Investigating the Role of Relational Responding and Relational Flexibility in Human Cognition

NUI MAYNOOTH
Ollscoil na hÉireann Má Nuad

Catriona O'Toole
Cert. Early Childhood Education, BA (psych)

Thesis submitted to the National University of Ireland Maynooth, in fulfillment of the requirements for a PhD in Psychology
October 2010

Supervised by Professor Dermot Barnes-Holmes
Department of Psychology

Head of Psychology Department: Dr Fiona Lyddy

CONTENTS

ABSTRACT

There is increasing recognition that relational processes are closely linked to, and may even provide the basis for many higher-order cognitive processes (e.g., Gentner, 2003b; Hayes, Barnes-Holmes, & Roche 2001). The primary aim of the current doctoral research was to explore the relationship between relational processes and human intelligence, focusing particularly on relational flexibility. The current research employed a relatively new methodology, called the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al, 2006) The IRAP is a computerised task, which requires participants to respond quickly and accurately in ways that are either consistent or inconsistent with previously learned relations. Response latencies are recorded on both consistent and inconsistent trials. A difference-score is also calculated by subtracting latencies on consistent trials from those on inconsistent trials. The difference-score therefore, provides a relatively "pure" measure of relational flexibility.

The current work comprises of four correlational studies. In Study 1 participants completed before/after and similar/different relational tasks, presented on the IRAP. They then completed the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990). Study 2 was a replication and extension of Study 1, in which, participants completed the same relational tasks, but were subsequently exposed to extensive and comprehensive cognitive abilities measures, including the WAIS-III^{UK} (Wechsler Adult Intelligence Scale – Third UK Edition: Wechsler, 1997). Study 3 involved the presentation of more complex relational tasks, including verbal and perceptual analogies and arithmetic. Study 4 returned once again to more basic relational frames. Specifically, in this study hierarchical and comparative relations were targeted, and subsequently the AH4 (Alice Heim 4; Heim, 1970) and the WAIS-III^{UK} were administered.

In general the results of the studies demonstrated that participants with higher scores on the intelligence test were not only faster at responding relationally, but also demonstrated a greater degree of relational flexibility than those with lower IQ scores. Interesting however, the more complex relational tasks presented in Study 3, produced fewer significant correlations with the intelligence measures than the relatively basic relational tasks. The results also indicated that particular types of relational frames predict performances on certain types of cognitive tasks.

Overall, the data obtained from the current research facilitate a greater understanding and greater specification of the processes underlying human intelligence. They also highlight the utility and sensitivity of the IRAP for investigating relational responding. Furthermore, the results suggest that targeting the fluid and flexible development of relational repertoires, may be crucially important in terms of promoting intelligent and creative behaviours in educational settings.

ACKNOWLEDGEMENTS

I would like to acknowledge and pay tribute to my supervisor, Professor Dermot Barnes-Holmes. Dermot introduced me to this exciting research area and with his incisive mind and clarity of thought he has been a constant source of inspiration. I am extremely grateful to the staff in the Psychology Department, especially Yvonne Barnes-Holmes, Anne Dooley, Victoria Thompson and Derek Walsh, whose help and support were invaluable throughout the course of my doctoral degree. I am also grateful to my fellow PhD students in the Department (both past and present) for the many laughs and the moral support. In addition, I would like to thank all those who volunteered as participants in this research, without whom the work could not have been completed. I owe a debt of gratitude to my parents, Peter and Catherine, who taught me patience, determination, and the value of a questioning mind – all qualities that were instrumental in my completion of this work. Finally, I am indebted to my husband, Andy, for his unwavering love and support, and my children, Freya and Alex, for ensuring I maintained a healthy work-life balance!

CHAPTER 1

CHAPTER 1

Introduction

For humans, the ability to identify relations between objects and events is an integral part of everyday life. Consider an individual who is attempting to solve a novel math problem. The solution will be more easily reached if the individual has encountered similar problems and relates what was previously learned to the current task. In this instance, the person is identifying a relation of similarity or coordination between a previous and current situation. The relation of similarity is a fundamental relational process, which can give rise to novel and sometimes extremely insightful ideas. It is widely reported, for example, that Newton conceived of the theory of gravity when he considered that the motion of an apple falling from a tree is somehow *similar* to the motion of the planets (e.g., Jameson & Gentner, 2003).

