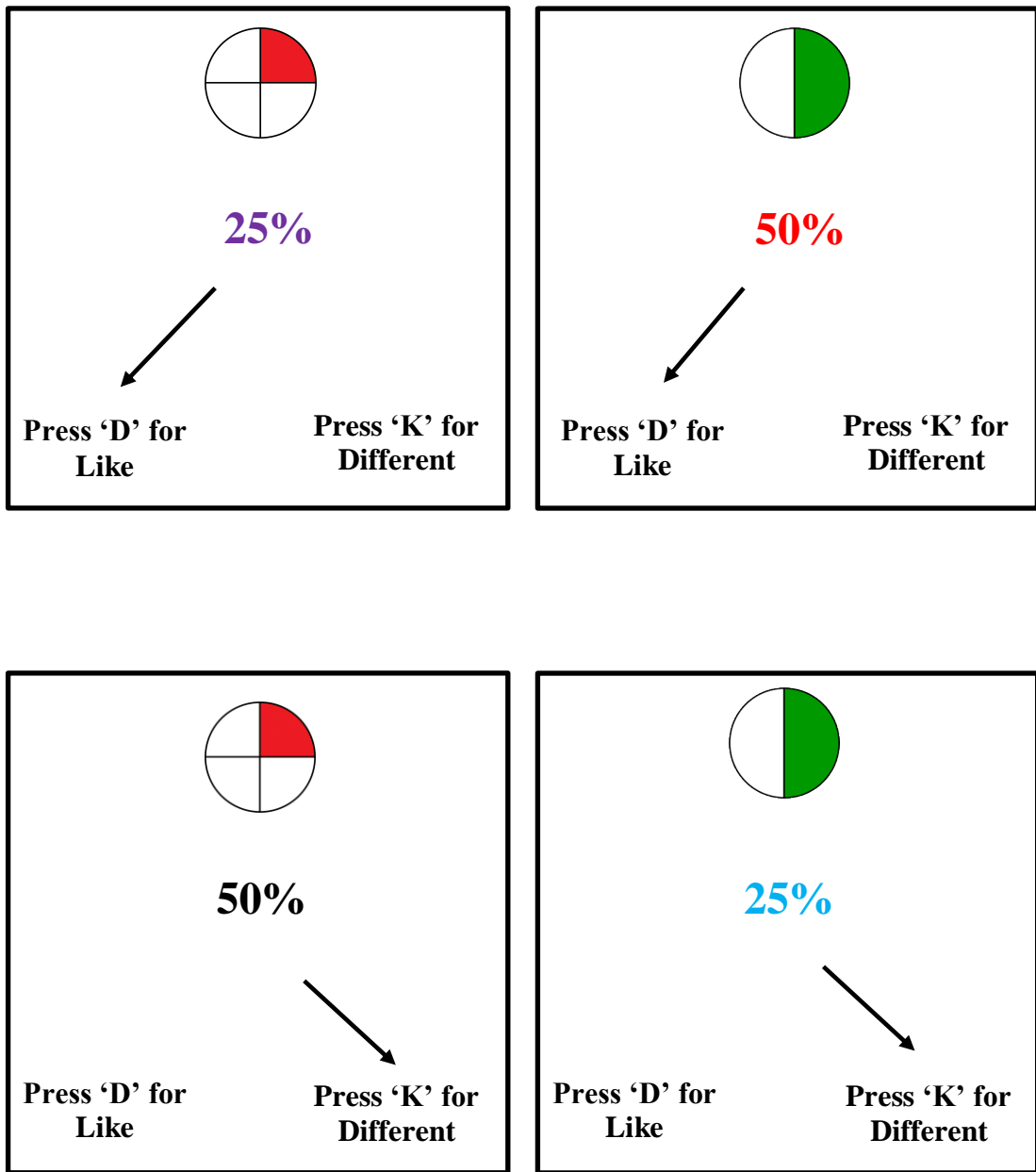


Figure 12. Examples of the four trial types used in the arbitrary stimulus set T-IRAP teaching procedure during “Like/Different” inconsistent complex relational flexibility training. Arrows signal the correct response; these responses were considered inconsistent with pre-established patterns of responding in the general population. The arrows did not appear on the screen during instruction.

Prompt Levels and Specific Alterations

Similar to studies 1 and 2 each participant may have been given different prompts depending on their individual abilities. Phases with the “Independent” or “IND” label above them required the participant to respond completely independently to all trials in that trial block no prompts were given. Phases with the label “Gestural prompt” or “GP” label included a gestural prompt for the first 50% of trials in that block. This prompt involved the researcher silently pointing to the correct response option prior to the participant making a selection for the first 12 of each 24 trials in a given trials block. The remaining 12 trials in these blocks required the participant to respond independently, without any prompt as per the “Independent” phases.

Phases with the label “Colour Prompt” provided an extra stimulus prompt by making the relations between stimuli more salient. This was done by assigning a unique colour to each category of stimuli, for example all squares were coloured blue while all triangles were coloured red. This increased the salience as each time two “Like” stimuli appeared not only where they related by being a different shape (for Study 3 the relation was reversed) they were also a different colour. Similarly, each time two “Different” stimuli came up they were the same colour as well as shape. For the animal/food stimulus sets it was not possible to change to colour of the actual picture, so the background of each set was coloured in the same manner as described above.

Results

Accuracy and Speed (Like-Different Non-arbitrary Double Reversals).

An alternating treatments design across five participants was used. The data for accuracy and for the speed of responding (duration of trial-blocks) using both TT and T-IRAP procedures with five participants are presented below in Figures 13 to 15. The data are presented in the same manner as studies 1 and 2.

Square/Triangle stimulus set. The data shown in Figure 13 for Andrew represents his performance across 5 distinct phases. Each phase represents a different prompt level or other variation within the experimental sequence. Andrew's accuracy data (percentage correct across trial-blocks) for initial independent trial-blocks during both TT and T-IRAP were moderately variable at low levels across five data paths with no trend evident, a moderate increase in the level of data for both TT and T-IRAP occurred in last data point of the phase. Andrew failed to demonstrate a criterion performance (80% for two consecutive data points or 100% once). Whereas the level of accuracy data was roughly similar in TT and T-IRAP conditions, the level of speed data were markedly different for both conditions. During initial independent trials speed of responding was greater for T-IRAP teaching compared with TT teaching; the speed/duration data in both conditions were variable with no trend evident.

A subsequent colour prompt procedure was implemented, accuracy data for TT and T-IRAP showed a steady increasing trend from low to relatively high levels. During the colour prompt procedure speed data were slightly variable at low-medium levels for both teaching conditions. The level of duration data was lower for the T-IRAP condition compared to TT teaching. Although accuracy data were ascending time constraints and an ethical obligation to finish with consistent trials

precluded continuing with the inconsistent trials and a return to independent consistent trials with no modifications was implemented. Accuracy data for both conditions were high and Andrew achieved a criterion accuracy performance after two data points. Speed data for TT and T-IRAP were stable at low levels with the later producing slightly faster responses.

Subsequently two probes were conducted to briefly assess the participants ability to independently identify and adapt to changes in the reinforcement contingency, firstly an inconsistent trials probe, which showed an immediate decrease in the level of accuracy data from previously high to medium levels. Speed data remained at levels similar to the previous phase for both conditions. Secondly a consistent trials probe, an immediate increase in level is evident in accuracy data for both TT and T-IRAP. Speed data for both TT and T-IRAP remained stable at levels similar to the two previous phases.

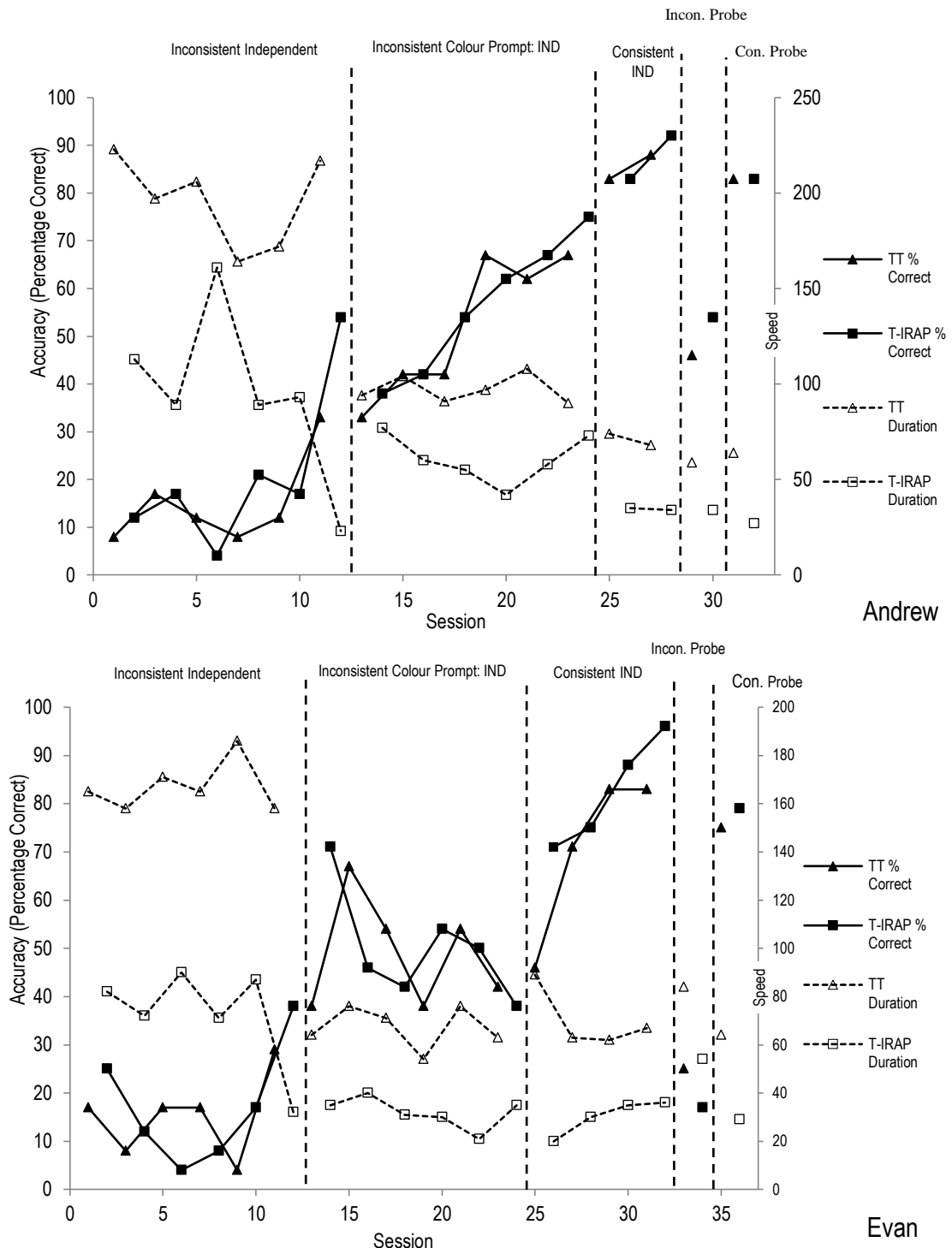


Figure 13. Table-Top and T-IRAP data for Andrew (top) and Evan (bottom) showing accuracy and speed of responding (duration) data for a double reversal Like/Different non-arbitrary inconsistent and consistent relations (i.e., shape: Square/Triangle), including inconsistent (Incon.) and consistent (Con.) probes

Overall, the level of speed data was again notably different for T-IRAP compared to TT conditions, and responding was faster in the former rather than the latter trials, however, this difference lessened as sessions progressed. Speed of responding was highly variable during early sessions becoming more stable in later sessions. Accuracy data for TT and T-IRAP are highly similar throughout all five phases.