In addition to identifying similarities between stimuli and events, there are many other ways in which we respond relationally. For instance, we have the ability to navigate through our environments, to learn from analogies, to understand a sequence of events, and to plan for the future. Each of these abilities relies on relational thought, to the extent that Gentner and Loewenstein (2002) referred to such thought as the *sine qua non* of human cognition. Thus relational processes appear to underpin our proficiency in many cognitive abilities (Gentner, 2003, Hayes, Barnes-Holmes & Roche, 2001).

Behaviour-Analytic and Cognitive Approaches to Psychology

Investigating the mental processes underlying cognitive abilities is often considered the subject matter of cognitive psychology. However, increasingly these types of processes are of interest to researchers from the field of behaviour analysis. Traditionally the fields of cognitive and behavioural science have been viewed as standing in direct opposition to each other. Cognitive psychology places emphasis on mental processes, whereas the behaviour-analytic tradition attempts to identify functional relations between behaviour and environment. In doing so, behaviourism rejects mentalistic terms and concepts. As a result, it is often seen as uninterested in cognitive phenomena (e.g., Pinker, 2004), and this was indeed the case with early forms of behaviourism (e.g., stimulus-response psychology, see Barnes & Holmes, 1991). Modern behaviour analysts, in contrast, actively research many of the same phenomena as cognitive researchers (e.g., Hayes, et.al, 2001). Where there is a divergence between the two intellectual traditions, it is mainly with regard to the philosophy upon which their respective analyses are based (see Chisea 1994 for a detailed treatise).

Cognitive psychology predominantly operates according to a mechanistic world view, in which the structure of the mind is viewed as analogous to a complex machine and its functions are likened to computational processes (Morris, 1989). Accordingly, hypothetical constructs, such as working memory, central executives, and phonological loops, are employed. The utility of these constructs is determined by the extent to which they correspond with, and predict real world behaviours (Hayes, Hayes, & Reese, 1988).

In contrast, behaviour analysis has recently adopted a philosophy of science known as functional contextualism (Hayes, Blackledge & Barnes-Holmes, 2001). In psychology, the goals of functional contextualism are the prediction and influence of psychological events. These events refer to both observable behaviours (e.g., walking, hitting) and unobservable behavioural processes (e.g., feeling, and thinking; Biglan & Hayes, 1996; Hayes, 1993b; Gifford & Hayes, 1999). In achieving the goals of prediction and influence, a functional contextualist focuses on environmental and historical factors. Thus, while the cognitive psychologist appeals to hypothetical constructs, the behaviour analyst focuses on manipulable variables within the environment.

The differences between cognitive and behavoural approaches can be illustrated in the following example. Say for instance, a cognitive and a behavioural researcher were interested in accounting for children's reading difficulties. On the one hand, a cognitive researcher might highlight a link between reading difficulties and working memory deficits (e.g., Gathercole, Alloway, Willis, & Adams, 2006). There is utility in this analysis, because working memory deficits might explain and predict additional difficulties beyond those associated with reading (e.g., Pickering, 2006). However, a functional contextualist would not appeal to working memory as an explanation. This is not necessarily because he/she rejects the notion of a working memory construct, rather it is because a functional contextualist would view *both* reading difficulties*, and* working memory deficits, as psychological events, and thus appealling to one psychological event in order to explain another, would not be acceptable (e.g., Hayes, Stosahl, &

Wilson, 1999). Therefore, in attempting to explain and predict reading ability, the analysis would be traced back to the situational or historical context in which the reading difficulties emerged. Consequently, a functional analysis would be conducted in order to identify the contextual factors that impact upon reading. This approach also allows the functional contextualist to alter particular aspects of the context in order to influence future reading ability.