Figure 13 represents Evans performance across 5 distinct phases. Evan's accuracy data for initial independent trial-blocks during both TT and T-IRAP were moderately variable at low levels across five data paths with no trend evident. Evan failed to demonstrate a criterion performance. The level of speed data were markedly different for TT and T-IRAP, with the latter producing faster times during initial independent trials; speed data in both conditions were variable with no trend evident. Following this a colour prompt procedure was implemented, accuracy data for TT and T-IRAP were highly variable at medium to low levels, with a downward trend evident in T-IRAP and no trend in TT. Evan did not achieve a criterion accuracy performance. Speed data remained slightly variable at low levels for the duration of the phase, with T-IRAP showing faster response times.

Subsequently a return to independent consistent trials with no modifications was implemented as time and ethical constraints precluded further teaching of inconsistent trials. An immediate increase in accuracy data for both conditions is evident following return to consistent trials, with both TT and T-IRAP showing a steady ascending trend to criterion after three data paths. Speed data for TT were slightly variable while T-IRAP data were stable at low levels with the latter producing slightly faster responses. Following this two probes were conducted, firstly an inconsistent trial probe, which showed an immediate decrease in the level

of accuracy data from previously high levels to very low levels in both conditions. Speed data remained at levels similar to the previous phase for both conditions. Secondly, a consistent trials probe where an immediate and dramatic increase from very low levels to relatively high levels is evident in the accuracy data for both TT and T-IRAP. Speed data for both conditions remained stable at levels similar to the two previous phases.

Overall the level of speed data was again notably different for T-IRAP compared to TT conditions, and responding was faster in the former rather than the latter trials, however, this difference decreased as sessions progressed. Speed of responding was highly variable during early sessions becoming more stable in later sessions. Accuracy data for TT and T-IRAP are highly similar throughout all five phases.

Animal/Food stimulus set. The data shown in Figure 14 for Steven represents his performance across 6 distinct phases. Steven's accuracy data for initial independent trial-blocks during both TT and T-IRAP showed a steady but slow increasing trend from low to medium levels across five data paths. Steven failed to demonstrate a criterion performance even though the data showed an upward trend the decision to introduce a colour prompt procedure was made, with a view to increasing the speed of acquisition. The level of speed data was markedly different for both conditions and speed of responding was greater during T-IRAP teaching compared with TT teaching; with TT speed data being quite variable with no trend evident and T-IRAP showing a stable downward trend.

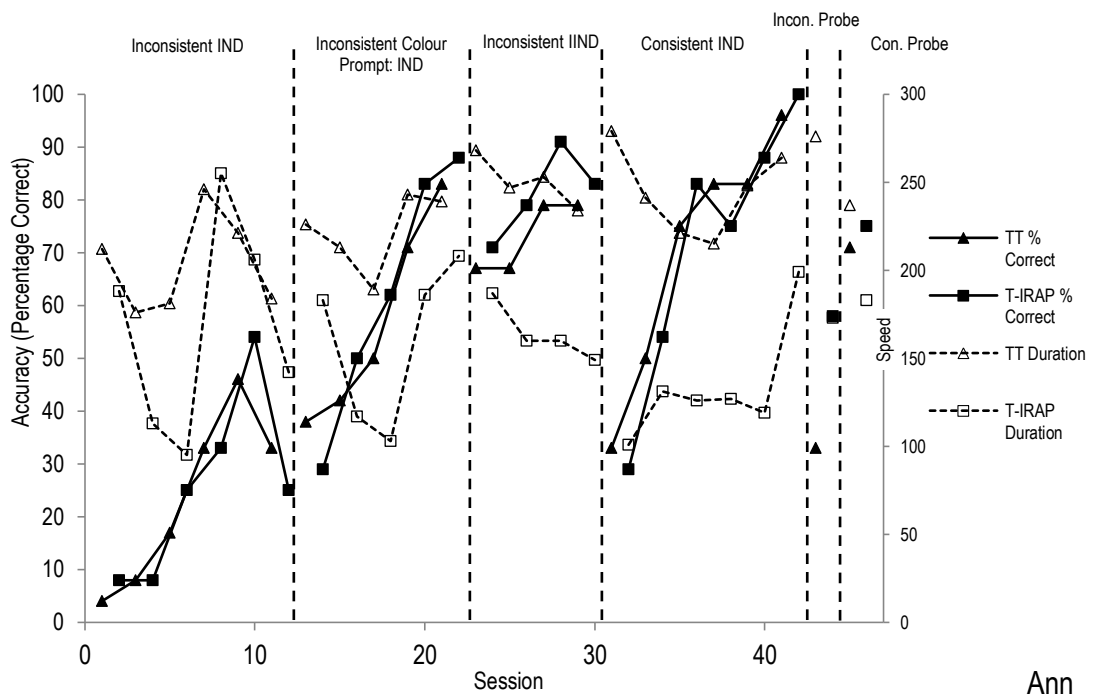
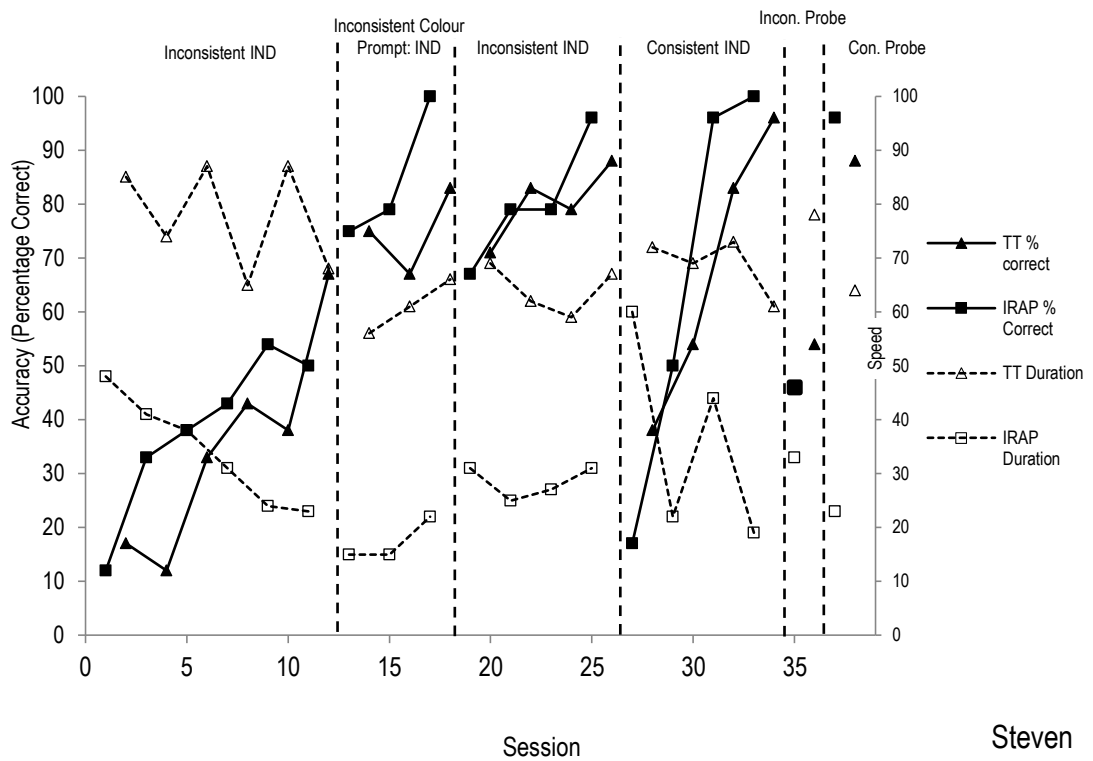


Figure 14. Table-Top and T-IRAP data for Steven (top) and Ann (bottom) showing accuracy and speed of responding (duration) data for a double reversal Like/Different non-arbitrary inconsistent and consistent relations (i.e., class: Animal/Food), including inconsistent (Incon.) and consistent (Con.) probes

The accuracy data for Steven when a colour prompt procedure was introduced were slightly variable at high levels for both teaching conditions with Steven achieving a criterion accuracy performance after two data paths. The level of duration data was lower for T-IRAP teaching, indicating greater speed of responding in this condition, as compared to TT teaching. Subsequently, the colour prompt procedure was removed and Steven progressed back to independent inconsistent trials with no modifications where he achieved a criterion accuracy performance after three data paths. Accuracy data for both T and T-IRAP are highly similar at high levels with both showing an ascending trend. Duration data were slightly variable for both T-IRAP and TT, with the former being at notability lower levels.

Subsequently a return to independent consistent trials with no modifications was implemented. An immediate decrease in accuracy data for both conditions is evident following return to consistent trials, with both TT and T-IRAP showing a steady and sharp ascending trend to criterion after three data paths. Speed data for TT were slightly variable at medium levels, while T-IRAP data were highly variable at medium to low levels. Following this two probes were conducted, firstly an inconsistent trial probe, which showed an immediate decrease in the level of accuracy data from previously high levels to medium levels in both conditions. The level of speed data remained at levels similar to the previous phase for both conditions. Secondly, a consistent trials probe was implemented where an immediate and increase from medium levels to high levels is evident in the accuracy data for both TT and T-ITAP. Speed data for both conditions remained at levels similar to the two previous phases.

Ann's data shown in Figure 14 represent her performance across 6 distinct phases. Each phase represents a different prompt level or other variation within the

experimental sequence. Ann's accuracy data for initial independent inconsistent trial-blocks during both TT and T-IRAP were a little variable at low levels with a slow upward trend across five data paths and failed to demonstrate a criterion performance. The level of speed data was highly variable in both conditions with no trend evident in either TT or T-IRAP during initial independent inconsistent trials.

The accuracy data for Ann when a colour prompt procedure was introduced showed a steady and sharp ascending trend from low levels to criterion after four data paths. Speed data were variable with no clear trend for both T-IRAP and TT at medium levels, with the former producing faster responding compared to the later. Ann progressed to independent inconsistent trials with no modifications and reached criterion after three data paths. Accuracy data for TT and T-IRAP was roughly similar with slight variability at very high levels and slight increasing trend. Speed data for T-IRAP was again at lower levels compared to TT with both showing slight downward trends.

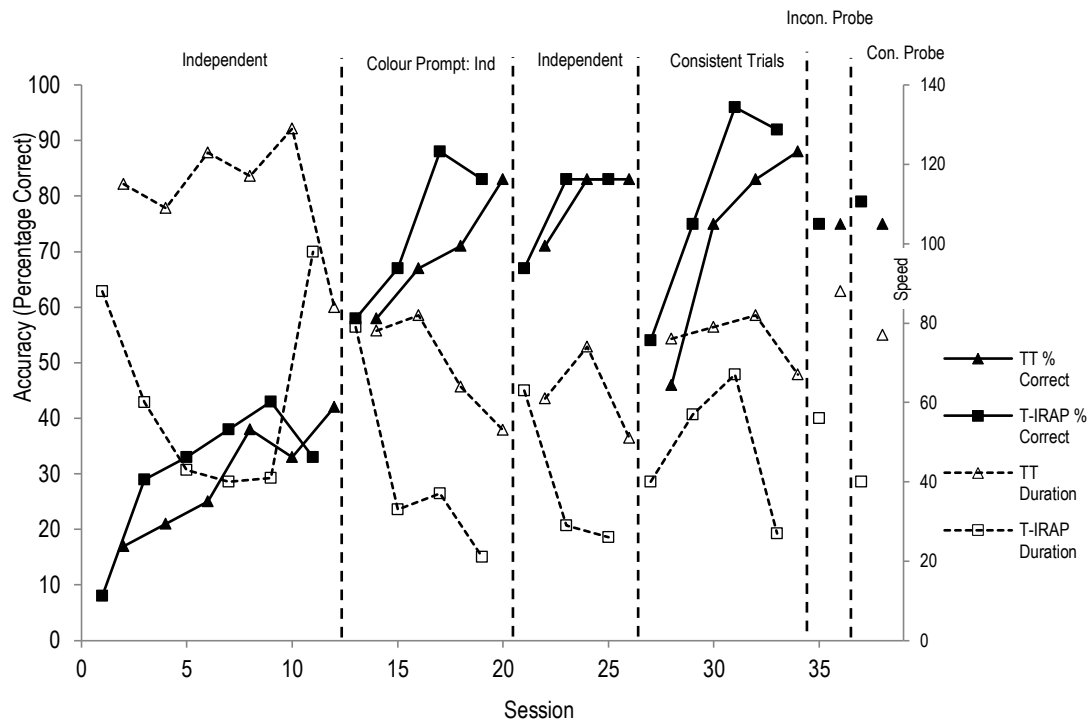
Subsequently Ann was exposed to a return to consistent independent trials with no modifications, an immediate decrease in the level of accuracy is evident compared to the previous phase, however, both TT and T-IRAP show a steady upward trend to a criterion accuracy performance after five data paths. Duration data is variable in both T-IRAP and TT with no clear trend evident, T-IRAP again produced a faster performance. Following this Ann was exposed to two probes, firstly an inconsistent trial probe, which showed an immediate decrease in the level of accuracy data from previously high levels to medium levels in T-IRAP and low levels in TT. The level of speed data remained at relatively high levels for TT and medium levels of T-IRAP. Secondly, a consistent trials probe was implemented where an immediate increase from low to medium levels is evident in the

accuracy data for TT, while T-IRAP accuracy data remained at medium levels. Speed data for both conditions was at medium levels with T-IRAP again being the faster of the two conditions.

Arbitrary stimulus set. The data shown in Figure 15 for Kevin represent his performance across 6 distinct phases. Kevin's accuracy data for initial independent inconsistent trial-blocks during both TT and T-IRAP are slightly variable with a slow increasing trend across five data paths and failed to demonstrate a criterion performance. The ascending trend was slight and the decision to introduce a colour prompt procedure was made with a view to increasing the speed of acquisition. Speed data were markedly different for both conditions and speed of responding was greater during T-IRAP teaching compared with TT teaching in initial independent inconsistent trials; with both conditions speed data being quite variable with no trend evident.

Consequently a colour prompt procedure was introduced; accuracy data for TT and T-IRAP were roughly similar and slightly variable at high levels with Kevin achieving a criterion accuracy performance after three data paths. The level of duration data was lower for T-IRAP teaching, indicating greater speed of responding in this condition compared to TT teaching with a downward trend evident in both conditions. Following achieving a criterion accuracy performance the colour prompt procedure was removed and Kevin progressed back to independent inconsistent trials with no modifications where he achieved a criterion accuracy performance after two data paths. Accuracy data for both T and T-IRAP are quite similar at high levels with both showing an ascending trend. Duration data were slightly variable for TT with no trend events, the level of T-IRAP data was lower compared to TT indicating faster responding, with T-IRAP showing a downward trend. Subsequently, a return

to independent consistent trials with no modifications was implemented. An immediate decrease in accuracy data for both conditions is evident following return to consistent trials, with both TT and T-IRAP showing a steady and sharp ascending trend to criterion after three data paths. Levels of speed data were again markedly different for both conditions, with T-IRAP producing faster times compared to TT, both conditions showed some variability.



Kevin

Figure 15. Table-Top and T-IRAP data for Kevin showing accuracy and speed of responding (duration) data for a double reversal Like/Different arbitrary inconsistent and consistent relations, including inconsistent (Incon.) and consistent (Con.) probes

Two probes were conducted following Kevin's criterion accuracy performance with independent consistent trials, firstly an inconsistent trial probe, which showed a slight decrease in the level of accuracy data, compared to the previous phases in both TT and T-IRAP conditions, to just below criterion

performance levels. The level of speed data showed a slight increase compared to the previous phase for both conditions. Secondly, a consistent trials probe was implemented; accuracy remained at high levels just below criterion performance levels for both TT and T-IRAP, while duration data showed a slight decrease in level compared to the previous phase for both conditions.

Mean Accuracies and Response Times for T-IRAP and Table-top

Mean response times and accuracies for Study 3 are presented in table 3 under the “double reversals” heading. A paired samples t-test was conducted to compare average duration in TT and T-IRAP condition. It was found that for Study 3 there was a significant difference in average duration for T-IRAP ($M=69.88$, $SD=50.38$) and TT ($M=122.81$, $SD=64.09$); $t(4) = -3.7448$, $p=0.002$. Paired samples t-test examining mean accuracy scores for Study 3 between T-IRAP ($M=55.87$, $SD=8.07$) and TT ($M=52.35$, $SD=8.73$) indicated a significant difference; $t(4) = 6.893$, $p=0.002$.

Mean response times for all sessions to date by participants and stimulus set are presented in Table 3. Mean response times were calculated by summing the duration of each trial block for each participants and dividing by the total number of trial blocks. Mean T-IRAP ($M=72.91$, $SD= 51.48$) durations were approximately 1.9 times shorter on average compared to TT ($M=144.35$, $SD=69.41$), paired-samples t-tests indicated that these differences were significant: $t(4) = -7.77$, $p=0.001$. Mean durations for T-IRAP ($M=69.88$, $SD=50.38$) during the double reversals were also shorter compared to TT ($M=122.81$, $SD=64.10$), paired samples t-tests indicated this difference was also significant; $t(4) = -7.77$, $p=0.002$. As predicted, the current study found the T-IRAP to be a faster method of stimulus presentation consistent with previous finding of Kilroe et al., (2011).

Paired samples t-test examining mean accuracy scores for all session to date between T-IRAP ($M=63.49$, $SD=5.72$) and TT ($M=61.08$, $SD=5.13$) indicated a significant difference; $t(4) = 8.46$, $p=.001$. Similarly, for the double reversals individually differences between T-IRAP ($M=55.87$, $SD=8.06$), and TT ($M=52.35$, $SD=8.73$) were not significant; $t(4) = 6.89$, $p=.002$. These results show that although accuracy data for both conditions appear roughly similar in for each participant across all stimulus sets T-IRAP produced significantly higher levels of accuracy during double reversals and the study as a whole.

Table 3

Mean Accuracies and Durations for all Sessions to Date by Participant and Stimulus

Set

	T-IRAP Mean Duration (Seconds)	Table-Top Mean Duration (Seconds)	T-IRAP Mean Accuracy (Percent)	Table-Top Mean Accuracy (Percent)
<i>Stimulus set 1</i>				
Andrew	48	166.8	71.6%	70.72%
Evan	67.8	138.1	66.67%	69.64%
Steven	30.7	148.5	75.07%	81.33%
Kevin	39.5	50.5	98%	94%
Ann	124.7	312.6	85.67%	81.33%
<i>Stimulus set 2</i>				
Steven	29.4	121.8	78.88%	81.13%
Kevin	41	56	84.6%	95%
Ann	185.9	285.3	71.31%	72.69%
<i>Stimulus set 3</i>				
Kevin	52.1	163.5	73.41%	67.93%
<i>Double reversals</i>				
Andrew	66.4	126.8	48.44%	43.94%
Evan	46.1	102	46.17%	42.67%
Steven	30.7	69.6	63.68%	61.52%
Ann	157	231	58.74%	54%
Kevin	49.7	84.7	62.32%	59.68%

Post-test Verbal Ability Measures

The IQ scores obtained in the current sample ranged from 40-62 on the K-BIT at pre-test and 20-50 on the PPVT-IV, at post-test the scores ranged from 40-74 and 20-60 respectively. Mean IQ score was well below average at pre ($N = 5$, $M = 44.4$) and post-test ($N=5$, $M=47.2$) for the K-BIT, as well as the PPVT-IV at both pre ($N=5$, $M=29.4$) and post ($N=5$, $M=30.2$). The IQ score generated by the K-BIT is a composite score, which corresponds to the sum of the standard scores on the vocabulary and matrices subtests.

Table 4 shows the standard score and raw for each of the participants on both the K-BIT and the PPVT-IV at pre and post-test. Some variation between pre and post test scores in the desired direction occurred however paired samples t-tests indicated no significant difference existed between pre (M=44.4, SD=9.84) and post(M= 47.2 , SD= 15) K-BIT standard scores for five participants; $t(4)=-1.20$, $p=.296$. Similarly paired samples t-tests indicated that no significant difference existed between PPVT-IV pre (M= 29.4, SD=12.88) and post-test (M=30.2, SD=17.18) standard scores; $t(4) =-.311$, $p=.757$. Furthermore, no significant difference was found between pre (M=17, SD=10.07) and post (M= 21, SD=15.15) K-BIT raw scores; $t(4) =-1.480$, $p=.213$. Finally, PPVT-IV pre (M38.6, SD=10.714) and post (M=45.8, SD=20.705) raw scores showed no significant difference; $t(4) =-1.18$, $p=.303$.

Table 4

Standard (S) and Raw(R) Scores for Each Participant on the K-BIT and PPVT-IV at Pre and Post Test

	K-BIT (S) Pre Post	K-BIT (R) Pre Post	PPVT-IV (S) Pre Post	PPVT-IV (R) Pre Post
Andrew	40 42	18 23	23 21	46 46
Ann	40 40	10 12	34 30	34 31
Evan	40 40	5 5	20 20	22 26
Kevin	62 74	31 45	50 60	48 79
Steven	40 40	21 20	20 20	43 47

Supplementary Analysis

In addition to visual analysis of data measures of flexibility and fluency were calculated based on methods used in previous behavioural research as a supplementary method of analysis. Differences scores, similar to those used by O'Toole et al., (2009) were calculated by subtracting accuracy scores during inconsistent probes from accuracy scores during consistent probes. Difference scores are considered a measure of flexibility, as the relational responding tasks involved changing contingencies so that relations initially taught were subsequently reversed and then double reversed back to the original relations. Participants who were able to maintain high levels of speed and accuracy throughout these procedures would have smaller difference scores. As such the lower the difference score the greater the flexibility demonstrated. Difference scores were calculated for both TT and T-IRAP conditions as well as the two combined.

It was found that difference scores for TT (i.e., greater flexibility in TT) a significant negative correlation existed between TT difference scores and gains in K-BIT raw scores ($r=-.965$, $p=.008$), K-BIT standard scores ($r=-.976$, $p=.005$) and gains in PPVT-IV raw scores ($r=-.896$, $p=.039$). T-IRAP difference scores did not show significant correlations with gains in IQ, while the overall difference scores across both TT and T-IRAP correlated with gains in K-BIT raw scores ($r=-.920$, $p=.027$).

A measure of fluency was also calculated based on the method used by Cassidy et al., (2011) by summing accuracy scores across all stimulus sets and dividing by the total number of trial blocks. As such higher fluency scores represent greater levels of fluency. A correlation analysis was carried out to examine if these measures predicted any of the variability in pre and post-test IQ scores. It was found

that greater fluency scores related to greater TT flexibility ($r=-.963$), post-test K-BIT standard scores ($r=.978$), post-test K-BIT raw scores ($r=.907$), post-test PPVT-IV standard scores ($r=.894$, $r=.964$), PPVT-IV post-test raw scores ($r=.964$), gains in K-BIT raw scores ($r=.881$), gains in K-BIT standard scores ($r=.958$) and gains in PPVT-IV standard scores ($r=.933$). Overall it may be said that measures of fluency correlated highly with a range of IQ measures.