The approaches adopted by cognitive and behavioural psychology are different; but that does not mean that one approach is better or worse than the other. No particular philosophy can be considered true or correct (Kuhn, 1962). As highlighted by Barnes and Holmes (1989), a philosophy is a set of assumptions or values, and scientists typically adopt one over the other, based on what they feel represents the most convincing set of assumptions for analysing the questions they are asking.

Behaviour Analysis and the Study of Relational Responding

Many have doubted whether behavioural psychology, with its emphasis on contextual analysis, is capable of answering the types of questions that are of interest to cognitive psychologists (e.g., Pinker, 2002). Indeed, within behavioural psychology the dominant accounts of learning (classical and operant conditioning) failed to adequately explain higher-level cognition. One notable example is provided by Skinner (1957). Skinner drew on behavioural theories in order to account for human language. His work had important implications in applied areas, for example, in training developmentally delayed individuals to name (tact) and request (mand) items (e.g., Barnes-Holmes, Barnes-Holmes & Cullinan, 2000; Mac Corquodale, 1969). However, as an attempt to account for the generative and productive nature of language, it ultimately failed to convince academics in the wider discipline of psychology (Hayes & Berens, 2004).

The inability of the world's most prominent learning theorist to provide a plausible and much-needed explanation of human language precipitated a shift in focus towards cognitive accounts of learning and cognition. However, in the 1980s Sidman and colleagues developed a methodology for examining a particular type of complex relational responding, which is known as stimulus equivalence (see Sidman, 1994, for a review). It was the phenomenon of stimulus equivalence that paved the way for a progressive and proliferate behaviour-analytic investigation of the inherently relational nature of human language and cognition.

A typical equivalence experiment involves explicit training in a number of interrelated stimulus relations, which then give rise to a number of predicted but untrained relations. Imagine, for example, that a participant is presented with one of two sample stimuli along with each of two comparison stimuli. The participant is trained to choose one comparison stimulus (call it B1) when presented with a particular sample stimulus (call it A1), and to choose the second comparison (B2) when presented with the second sample (A2). Subsequently, when the participant is presented with B1 as sample, he/she will choose A1 as comparison, and when presented with B2 as a sample, he/she will choose A2 as comparison. Thus the two relations A1-B1 and A2-B2 are directly trained but the relations B1-A1 and B2-A2 are derived without explicit training. If two further conditional discriminations are then taught, B1-C1 and B2-C2, the number of relations that can be derived increases dramatically. In fact, training just two three-member relations, A1-B1/B1-C1, and A2-B2/B2-C2, will result in the emergence of the following eight derived relations: B1-A1, B2-A2, C1-B1, C2-B2, A1-C1, A2-C2, C1-A1, C2-A2. When these derived relational performances emerge, the stimuli involved are said to participate in an equivalence relation or a relation of sameness (Barnes, 1994; Sidman & Tailby, 1982).

Relational Frame Theory

In accounting for the phenomena of stimulus equivalence, Hayes (1991) provided a new analysis based on relational control. Specifically, he suggested that the stimulus equivalence phenonmenon outline above reflects an underlying pattern of relational responding on the basis of sameness or coordination. It was from this analysis that Relational Frame Theory (RFT; Hayes, et al., 2001) subsequently emerged. One reason for the development of this theory was that stimulus equivalence appeared to provide a behavioural analogue of the generative and productive nature of human language. Thus, it provided an avenue for the empirical investigation of language, which had not been available to Skinner when he published his 1957 account. Indeed, there is now substantive evidence in support of the view that equivalence and human language are closely related (see Barnes-Holmes, et al., 2004 for a review of the relevant evidence).

Another reason for the development of RFT was the recognition that, in addition to equivalence, there a many other ways in which stimuli can be related.