Further analysis revealed that, of the IQ scores, pre-test K-BIT standard scores were the best predictor of overall fluency ($r=.984$), and flexibility in the TT teaching procedure ($r=-.964$), as well as gains in K-BIT raw scores ($r=.925$), standard scores ($r=.986$) and PPVT-IV raw scores ($r=.976$). Finally, pre-test scores on the vocabulary subtest of the K-BIT predicted greater levels of overall flexibility ($r=-.910$)

Summary

Findings in Study 3 show that three of five participants with diagnosed autism learned to respond correctly to a complex “Like/Different” relational flexibility task during TT and T-IRAP teaching procedures with T-IRAP being a more efficient method of stimulus presentation. Responding during T-IRAP was found to be significantly faster compared to TT for all stimulus sets taken together and each individually. Two participants did not acquire the inconsistent response, it may be said that these participants displayed lower levels of relational flexibility compared to those who achieved a criterion performance with inconsistent trials.

The current study attempted to teach inconsistent responding to five participants with autism and low IQ scores using an alternating treatments design. Findings show it was successful in doing so with three of the five participants who all learned the inconsistent response and successfully reverted to the consistent

response relatively quickly. All participants achieved a criterion accuracy performance upon return to consistent trials. Subsequently two probes were carried out, first and inconsistent then a consistent probe. It is clear from the data shown in Figures 13 to 15 that some participants showed greater flexibility and readily adapted to the changing contingencies, achieving high levels of accuracy in both inconsistent and consistent trials. Difference scores were calculated as a measure of relational flexibility, it was found that difference scores rather than speed predicted gains in IQ.

A slight caveat is worth noting here, Andrew failed to demonstrate a criterion accuracy performance during inconsistent trials. However, the data show an ascending trend during the gestural prompt phase (phase 2; Figure 10), despite this, a reversal to consistent trials implemented due to time constraints. If this phase was continued, it may have transpired that Andrew achieved a criterion performance with the gestural prompt and subsequently with inconsistent trials. This is speculation, but given the trend evident in the data, it is nonetheless a possible alternative scenario.

-Chapter 5-

General Discussion

The current thesis successfully taught flexible relational responding to four boys and one girl with autism and extremely low ability scores as measured by the PPVT and the K-BIT. IQ assessments were carried out prior to experimental procedures, it was found that across both tests all participants fell into the descriptive category of “extremely low” placing them within the 1st percentile when compared to age based norms. The test scores were measured pre and post the teaching intervention, and some gains in IQ were observed for individual participants. At a group level there was a non-significant difference between pre and post IQ measures but the very small sample size may be considered relevant in this regard. The study successfully combined ABA methods for example positive reinforcement, prompting, fading, with RFT-based instructional procedures to teach flexible relational responding, and children with autism learned to use a computerised interactive teaching procedure (T-IRAP). The utility of TT and T-IRAP teaching procedures at teaching an RFT based intervention to teach flexible relational responding was compared using an alternating treatments design. T-IRAP was found to produce significantly faster speeds of responding for all participants, it was also found to produce significantly higher levels of accuracy for all participants across the entire study compared to TT teaching. It should be emphasised once more that the author is not suggesting the replacement of TT teaching methods with computerised teaching programmes, but rather that the T-IRAP could be used to supplement TT teaching and students could use it independently for self-practice and maintenance purposes.

Summary of Findings

Study 1 employed an alternating treatments design to compare the utility of traditional TT and T-IRAP at teaching an RFT based intervention to teach basic non-arbitrary flexible relational responding with five children with autism and extremely low IQ scores. The research thus extends the study by Kilroe et al. (2011) which used a multiple baseline design across participants to demonstrate that on each occasion that the T-IRAP was introduced, responding for participants with autism was faster and more accurate. The current findings support those of Kilroe et al. in that similar effects on participant responding were demonstrated, albeit that the positive impact on accuracy was less pronounced than for participants in the Kilroe et al. study. This latter effect may be related to the level of functioning of the children, as the current study extended the Kilroe et al. research by successfully using the T-IRAP to teach relational flexibility with children with lower levels of functioning than participants in the former study. Whether the level of functioning was influential in lessening the positive influence of the T-IRAP procedure regarding accuracy of responding remains speculative however, and further research would be needed to confirm or disconfirm as the case may be. The data show that an RFT based intervention was successful in teaching the participant flexible relational responding with both TT and T-IRAP and also show that the T-IRAP produced response speeds that were significantly faster than TT.

Study 2 was conducted with one participant Kevin and aimed to teach Kevin (diagnosed with autism, participated in Study 1 also) flexible relational responding with arbitrary stimuli. An alternating treatment design was again used to compare TT with T-IRAP procedures in terms of speed and accuracy of participant responding shown with each procedures. The data show T-IRAP was an effective method of

teaching the target relations. It is also clear the T-IRAP again produced response speeds that were notably faster than TT, however, levels of accuracy data remained quite similar during both TT and T-IRAP.

An aim of Study 3 was to increase complexity in relational flexibility for five participants with autism and extremely low IQ scores. Pre and post-intervention measures of IQ were taken and compared to examine if any gains occurred following relational flexibility training. It was hypothesized that participants who learned to respond to relational stimuli in a flexible manner would show gains in IQ from pre to post-intervention. Some differences in IQ scores in a mostly positive direction were evident for 3 participants. The differences shown represent important improvements in measures of cognitive abilities for each individual.

One of the main research aims of the current study was to compare the efficacy of two methods, T-IRAP and TT, at teaching the same relational targets. With the exception of Study 2 (as statistical analysis was not possible with only one participant) all of the above differences in speed of responding between T-IRAP and TT were statistically significant. There was also a significant difference between mean TT and T-IRAP duration across all three studies collectively. Consequently it may be said that speed of responding during T-IRAP was significantly faster than during TT across the entire study, a possible reason for this is that T-IRAP as a method of stimulus presentation may be more conducive to fluency in relational responding.

It was also found that for the study as whole T-IRAP produced significantly higher mean accuracies compared to TT; this difference was not readily apparent upon visual analysis of graphed data, thus it was small, however it was found to be statistically significant and in this sense provides some minimal support for accuracy

findings in Kilroe et al. (2011) It can be concluded that T-IRAP, which produced significantly faster response speeds and higher levels of accuracy across the study as a whole, is a more effective and efficient method of teaching children with autism and extremely low IQ scores relational flexibility. These findings support those of Kilroe et al. (2011) which found T-IRAP to produce higher levels of accuracy and faster responding compared to TT. This adds further support to the suggestions of Kilroe et al., that the T-IRAP may, at the very least, be a useful supplementary teaching tool. Continued research is needed to further develop the T-IRAP as a teaching tool, and large scale studies with greater numbers of relations, different populations and tests of generalisation are needed to fully evaluate its potential.

As stated previously, the study assessed participant's pre and post intervention IQ scores. The aim was to examine if teaching flexible relational responding would result in gains in intelligent behaviour measured by IQ tests in line with predictions made by RFT and the findings of Cassidy et al., (2011). No significant differences were found from pre to post-intervention assessments at a group level, that is not to say that individual gains do not represent substantial improvements. At pre-intervention assessment all participants fell within the 1st percentile, ranking them in the lowest 1% of the population in terms of verbal ability.

The largest increases found were 12 points in K-BIT standard scores, and 10 points in PPVT-IV standards, both for Kevin. Kevin's 12 point increase in K-BIT standard scores represents a remarkable increase in his individual performance taking him from the 1st percentile at pre-intervention to the 4th percentile at post-intervention. Similarly his increase from a standard score of 50 on the PPVT at pre-intervention assessment, which placed him in the lowest 0.1% of the general population, to his post-intervention score of 60, which brings him up to the lowest

2.2% of the population, represents a substantial improvement. It is worth noting that Kevin, who achieved the largest gains in IQ post-intervention, demonstrated the greatest levels of flexibility overall and was the only participant to be exposed to arbitrary stimulus relations. These findings are impressive and in line with Cassidy et al., (2011) however, a caveat should again be sounded because of the preliminary nature of the data from the current study and from Cassidy et al., and much more extensive research will be necessary before anything more than tentative suggestion can be put forward. It should also be noted that although IQ measures tend to remain stable across the lifespan of an individual, raised IQ levels have been shown previously with ABA measures that did not specifically target relational responding (e.g., Lovass, 1987).

These gains are substantial increases for the individual and may show potential for greater increases in IQ over time however, the author emphasises that it is not possible to attribute these gains purely to teaching flexible relational responding, as Kevin was receiving ongoing ABA instruction in a number of academic areas over the course of the study. There were no previously recorded IQ measures to examine whether changes in IQ level had occurred at any point previously for Kevin as a result of ABA teaching methods or other factors. In the current investigation, Kevin was the only participant to demonstrate increases of this magnitude, with the remaining participants showing marginal changes overall. Andrew showed a 2 point gain in K-BIT standard scores, while the remaining three participants showed no difference in standard scores from pre to post-intervention. Both Andrew and Ann showed small decreases in PPVT standard scores of 2 and 4 points respectively, while data for Evan and Steven showed no difference.

Participants who achieved greater accuracy with a greater number of stimulus sets in the shortest number of trials were considered to have demonstrated the greatest levels of flexibility. The data show a range of competencies in relational to flexibility across participants, Kevin who achieved mastery on a greater number of stimulus sets, and displayed highly accurate and fluent responding on stimulus sets 1 and 2 may be said to have displayed the greatest levels of flexibility overall. Kevin was also the only participants reach criterion with arbitrary stimulus relations. On the other hand, Evan and Andrew who achieved mastery with stimulus set 1 only and struggled to do so, may be considered to have shown the lowest levels of flexibility. These finding emphasise the importance of flexibility within relational responding as those who demonstrated greatest levels of flexibility on early stimulus sets maintained greater levels of accuracy across trials.

Supplementary correlational analysis revealed that K-BIT pre-test standard scores predicted greater fluency (total summed performances divided by total number of trial blocks) and flexibility (subtracting the accuracy scores for inconsistent probes from consistent probes). This may provide tentative support for RFT assertions that fluency and flexibility in relational responding are an important part of intelligent behaviour. Those who scored higher on the K-BIT produced more flexible and fluent performances. However, these data are very preliminary and must be taken cautiously, the correlation may also have been skewed by Kevin's performance showing relatively large gains compared to the other participants. Findings show the pre-test vocabulary subtest of the K-BIT correlated with greater T-IRAP fluency and that gains in the vocabulary subtest correlated with greater overall flexibility in relational responding tasks, which support the findings of

O'Hora et al., (2005) who found a correlation between vocabulary subtests of the WAIS-III and the number of correct responses on a relational task.

With regard to vocabulary, RFT suggests that it is not the verbal content itself but the relational skills involved in developing that content that are important, as such those with highly elaborated vocabulary tend to have highly elaborated relational skills (Barnes-Holmes et al., 2001) Although these findings are not surprising from an RFT perspective, what is interesting is that the matrices subtest of the K-BIT did not show a significant correlation with measures of relational ability. It would be expected that a relationally rich subtest like the matrices included in the K-BIT would predict relational ability, however this was not the case in the current study and further investigation is needed. It may be the case that given the low levels of ability within the current population, and that the matrices represent more complex relational abilities such as analogy, that the matrices were simply too complex for the population. Thus, no correlation was found, while the relational abilities involved in vocabulary subtests were more on par with the level of the current population. Furthermore, the gains achieved by Kevin on the K-BIT were more impacted by gains in the matrices subtest rather than the area of vocabulary. Intuitively, it may be unsurprising that the more complex forms of relational responding which Kevin demonstrated during intervention may be related to measures of fluid intelligence like analogy matrices, while more basic relational responding may be more sensitive to relationally rich measures of crystallised intelligence such as vocabulary.