The term *relational frame* has been used to highlight these particular patterns of relational responding. A number of relational frames have so far been identified. These include frames of coordination, distinction, opposition, and hierarchy, as well as temporal, spatial, causal, and deictic relations. Empirical evidence in support of relational frames has also been offered (e.g., Dymond & Barnes, 1995; McHugh, Barnes-Holmes, & Barnes-Holmes, 2004; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGeady, 2000; Steele & Hayes, 1991). According to RFT, these families of relational frames play different roles in different types of verbal or cognitive behaviour.

Arbitrarily Applicable Relational Responding

RFT acknowledges that most species are capable of responding to relations among the physical properties of two or more stimuli. For example, mammals, birds, and even insects can all be trained to select the dimmer of two stimuli (Reese, 1968). Humans however are particularly proficient at identifying relations that extend beyond the formal properties of the relata. That is, humans can respond to objects and events where the relation between them is defined *not* by their physical properties, but by arbitrary contextual cues (i.e. cues that may be applied on the basis of social whim or convention). According to RFT this type of responding is termed arbitrarily applicable relational responding (AARR).

An early example of this type of responding is learning to name objects and events. For instance, a caregiver will often utter the name of an object in the presence of a young child and then reinforce any orienting response toward that object. Thus, upon hearing the word *Apple* a child will be rewarded for pointing to, looking at, or selecting an actual *Apple*. On other occasions the caregiver may present the apple to the child and then model and reinforce the appropriate response (i.e., the word *apple*). In the early stages of language development a number of these name-object and object-name exemplars will be explicitly trained. Gradually, however, the child no longer needs explicit reinforcement for each naming response. After a sufficient number of exemplars the child learns to *abstract out* the specific contextual cues as discriminative for the derived naming response and can therefore respond appropriately in novel instances. Thus, now when the child is presented with an unfamiliar object and told, "*this is a guava*, for example (i.e., when an object-name relation is trained), he/she will spontaneously reverse this relation without further training (i.e., provide a name-object relational response).

This example also serves to illustrate a number of features of AARR. First, the example highlights the arbitrary nature of these early language interactions. Clearly, there are no formal similarities between words and their referents (i.e. an actual apple and the word *apple* share no physical similarity). Thus the relation between them is applied arbitrarily on the basis of social convention. Second, the example highlights that derived relational performances come about as a result of multiple exemplar training. The child in the example was provided with multiple exemplars of the object-name and name-object relational instruction. With a sufficient number of exemplars, the process of relating becomes an overarching or generalised operant response class and extends to specific novel instances.

The literature on RFT has identified three defining properties of AARR; mutual entailment, combinatorial entailment, and the transformation of stimulus function (see Hayes, Fox, Gifford, Wilson, Barnes-Holmes, & Healy, 2001). Mutual entailment describes the relations between two stimuli or events. For example, if A in a specific context is related in a particular way to B, then a relation between B and A is also entailed in that context. Mutually entailed relations may or may not be symmetrical. For instance, if A is the same as B, then the derived mutually entailed relation between B and A is also one of sameness (i.e., B = A). However, if A is *more than* B, then a *less than* relation is entailed between B and A.

Combinatorially entailed relations pertain to three or more related stimuli. Consider the following example: ‘A is larger than B and B is larger than C’. In this case, a *larger than* relation is entailed between A and C, but a *smaller than* relation is entailed between C and A. Combinatorially entailed relations also differ from mutually entailed relations with regard to their specificity. For instance, ‘if A is smaller than B and A is smaller than C’, then the entailed relations between B and C and between C and B remain unspecified (i.e., B and C might be the same, or one might be smaller/larger than the other).

The transformation of stimulus functions is the behavioural property that provides stimulus relations with psychological content (Dougher, Hamilton, Fink, & Harrington, 2003). Consider again the ‘A larger than B’ example. If in certain contexts A acquires anxiety eliciting functions, then by virtue of the comparative relation, B will acquire reduced anxiety eliciting functions, relative to A. In effect, the different functions possessed by A and B are determined by the nature of the relation between those stimuli.