Table-Top and T-IRAP Procedures

It is likely that the greater overall speed of responding produced by the T-IRAP was related to the rapid speed at which the computerised program can present

stimuli as compared to TT methods. TT methods were set up to present stimuli as fast as possible, yet were still significantly slower. As such the T-IRAP by its very nature may be more efficient tool for teaching fluent responding, as children may find the computer presentation format more appealing. Although not examined directly, it is unlikely that the greater durations found in TT conditions are solely due to time taken to manipulate the stimuli by hand. For example, Ann took an average 155 seconds to complete a T-IRAP trial block and 276 seconds to complete a TT trial block. This difference is appears too large to be accounted for solely by the additional time taken to manually manipulate the stimuli, maybe then that shorter durations are a result of the T-IRAP being more conducive to speedy and fluent responding due to the rapid presentation of stimuli. Thus the T-IRAP shows great potential as a tool for developing fluency in relational responding, by actively promoting faster responding and facilitating independent self-training and fluency. Previous researchers have highlighted the need for such a tool, which can be employed to increase fluency in specific behaviour repertoires (Cassidy et al., 2011; Williams, Myerson, & Halle, 2008).

Specifically an efficient and effective automated teaching tool like the T-IRAP, could be used independently by students to practice topics being taught in class and build fluency. This could reduce the need for one-to-one instruction in some cases and could also be a very useful for self-training and maintenance of previously taught material. The use of the T-IRAP would allow for better learning outcomes using minimal resources and without additional teaching hours. Given the speed capacity of the T-IRAP a large number of trials can be administered in a relatively short period of time, allowing students to be exposed to multiple exemplar training to build fluency in target concepts without extensive teaching resources.

Binder, (1996) points to fluency as beneficial in terms of acquisition, retention, generalisation and resilience to distraction, similarly RFT posits that fluent and flexible relational responding is a key aspect of higher order human cognition and intelligence.

Flexibility in Relational Responding

The present findings support the RFT view of the importance of flexibility in relational responding regarding advanced cognitive and intelligent behaviour. The combining of multiple relations, and the subsequent reversals carried out in the current study are seen as flexible behavioural repertoires. Similarly to Cassidy et al., (2011) one might conceptualise the current study as explicit training in relational flexibility, including both basic and complex levels of flexibility. The importance of relational flexibility lies within the underlying behavioural process rather than flexibility with the content itself (Barnes-Holmes et al., 2001). The facilitation of flexibility in relational repertoires has been called for by a number of authors; specifically there is a need to develop programs which foster flexibility in relational responding in autistic populations whom characteristically display cognitive rigidity (Cassidy, et al., 2011; O’Hora et al., 2005, O’Toole et al., 2009). It has been suggested that children must learn to switch relational responding between like and different, correct and incorrect, more and less, and so on as many aspects of standard educational curricula require similar responses (Cassidy et al., 2011). The current study was successful at teaching five participants to respond in a flexible manner to “Like/Different” relational stimuli, as well correct and incorrect (consistent and inconsistent trials).

Participants’ ability to adapt to inconsistent trials shown in Study 3 represents added complexity in the level of relational flexibility required. During Study 3

participants were not only required to switch between coordination and distinction relations but also between inconsistent and consistent responding (responding correctly or incorrectly in accordance with changing contingencies). Again those who successfully maintained high levels of responding across both inconsistent and consistent trials were considered to have shown the greatest levels of flexibility. An examination of the inconsistent and consistent probes included in Study 3, readily reveals that Kevin demonstrated the greatest levels of flexibility maintaining high levels of accuracy across both inconsistent and consistent probes. As a brief measure of relational flexibility a “difference score” was calculated by subtracting the accuracy scores for inconsistent probes from consistent probes, lower difference scores represented greater flexibility. For example, if a participant correctly answered 20% of inconsistent trials and 80% consistent trials (a difference scores of 60), they would be considered less flexible than another participant who got 78% of inconsistent trials and 80% of consistent trials correct (a difference score of 2). Interestingly, difference scores correlated to measures of fluency and significantly predicted gains in K-BIT standard scores. Difference scores represent each participant’s ability to rapidly switch between two patterns of responding when contingencies change. This may suggest that greater flexibility in relational responding is related to greater fluency (total summed performances divided by total number of trial blocks), although the nature of this relation was needs more extensive exploration. Caveats notwithstanding, those who learned greater levels of relational flexibility showed greater gains across K-BIT standard scores. Taken together, it suggests that increasing fluency in flexible relational responding may be a potential avenue for increasing intelligent behavioural repertoires.

Effectiveness of Prompting Procedures

The data showed a range of competencies between participants with three of the five participants achieving mastery on stimulus sets 1 and 2, while two participants only achieved mastery on stimulus sets 1. It was also necessary to implement a number of specific alterations to the experimental sequence to remediate difficulties some participants experienced. These involved a gestural prompt procedure, colour coding the stimuli in a colour prompt procedure to increase the salience of the relations being taught and a block trials procedure which kept the target response constant over a number of trials to aid in the acquisition of the conditional discriminations required was also used (Slocum et al., 2012). Each of the alterations used proved to be effective at remediating the difficulties each participant was experiencing; the current findings provided support for their use as possible interventions for similar tasks

The implementation of a gestural prompt where participants received an additional prompt in the form of the researcher pointing to the correct answer for the first 50% of trials in each trial block proved to be an effective intervention to increase accuracy of responding. A gestural prompt procedure was used 10 times and was effective each time, with participants progressing to and subsequently reaching mastery on independent trial blocks following its implementation. The gestural prompt serves to increase the salience of the correct answer and was quickly faded allowing the stimulus to take control of the response.

Similarly the implementation of a colour prompt procedure which increased the salience of the relation until the target dimension acquired stimulus control proved to be an effective intervention also. The colour prompt procedure was used 7 times during the study and was effective 5 of those 7 times. The procedure added an

additional non-arbitrary visual cue to the stimuli increasing the salience of both “Like” and “Different”. The colour prompt procedure proved effective at highlighting the appropriate similarities and difference which then acquired stimulus control of the response once the colour prompt procedure was faded. Interestingly during Study 2 where arbitrary stimuli were used, it was necessary to add a colour prompt procedure to the experimental sequence before Kevin achieved mastery on the arbitrary stimuli. This colour prompt procedure effectively made the arbitrary stimuli non-arbitrary by adding a physical dimension (colour) along which Kevin could discriminate. This intervention was successful and the prompt was faded; subsequently Kevin achieved mastery with the purely arbitrary stimuli. These findings support those of Berens and Hayes (2007) that mastery on non-arbitrary stimuli may be a prerequisite for establishing arbitrary stimulus relations

Data indicated that an intervention with the blocked trials presentation procedure was very effective in facilitating acquisition of relational responding for two participants during Study 1. This involved presenting the same target relation repeatedly and had been previously shown to be effective at helping students learn conditional discrimination (Slocum et al., 2012). During blocked trials the same relations target was presented six times in a row before the second target was presented, for example six “Like”, followed by six “Different” and so on. This procedure was used twice and proved effective both times.

Accounting for Failure to Generalise a Previously Learned Response

With the exception of Kevin all participants required a large number of trials to reach mastery of even the most basic relations with non-arbitrary stimuli (e.g., Like/Different). This is of particular interest because a review of academic programmes showed recorded data for all participants indicating that they had met

the learning criteria for coordination relations, and records accorded with reports by staff. For example, all participants could discriminate identical and non-identical objects, and match written words to pictures of objects. Yet, when presented with a matching task involving physically identical and non-identical objects in the current study, participants could not discriminate Like/Different in either the TT or the T-IRAP conditions, and required extensive training before meeting the learning criterion. A number of factors may have hindered generalisation of the previously learned matching skill to the novel context. The procedure in the current study required a slightly different response topography (e.g., pressing a key on the keyboard during the T-IRAP); the stimuli in both teaching procedures were presented in a vertical as opposed to horizontal order, and the children may be more used to the latter, more common form of presentation. The word “Like” rather than “Same” was used. Moreover, it was found in the current study that some participants with a documented proficiency with “Same” stimuli required a large number of training trials when “Like” was taught alongside “Different” in the same trial-blocks. In other words for these children, learning to discriminate coordination relations as “Same” or “Like” may not have entailed a bi-directional derived response discriminating non-identical or unlike stimuli as “Different” or “Not-same”. This may suggest that children with learning difficulties may require explicit training for the constituent parts of a relational frame, for example teaching both “Different” as well as “Like” simultaneously in coordination relations, and teaching more-than and less-than simultaneously within a frame of comparison.

Generalisation has long since been a challenge within traditional ABA teaching and specific approaches have been developed to combat this for example incidental and natural environment teaching (Charlop-Christy & Carpenter, 2000).

There has been suggestion that highly structured learning environments and language programs commonly used in ABA may ultimately hinder generativity, and students frequently require specific “generalisation programs” following mastery of a target in one context (Chandler, Lubeck, & Fowler, 1992; Jones, Lerman, & Lechago, 2014; O’Connor et al., 2011). Studies that include a derived relational responding component have been thought to have potential utility in this area and may facilitate an effective technology for increasing generalisation across the board (Kilroe et al., 2011; Murphy et al., 2005, 2010).

A third related factor that may have influenced generalisation of the previously learned matching response to the current study is including both “Like and “Different” relations in the same trial block, in a rapidly and randomly changing sequence. This type of sequence requires a level of flexibility, to be able to adapt to the repeatedly changing contingencies, something which children with ASD may find difficult as rigidity is characteristic and symptomatic of ASD (Cassidy, et al., 2011; O’Hora et al., 2005, O’Toole et al., 2009). The combination involved in the current study may have been hindered by a significant learning history in which the participants were only required to respond to one target relation at a time.

Limitations and Future Research

For IQ measures, it was not possible to use an assessor who was blind to the purpose of the study, which would have been preferable for pre and post intervention IQ assessments to eliminate possibility of bias. Attempts to control for other sources of error were made by taking frequent inter-observer agreement measures (IOA), and high levels of IOA throughout all three studies suggest that sources of error in experimental data were kept to a minimum.

With regards to the limited sample size, statistical analyses showing significant differences must be viewed with caution as the sample may not be representative and generalisability of findings is therefore limited, nonetheless, the use of a single subject design (alternating treatments design) in which the effects are replicated on each occasion that the intervention changes provides a more robust measurement of behaviour change across conditions with individual participants. As such the conclusions drawn from the analysis of data using an alternating treatments design are both reliable and robust, and statistical analyses were seen as supplementary and thus provide additional support to that which is already apparent from visual analysis of the data. A number of participants achieved the lowest score possible on both IQ assessments, for example on the K-BIT all raw scores below a certain level are given a standard score of 40 regardless of individual differences. This posed a challenge to the current study as it may be the case that the participants varied in ways that the IQ tests were not sensitive to, and this problem has been recognised in particular with regard to populations with autism who may not be motivated to respond correctly. Also with a number of participants receiving highly similar IQ scores at both pre and post-test, Kevin's relatively large gains may have disproportionality skewed tests of statistical significance.

Future research would benefit from examining if RFT-based training programs impact on the generalisation of targets compared to standard interventions. An interesting feature of the data from the current study is that Kevin the only participant to achieve a criterion performance with arbitrary stimuli achieved the largest gains in IQ scores from pre to post intervention assessments. Further research is needed to fully explore this relation and to test if it is necessary to train arbitrary

stimulus relations in order to achieve large scale gains in IQ scores or can similar results be achieved with non-arbitrary relations alone.

Conclusion

The current findings provide support to previous claims that the T-IRAP is an efficient and effective method of teaching relational concepts, suggesting that the T-IRAP would be a highly beneficial supplementary teaching tool. The examination of the T-IRAP as a self-administered supplementary teaching tool to build fluency in the teaching targets may also prove beneficial in an applied sense. With regard to basic research, further investigation into relational flexibility with a greater number of relations up to an including analogy is needed, particularly with educationally disadvantaged learners.