Another level of complexity provided by RFT is the concept of relational networks and the relating of such networks. The term network is used to describe relations between or among relational frames. For example, if A is more than B, and C is more than D, then the relation between A and B participates in a frame of coordination with the relation between C and D. Throughout our everyday interaction with the world, we continually develop relational networks, and subsequently, these networks of stimulus relations can be related to other networks. This ability to relate entire networks of relations underlies the development and use of analogies, metaphors, and other aspects of higher cognition (see Stewart, Barnes-Holmes, Hayes, & Lipkins, 2001).

Intelligence

In the first book-length treatment of RFT (Hayes, et al., 2001), a section was devoted to the domain of human intelligence (Barnes-Holmes, et al.). Specifically, the authors argued that “a small number of psychological processes are sufficient to yield the full gamut of cognitive skills” (p.160). In effect, therefore, RFT attempts to identify the core processes that underlie higher level cognition. However, before considering the implications of the RFT view of intelligence, we will first review briefly the dominant mainstream approaches to this area.

The concept of intelligence has long been one of the most controversial topics in psychology (Jensen, 1987). The controversy stems, not just from the emphasis on individual differences in cognitive ability, but also from claims of gender and race differences as well (Hernstein & Murray, 1994). Also divisive are issues surrounding the malleability (e.g., Wahlsten, 1997) and the measurement of intelligence (Gould, 1981). These controversies become all the more pronounced, given the findings that intelligence (as assessed by intelligence tests) appears to be an important predictor of critical life outcomes, including financial income, job performance, socioeconomic status (e.g., Hernstein & Murray, 1994), not to mention health, and life expectancy (Kilgour, Starr, & Whalley, 2010).

Currently, there are a number of different intelligence theories that dominate the literature. These can be loosely classified as biological models, hierarchical models, and complex systems models. The following is a very brief account of some of the main view points within each of these three camps.

Biological models

The biological models are based on the premise that highly intelligent people have brains that operate more accurately and efficiently (i.e., have greater neural efficiency) than less intelligent individuals (Davidson & Downing, 2000). In support of this idea, Haier and colleagues (1988) used positron emission tomography (PET) to measure cerebral glucose metabolic rates while individuals performed a range of cognitive tasks. The study found that individuals, who obtained higher IQ scores, had brains that expended less energy than individuals with lower IQ scores. Studies of nerve conduction velocity have also been conducted (e.g., Vernon & Mori, 1992). Nerve conduction velocity measures the transmission speed as electrical impulses travel form one part of the body to another. Faster conduction velocities were related to higher IQ scores.

Further support for the neural efficiency premise comes from Hendrickson (1982) who employed event related potentials to record the electrophysiological activity of the brain. Hendrickson developed a method of data analysis that essentially simulated a string being place on top of the wave form, generated from the event-related potential. Long string lengths reflected consistent and efficient electrical activity and were associated with higher IQ scores. It is worth noting, however, that significant correlations between string length and IQ has not been replicated consistently across studies (Davidson & Downing, 2000).

Hierarchical models.

Hierarchical models of intelligence are based on psychometric analyses. The assumption underlying these models is that the structure of intelligence can be discovered by analysing the inter-relationship of scores on cognitive abilities tests. Current hierarchical models have influenced the development of intelligence tests, and they can be best understood in terms of two earlier psychometric theories, one offered by Spearman and the other by Thurstone.

Spearman (1904) noted that there was a strong tendency for performances on various ability measures to be intercorrelated. He assumed there must be a common intellectual ability that accounted for the positive manifold of correlations, and he labelled this *g* for general intelligence. However, Thurstone (1938), using slightly different factor analysis techniques, found clusters of independent abilities rather than a unified *g*. His analysis led him to assume the existence of several primary ability factors, including verbal comprehension, number facility, spatial reasoning, memory, deduction, and inductive abilities.

Despite the diametrically opposed emphases of Spearman and Thurstone, the hierarchical models, which are dominant today, represent something of a reconciliation between the unity and diversity positions. Among the most dominant of these models was a theory proposed by Cattel (e.g. 1963) and later modified by Horn and Cattell (1966). Essentially this theory distinguishes between two types of ability: fluid intelligence or *gf*, and crystallised intelligence or *gc*. Fluid intelligence is viewed as the biologically influenced dimension of *g.* It is a kind of capacity or potential, which is most clearly manifest in novel, complex, or challenging environments. Cattell (1987) saw *gf* as the ability of an individual to apprehend the "complexity of relationships" and to act on them "when he does not have recourse to answers to such complex issues already stored in memory" (p115).