The current study found that the T-IRAP may be a more effective and efficient method of stimulus presentation that can produce significantly faster response speeds and higher accuracies compared to TT. This supports the findings of Kilroe et al., (2011) which showed that participants responded with greater speed and accuracy during T-IRAP teaching procedures compared to TT. A further extension is that the current study taught more complex relational responding with a focus on relational flexibility using double contingency reversals with autistic children with extremely low IQ scores.

Flexibility in relational responding was found to significantly relate to measures of fluency and to gains in IQ scores from pre to post-intervention assessments. The findings suggest that training to fluency in flexible relational responding may positively impact on global cognitive functioning, providing further, although tentative support to the findings of Cassidy et al., (2011) that training in complex forms of relational responding can positively impact on cognitive ability as

measured by IQ testing. Research in this area, including the current study, is still highly preliminary and should be viewed with caution. The current findings, however, may represent another small step towards the development of an effective technology for increasing global cognitive abilities and intelligent behaviour with children with autism and learning deficits. Future research could aim to build upon the current findings by, examining if training in specific relational frames have a greater impact, for example, data from Cassidy et al. suggests that learning complex relational responding, but not more basic relational responding, resulted in improved IQ scores. In this regard future research could focus on examining the order, type and number of relational frames required to impact positively upon cognitive abilities.

References

- American Psychological Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text rev.)*. Washington, DC: Author.
- Azrin, N., & Foxx, R. (1971). A rapid method of toilet training the institutionalized retarded. *Journal of Applied Behavior Analysis, 4*(2), 89-99.
- Baer, D., Wolf, M., & Risley, T. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91-97.
- Barnes, D., McCullagh, P., & Keenana, M. (1990). Equivalence class formation in non-hearing impaired children and hearing impaired children. *Analysis of Verbal Behavior, 8*, 1-11.
- Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's Verbal Behavior: A possible synthesis. *The behavior Analyst, 23*, 66-84.
- 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 (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*, 169-177.
- Barnes-Holmes, Y., Barnes-Holmes, D., & Mchugh, L. (2004). Teaching derived relational responding to young children. *Journal of Early and Intensive Behavior Analysis, 1*(1), 3-13.
- Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., & Smeets, P. (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., & Smeets, P. (2001). Exemplar training and a derived transformation of function in accordance with symmetry: II. *The Psychological Record*, *51*, 589-604.
- Barnes-Holmes, Y., Barnes-Holmes, D., Roche, B., Healy, O., Lyddy, F., Cullinan, V., & Hayes, S. (2001). Psychological development. In S. Hayes, D. Barnes-Holmes, & B. Roche, *Relational Frame Theory: A post-Skinnerian account of human language and cognition* (pp. 157-181). New York: Plenum.
- Berens, N., & Hayes, S. (2007). Arbitrarily applicable comparative relations: experimental evidence for a relational operant. *Journal of Applied Behavior Analysis*, *40*(1), 45-71.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, *2*, 163-197.
- Brown, L., Sherbeno, R., & Johnsen, S. (1990). *Test of Nonverbal Intelligence-Revised*. Austin: Pro-Ed.
- Carr, E., Binkof, J., & Kologinsky, E. (1978). Acquisition of sign language by autistic children. 1: Expressive labelling. *Journal of Applied Behavior Analysis*, *4*, 489-501.
- Carroll, J. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language*. Circle Pines: American Guidance Service.
- Cassidy, S., Roche, B., & Hayes, S. (2011). A Relational Frame Training intervention to raise intelligence quotients: A pilot Study. *The Psychological Record*, *61*, 173-198.

- Chandler, L., Lubeck, R., & Fowler, S. (1992). Generalization and maintenance of preschool children skills: A critical review and analysis. *Journal of Applied Behavior Analysis, 25*(2), 415-428.
- Charlop-Christy, M., & Carpenter, M. (2000). Modified incidental teaching sessions; A procedure for parents to increase spontaneous speech in their children with autism. *Journal of Positive Behaviour Interventions, 2*(2), 98-112.
- Cohen, H., Amerne-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: replication of the UCLA model in a community setting. *Journal of Developmental and Behavioral Paediatrics, 27*(2), 45-55.
- DeLuca, R., & Holborn, S. (1992). Effects of a variable-ratio reinforcement schedule with changing criteria on exercise in obese and nonobese boys. *Journal of Applied Behavior Analysis, 25*(3), 671-679.
- Devany, J., Hayes, S., & Nelson, R. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 3*, 243-257.
- Dunn, L., & Dunn, D. (2007). *Peabody Picture Vocabulary Test* (4th ed.). San Antonio: Pearson.
- Grandin, T. (2008). *The way I see it: A personal look at autism and Asperger's*. Arlington: Future Horizons.
- Grant, D., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*, 404-411.
- Greer, J. (2002). *Designing teaching strategies: An applied behavior analysis systems approach*. San Diego: Academic Press.

- Greer, R., Keohane, D., & Healy, O. (2002). Quality and comprehensive applications of behavior analysis to schooling. *The Behavior Analyst Today*, 3, 120-132.
- Hall, G., & Sundberg, M. L. (1987). Teaching mands by manipulating conditioned establishing operations. *The Analysis of Verbal Behavior*, 5, 41-53.
- Hayes, S. (2011, December). *The implications of Relational Frame Theory (RFT) for Precision Teaching (PT)*. Retrieved January 31, 2013, from Centextualscience.org:
http://contextualscience.org/the_implications_of_relational_frame_theory_rft_fo
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (. (2001). *Relational Frame Theory: A Post-Skinnerian account of human language and cognition*. New York: Plenum Press.
- Hayes, S., Fox, E., Gifford, E., Wilson, K., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. Hayes, D. Barnes-Holmes, & B. Roche, *Relational frame theory: A post skinnerian account of human language and cognition* (pp. 21-50). New York: Plenum.
- Horn, J. L., & Cattell, R. B. (1966a). Age differences in primary mental ability factors. *Journal of Gerontology*, 21, 210-220.
- Horn, J. L., & Cattell, R. B. (1966b). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 53, 253-270.
- Howard, J., Sparkman, C., Cohen, H., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young

children with autism. *Research in Developmental Disabilities*, 26(4), 359-383.

Jensen, J., & Armstrong, R. (1985). *Slosson Intelligence Test (SIT) for children and adults: Expanded norms tables application and development*. East Aurora, N.Y.: Slosson Educational Publications.

Jones, J., Lerman, D., & Lechago, S. (2014). Assessing stimulus control promoting generalization via video modelling when teaching social responses to children with autism. *Journal of Applied Behavior Analysis*, 47(1), 37-50.

Kaufman, A., & Kaufman, N. (1990). *Kaufman Brief Intelligence Test (K-BIT)*. Circle Pines: American Guidance Service Inc.

Keohane, D., & Greer, D. (2005). Teachers use of verbally governed algorithm and student learning. *Journal of Behavioural and Consultation Therapy*, 246-259.

Kilroe, H., Murphy, C., & Barnes-Holmes, D. (2011). Using the IRAP computer programme and applied behaviour analysis to teach relational responding to children with diagnosed autism. *Unpublished doctoral thesis*.

Koegel, L., Koegel, R., & Carter, C. (1998). Pivotal responses and the natural language teaching paradigm. *Seminars in Speech and Language*, 19(4), 355-372.

Koffarnus, M., Wong, C., Fingerhood, M., Svikis, D., Bigelow, G., & Silverman, K. (2013). Monetary incentives to reinforce engagement and achievement in a job skills training program for homeless, unemployed adults. *Journal of Applied Behavior Analysis*, 46(3), 582-591.

Larsson, E. (2012). *Analysis of the Evidence Base for ABA and EIBI for Autism*.

Retrieved from www.Behavior.org:

<http://www.behavior.org/resources/649.pdf>

- Le Blanc, L., Esch, J., Sidener, T., & Firth, A. (2006). Behavioral language interventions for children with autism: Comparing applied verbal behavior and naturalistic teaching approaches. *The Analysis of Verbal Behavior*, 22(1), 49-60.
- Lewis, S. (2004). Using precision teaching to prepare students with learning differences for the SAT. *Journal of Precision Teaching and Celeration*, 20(1), 28-29.
- Lindsley, O. (1990a). Our aims, discoveries, failures, and problem. *Journal of Precision Teaching*, 7, 7-17.
- Lindsley, O. (1990b). Precision teaching: By teachers for children. *Teaching Exceptional Children*, 22(3), 10-15.
- Lindsley, O. (1991). Precision teaching's unique legacy from B. F. Skinner. *Journal of Behavioral Education*, 1, 253-266.
- Lindsley, O. (1992a). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis*, 25, 51-57.
- Lindsley, O. (1992b). Why aren't effective teaching tools widely adopted? *Journal of Applied Behavior Analysis*, 25, 21-26.
- Lindsley, O. (1995). Ten products of fluency. *Journal of Precision Teaching*, 13(1), 2-11.
- Lipkens, G., Hayes, S., & Hayes, L. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, 56, 201-239.
- Lovaas, O. (1977). *The autistic child: Language development through behavior modification*. New York: Irvington.

- Lovaas, O. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(3), 3-9.
- Lovaas, O. I., Berberich, J. P., Perloff, B. F., & Schaeffer, B. (1966). Acquisition of imitative speech in schizophrenic children. *Science, 151*, 705-707.
- Luciano, C., Gomez Becerra, I., & Rodriguez Valverde, M. (2007). The role of multiple- exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior, 87*, 349-365.
- Matson, J., & Coe, D. (1992). Applied behavior analysis: Its impact on the treatment of mentally retarded emotionally disturbed people. *Research in Developmental Disabilities, 13*, 171-189.
- McDade, C. (1992). Computer-Based precision Learning: A course builder application. *Behavior Research Methods, Instruments, & Computers, 24*(2), 269-272.
- McDade, C., Brown, J., & Vance, R. (1993). Can computer-based precision learning increase test scores? *Journal of Precision Teaching, 10*(2), 75-78.
- 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*(4), 445-462.
- Murphy, C., & Barnes-Holmes, D. (2010). Establishing complex derived manding with children with and without a diagnosis of autism. *The Psychological Record, 60*, 489-504.

- Nuzzolo-Gomez, R., & Greer, R. (2004). Emergence of untaught mands or tacts of novel adjective-object pairs as a function of instructional history. *The Analysis of Verbal Behavior*, 20, 63-76.
- O'Connor, J., Barnes-Holmes, Y., & Barnes-Holmes, D. (2011). Establishing contextual control over symmetry and asymmetry performances in typically developing children and children with autism. *The Psychological Record*(61), 287-312.
- O'Hora, D., Peláez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-II. *The Psychological Record*, 55, 155-175.
- Organisation, W. H. (1992). *International Statistical Classification of Diseases and Related Health Problems* (10th Revision (ICD-10) ed.). Geneva: WHO.
- O'Toole, C., & Barnes-Holmes, D. (2009). Three chronometric indices of relational responding as predictors of performance on a brief intelligence test: The importance of relational flexibility. *The Psychological Record*, 59, 119-132.
- Raybould, E., & Solity, J. (1988). Precision teaching ... and all that! *British Journal of Special Education*, 15(1), 32-33.
- Rehfeldt, R., & Barnes-Holmes, Y. (2009). *Derived Relational Responding Applications for Learners with Autism and Other Developmental Disabilities : A Progressive Guide to Change*. Oakland: New Harbinger Publications.
- Rehfeldt, R., & Root, S. (2005). Establishing derived requesting skills in adults with severe developmental disabilities. *Journal of Applied Behavior Analysis*, 38(1), 101-105.
- Rosales, R., & Rehfeldt, R. (2007). Contriving transitive conditioned establishing operations to establish derived manding skills in adults with severe

- developmental disabilities. *Journal of Applied Behavior Analysis*, 40(1), 105-121.
- Rosales, R., Rehfeldt, R. A., & Lovett, S. (2011). Effects of multiple exemplar training on the emergence of derived relations in preschool children learning a second language. *The Analysis of Verbal Behavior*, 27(1), 61-74.
- Sallows, G., & Graupner, T. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation*, 110(6), 417-438.
- Saunders, K. J., & Spradlin, J. E. (1993). Conditional discrimination in mentally retarded adults: Programing acquisition and learning set. *Journal of the Experiemntal Analysis of Behaviour*, 60, 571-585.
- Saunders, K., & Spradlin, J. (1989). Conditional discrimination in mentally retarded adults: The effects of training the component simple discriminations. *Journal of the Experiemntal Analysis of Behavior*, 52, 1-12.
- Saunders, K., & Spradlin, J. (1990). Conditional discrimination in mentally retarded adults: The development of generalized skills. *Journal of the Experimental Analysis of Behavior*, 54, 239-250.
- Semel, E., Wiig, E., & Wayne, S. (2003). *Clinical Evaluation of Language Fundamentals - Fourth Edition*. San Antonio: Psychological Corporation.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14, 5-13.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- Sidman, M. (2009). Equivalence Relations and Behavior: An Introductory Tutorial. *The Analysis of Verbal Behavior*, 25(1), 5-17.

- Skinner, B. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.
- Skinner, B. (1953). *Science and human Behavior*. New York: The Free Press.
- Skinner, B. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Slocum, S., Miller, S., & Tiger, J. (2012). Using a blocked-trials procedure to teach identity matching to a child with autism. *Journal of Applied Behaviour Analysis, 45*(3), 619-614.
- Smeets, P., Lancioni, G., & Striefel, S. (1987). Stimulus manipulation versus delayed feedback for teaching missing minuend problems to difficult-to-teach students. *Research in Developmental Disabilities, 261-282*.
- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 643-662*.
- Sundberg, M. (2004). A behavioral analysis of motivation and its relation to mand training. In L. Williams, *Developmental disabilities: Etiology, assessment, intervention and integration* (pp. 199-220). Reno: Context Press.
- Sundberg, M. L., & Michael, J. (2001). The value of Skinner's analysis of verbal behavior for teaching children with autism. *Behavior Modification, 25*, 698-724.
- Sundberg, M., & Partington, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25*(5), 698-724.
- Tincani, M. (2004). Comparing the Picture Exchange Communication System and sing language training for children with autism. *Focus on Autism and Other Developmental Disabilities, 19*(3), 152-163.
- von Hahn, E., & Murphy Bentley, B. (2013, December 9). *Executive function in children with ASD*. Retrieved from www.Autismconsortium.com:

<http://www.autismconsortium.org/blog/detail/executive-functions-in-children-with-asds>

- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children—Revised*. New York: Psychological Corporation.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.
- Wechsler, D. (2004). *The Wechsler Intelligence Scale for Children - Fourth edition*. London: Pearson Assessment.
- Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale-Third Edition*. San Antonio: The Psychological Corporation.
- White, O. R. (1986). Precision teaching—Precision learning. *Exceptional Children*, 52, 522-534.
- Williams, B., Myerson, J., & Halle, S. (2008). Individual differences, intelligence and behaviour analysis. *Journal of the Experimental Analysis of Behavior*, 219-231.
- Williams, K. (2001). *Group Reading Assessment and Diagnostic Evaluation*. Circle Pines: American Guidance Service.
- Williams, K. (2007). *Expressive Vocabulary Test, Second Edition*. Circle Pines: AGS Publishing.

Table of Appendices

Appendix 1:	Informed Consent Form For Participant In Doctoral Research	136
Appendix 2:	Parent Information Sheet	138
Appendix 3:	Continued Consent Form	141
Appendix 4:	Correspondence with Sigma Assessment Systems	142
Appendix 5:	APA Guidelines For Student Use of Psychometric Test	143
Appendix 6:	Samples of Stimuli Used	145

**Appendix 1: INFORMED CONSENT FORM FOR PARTICIPATION IN
DOCTORAL RESEACRH**

Title: Teaching flexible relational responding to children with diagnosed autism and/or developmental delay: A comparison of teaching IRAP with Table-Top procedures.

The current research will be conducted by Keith Lyons, B.A. (Hons) Psych., who is a doctoral student at the Department of Psychology, National University of Ireland, Maynooth, Co. Kildare. Keith has been a teaching assistant at NAME REMOVED in a school for children with diagnosed autism and intellectual disabilities for the past year, and can be contacted via telephone: 01 2819162, or email: Keith.j.Lyons@nuim.ie. The research will be supervised by Dr. Carol Murphy, B.A., Ph.D., B.C.B.A.-D. Dr Murphy is Course Manager on the Doctorate in Psychological Science at the Department of Psychology, NUI Maynooth, and can be contacted via telephone: 01 7086723 or email: Carol.A.Murphy@nuim.ie

In agreeing that my child participates in a research study carried out by a Doctoral student at the National University of Ireland, Maynooth, I _____ understand the following:

- In conducting the current research the student and supervisor are responsible for adhering to ethical guidelines set out by the Psychological society of Ireland and the Behaviour Analyst Certification Board in all dealings with my child.
- That the attached information sheet will tell me what procedures will be completed with my child. My child's identity will not be provided in any subsequent presentation or publication of data. All data will be assigned false names and will be stored in encrypted (protected) files on the researcher's computer for a period of five years after which the files will be deleted.
- My child will have their verbal ability assessed using subscales of two standardised tests (The Peabody Picture Vocabulary Test- fourth editions and the Kaufman Brief Intelligence Test). Individual results from the PPVT-IV and the K-BIT will not be made available to either the school or parents Should I (parent/caregiver) request access to the test results, I will be asked to make a formal written request and access will be provided (in accordance with current Freedom of Information legislation) with formal written advice from the researcher and supervisor that the test scores should not be used to guide clinical or other important decisions, because the researcher is insufficiently experienced to interpret test results for this purpose.
- If I have any concerns about my child's participation I understand that I may refuse consent to participate, or may withdraw my consent at any stage without any negative consequences for my child or me in either case.
- I will be reminded that I can withdraw my child's participation after a period of approximately 10 weeks The research study will use both Table-Top and

computerised teaching methods with the aim of teaching my child to respond flexibly to a number of relational concepts (such as above-below, same-opposite etc.) and to changing contingencies (e.g., match A with B but not C; subsequently match A with C but not B).

Important:

- If my child has a history of any seizure disorder or has experienced discomfort when viewing a computer or television screen, I should notify the classroom staff of this condition and consider carefully my child's involvement in this project.
- If I plan to have my child's ability or IQ tested for clinical reasons, the assessments to be carried out in this study may interfere with any other assessments carried out within the following 6-12 months. In this case I should exclude your child from participating in the research.

I confirm that I have read and understand the accompanying information sheet and that I agree to allow my child to participate in this study.

I understand that this research should not be considered to be a treatment of any description.

Signed:

_____ Participant/Parent

_____ Participant/Parent

_____ Researcher

_____ Date

Please note

Should you have any further questions please do not hesitate to contact either the researcher, Keith Lyons, at Keith.j.Lyons@nuim.ie or Dr. Carol Murphy, Department of Psychology National University of Ireland Maynooth, Maynooth, Co. Kildare (017086723) or at Carol.A.Murphy@nuim.ie

If during your 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 Chair of Departmental Ethics Committee, Dr. Bryan Roche. E-mail:Bryan.T.Roche@nuim.ie Please be assured that your concerns will be dealt with in a sensitive manner.

Appendix 2: Parent Information Sheet

Information about research being conducted at NAME REMOVED

Details about Researchers

Names: Keith Lyons, B.Sc. Psych.;
Doctoral Student

Address: ADDRESS REMOVED

Contact no. REMOVED

Email: Keith.j.lyons@nuim.ie

Details about Supervisor

Name: Dr Carol Murphy,
BCBA-D

Address: Department of Psychology,
NUI Maynooth

Contact no.: 01 7086723

Email: Carol.A.Murphy@nuim.ie

Please note that this research should not be considered to be a treatment of any description.

What is the purpose of the research?

Relational responding (e.g., same/different; more/less; before/after; I/You) is thought to be very important in cognitive and language skills. The current research aims to address the following questions: Are there correlations between relational responding abilities and verbal ability? Can children with diagnosed autism or developmental disorder learn rapid and flexible relational responding via a computerised teaching programme (T-IRAP) compared to a Table-Top teaching procedure?

What will the research involve for my child ?

Firstly, the Investigator will conduct assessments of your child's verbal ability using subscales of the Peabody Picture Vocabulary Test – fourth edition, and the Kaufman Brief Intelligence Test. These tests will take approximately 35 minutes to complete. Assessments will also be conducted to determine your child's ability to relate, for example, same/different stimuli, more/less, opposite. After the assessments, your child will be taught to relate stimuli, starting with same/different and gradually building up complexity (e.g., more/less, opposition and reversed relations in which the contingencies are changed (e.g., match A-B but not A-C, now match A-C but not A-B). A computerised teaching programme called the T-IRAP has been developed for the purpose, and learning will be compared using the T-IRAP and Table-Top

teaching. Positive reinforcement will be used throughout all teaching procedures, and frequent short breaks will be provided to your child to prevent boredom.

When will the research be conducted?

Research will commence in early October 2012 and will end no later than January 2014. Assessments will be conducted prior to and following teaching procedures and these will be conducted outside of school hours. Teaching sessions will follow the first assessments and will take place either during school or immediately after school; according to your preference. Assessment sessions will take approximately 1 hr. Teaching sessions will be conducted 2-3 times a week for approximately 40 minutes. The total estimated time for teaching sessions is expected to be 35 hours on average.

Where will the research be conducted?

The research will be carried out in the following schools: NAMES REMOVED, and will be carried out in your child's usual classroom setting with the classroom teacher present. If you prefer, all sessions can take place after school and you (parent) will be required to be present in this case.

What if I don't want my child to participate?

If you prefer not to volunteer your child's participation, please be assured that there is no obligation, nor is there a penalty of any kind for not participating. Furthermore, if you do consent to your child's participation, please be assured that we will monitor your child throughout to ensure that child participation is voluntary (s/he will be asked if s/he would like to work with the Researcher or to continue with other work) and that your child is not distressed in any way. If you consent now and later change your mind, please note that you are free to withdraw your child's participation at any time. In order to withdraw consent for your child's continued participation please contact the researcher immediately using the details above.

How will my child's data be kept safe?

Your child's data will be given a false name and will be stored in encrypted (protected) files on the researcher's computer with false names for a period of five years and only the researcher will have access to the data. All data will be encrypted

using Microsoft's "*encryption file system*". A code will be assigned to each child's data when the research commences. A key to identify the anonymised data will be stored on an encrypted computer hard drive should and this key to identifying your child's data will be destroyed when data analyses are complete. Individual results from the PPVT-IV and the K-BIT will not be made available to either the school or parents as doing so may result in clinical decisions being made based upon them. It is not the intention of this research project to guide any clinical or teaching decision. Should a parent request access to the test results, they will be asked to make a formal written request and access will be provided (in accordance with current Freedom of Information legislation) with formal written advice from the researcher and supervisor that the test scores should not be used to guide clinical or other important decisions because the researcher is insufficiently experienced to interpret test results for this purpose.

Thank you for taking the time to read this information sheet

Appendix 3: Continued Consent Form

Details about Researchers

Names: Keith Lyons, B.Sc. Psych.;
Doctoral Student

Address: ADDRESS REMOVED

Contact no. REMOVED

Email: Keith.j.lyons@nuim.ie

Details about Supervisor

Name: Dr Carol Murphy,
BCBA-D

Address: Department of Psychology,
NUI Maynooth

Contact no.: 01 7086723

Email: Carol.A.Murphy@nuim.ie

We would like to thank you for your cooperation with the current piece of research for which you have provided consent for your child to part-take in. At this point in the research programme, which we are approximately half way through we would like to make sure you are still comfortable with your child's continued participation. If you have any concerns please do not hesitate to contact the researcher using the above details. The researcher is always willing to answer questions you may have and will try to address any issues which may have arisen since the research has commenced. Given your child's progress through the first half of the research procedure we estimate that a further **X (approx.)** sessions will be required to complete the procedure. This is just an estimation of the number sessions required to complete the research procedure as it is difficult to predict the rate at which any child will learn. Please note that the research procedure will not carry on past the 31st of January 2014.

If you wish to *WITHDRAW* consent for your child's participation, Please sign below and return this form immediately. If you wish to allow your child to continue you do not need to do anything further.

Please only sign below if you wish to **WITHDRAW** your child from the study.

Signed:

_____ Participant/Parent

_____ Participant/Parent

_____ Researcher

_____ Date

Appendix 4: Correspondence with Sigma Assessment Systems

Dear Keith Lyons

Thank you for taking the time to contact SIGMA regarding possible training opportunities. Unfortunately, SIGMA does not offer training for learning how to administer and score level B assessments. It sounds as though you are already qualified for Level B. To score and administer Level C assessments, you may only need a supervisor to help with any interpretation problems you might have. Please complete the Qualifications form to be sure.

If you are interested in using any of our testing products for research, please complete the Application for Research Discount online as well as the Test User Qualifications form. Links to both forms can be found at

<http://sigmaassessmentssystems.com/departments/research.asp>.

If you have any further questions, please contact me.

Yours sincerely,
Sharon Van Duynhoven

SIGMA Assessment Systems, Inc.
PO Box 610757
Port Huron, MI
48061-0757
United States: 1.800.265.1285 x 222
International: 1.519.673.0833 x 222
svanduynhoven@SigmaHR.com
www.SigmaAssessmentSystems.com
www.SigmaHR.com

Research Psychologists Press, Inc.
PO Box 3292 , Station B
London, ON N6A 4K3
svanduynhoven@rpp.on.ca
www.rpp.on.ca
www.SigmaHR.com
Canada: 1.800.401.4480 x 222
International: 1.519.673.0833 x 222

Appendix 5: APA Guidelines For Student Use of Psychometric Test

Please see: <http://www.apa.org/science/programs/testing/qualifications.pdf>

With particular attention to the following points;

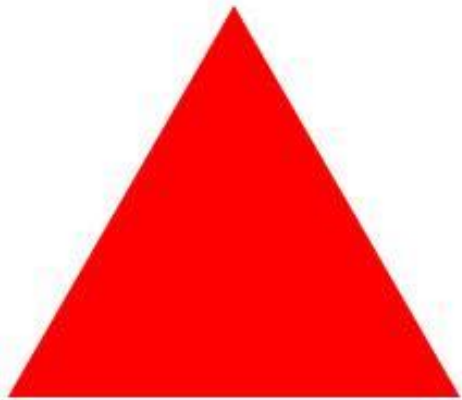
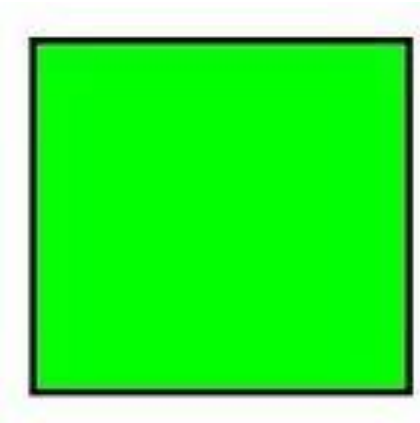
- “The APA’s purpose in developing these guidelines is to inform test users as well as individuals involved with training programs, regulatory and credentialing bodies, and the public about the qualifications that the APA considers important for the optimal use of tests. These guidelines describe two types of test user qualifications: (a) generic qualifications that serve as a basis for most of the typical uses of tests and (b) specific qualifications for the optimal use of tests in particular settings or for specific purposes. They are aspirational because they identify qualifications for the optimal use of tests in a competent and responsible manner. These guidelines describe qualifications that apply to a variety of testing settings and for multiple purposes; therefore, it is unlikely that a single test user possesses all the qualifications described here. The qualifications should also be considered in relation to the context, setting, and purpose of test use.” (p.8)
- “Various activities included in the testing process may be appropriately conducted by different people working collaboratively. Each participant should possess the knowledge, skills, and abilities relevant to his or her role. For example, different individuals may be responsible for deciding what constructs, conditions, or characteristics need to be assessed; selecting the appropriate tests; administering and scoring tests; and interpreting and communicating the results. Moreover, some testing activities may involve tasks that require limited professional knowledge (e.g., administering or scoring some tests, communicating simple test results). In such circumstances, test use should be directed by a qualified test user. It is this test user to whom these guidelines apply. Persons whose psychological test use is confined to research will find that the degree to which these guidelines apply to their work depends on their research focus and the research setting. The sections that address knowledge and skills in relation to psychometrics, statistics, test administration, and scoring are

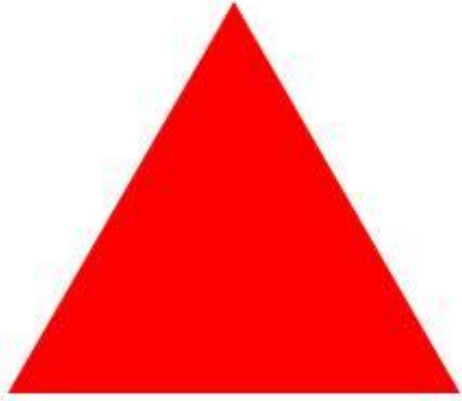
applicable to research that uses psychological tests.” (p.9)

- “This three-tiered system labelled some tests Level A (e.g., vocational proficiency tests) and designated them as appropriate for administration and interpretation by non-psychologists. The next level of tests (e.g., general intelligence tests and interest inventories) was labelled Level B. Qualifications for administering them included “some technical knowledge of Test construction and use, and of supporting psychological and educational subjects such as statistics, individual differences, the psychology of adjustment, personnel psychology, and guidance” (APA, 1950, p. 622). Over time, however, all those sanctioned “by an established school, government agency, or business enterprise” (APA, 1950, p. 622) were reclassified as eligible test users of Level B tests. Subsequent evidence suggests that those institutions did not provide the oversight necessary to ensure that these test users were in fact qualified (Eyde et al., 1993). Finally, qualifications for the use of Level C tests (e.g., individually administered tests of intelligence, personality tests, and projective methods) restricted their use to “persons with at least a Master’s degree in psychology, who have had at least one year of supervised experience under a psychologist” (APA, 1950, p. 622). The Level C qualification also had some exceptions. The reference to the three-tiered system was dropped from the 1974 (and subsequent). Standards without a replacement, but casual inspection of test publishers’ current catalogues reveals that it is still in widespread use (cf. Robertson & Eyde, 1986).” (p.11).

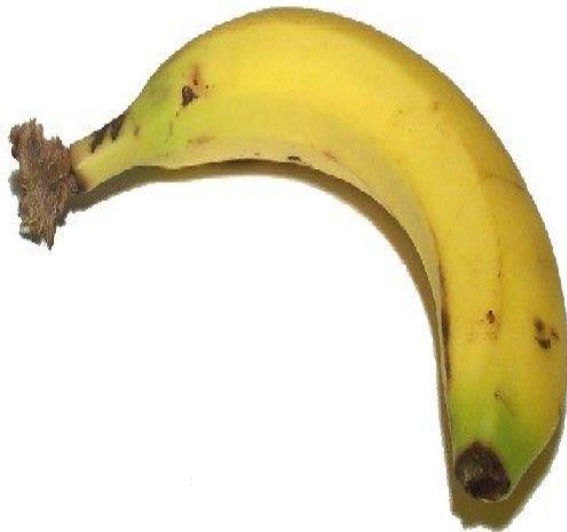
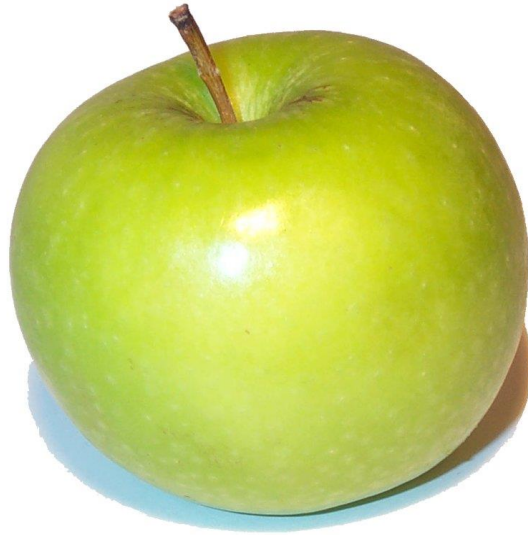
Appendix 6: Samples of Stimuli used

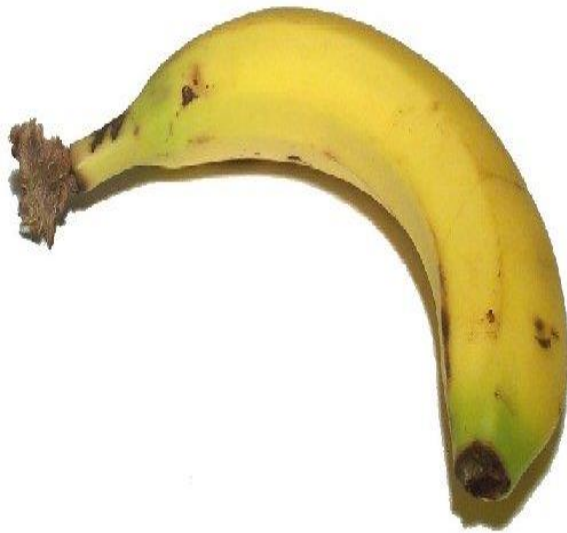
Stimulus set 1



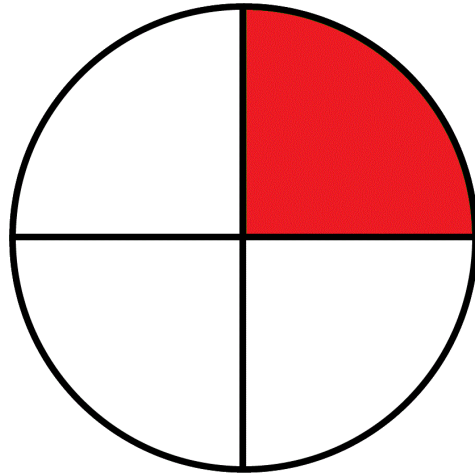


Stimulus set 2

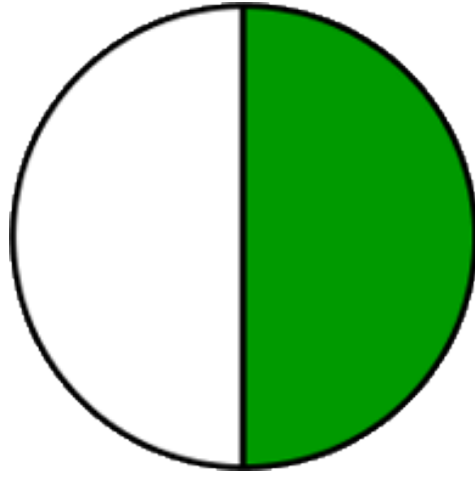




Stimulus set 3



50%



50%