Crystallised intelligence, on the other hand, is believed to be influenced by education and culture (Cattell, 1987). It is intelligence that takes shape as a person acquires knowledge (often domain specific knowledge in verbal form) and integrates that knowledge with information already stored in long-term memory. Fluid and crystalised intelligence are functionally related in that *gf* is often understood to be the raw material for the formation of gc. Cattell (1987) suggests that *gf* can be considered a process, and *gc* as a product, resulting from *gf* and experience. The *gf-gc* theory has proven to be one of the most powerful organising theories in the intelligence literature (Sternberg, 2000).

The case for the hierarchical structure of intelligence was bolstered by the work of Carroll, published posthumously in 1993. Carroll conducted a massive analysis of over 460 data sets, and from this developed the three-stratum structure of intelligence. In his analysis, *g* is represented as the single third-order stratum at the apex of the hierarchy. Below *g* lie eight broad second-stratum abilities including *gf* and *gc*. Each of the second-stratum abilities is related to several narrow first-stratum dimensions of ability, which are represented at the base of the hierarchy.

Given the similarities between the *gf*-*gc* theory and the three-stratum theory, a synthesis has been offered by McGrew (1997, 2009) and McGrew and Flanagan (1998). This is commonly referred to as the Cattell-Horn-Carroll (CHC) theory. The integrated CHC model posits 10 broad stratum II ability domains. The recognition and influence of the CHC model has steadily increased over the past 10 years. Most of the major tests of intelligence have changed to incorporate CHC theory as their foundation for specifying and operationalising cognitive abilities/processes (McGrew and Flanagan 1998). McGrew (2009) expressed the hope that the language used to describe cognitive abilities and processes in the CHC model, would become the common nomenclature in the field of intelligence research. The CHC model represents a first effort to create a single *Gf*-*Gc* taxonomy for use in the evaluation and interpretation of intelligence

Complex Systems Models.

These models view intelligence as a complex system that includes interactions between mental processes, contextual factors, and multiple abilities

(Davidson & Downing, 2000). According to these models intelligence is dynamic, and can change when contextual conditions change. Most notable of these models is Howard Gardner’s (1993, 1998) multiple intelligences theory, which focuses on domains of ability rather than processes. Gardner rejects the notion of intelligence as a unitary ability. He suggests that there are at least eight fairly independent, equally important types of intelligences. Three of these intelligences - linguistic, logical-mathematical, and spatial -- are related to abilities measured by conventional intelligence tests. The remaining five types – musical, bodily-kinaesthetic, intrapersonal, interpersonal, and naturalist – are valued in most cultures but are not measured in tests of ability.

The bioecological model put forward by Ceci (e.g., 1996) proposes that intelligence is a function of the interactions between innate potential abilities, environmental contexts and internal motivation. Like Gardner, Ceci believed that there are multiple innate potential abilities that can be fostered by specific contexts. Interaction with environmental resources determines whether an innate cognitive potential will succeed or fail in its development. The model emphasises that an individual’s context helps shape and develop his or her biological predispositions, and in turn the predispositions help shape an individual’s context. This ongoing interplay between biological endowments and ecological contexts help shape a person’s development. This model also highlights that the timing of the interplay can be crucial. Some neural connections disappear if they do not receive specific contextual stimulation during sensitive periods of development. And when this happens certain potential abilities do not fully develop. Internal motivation is also considered important. According to Ceci, an individual must be motivated in order to capitalise on their innate abilities. Motivation can therefore contribute to uneven intellectual performances across domains.

Intelligence from an RFT Perspective

Each of the models, presented above, has advanced our understanding of intelligence. The hierarchical models provide a useful taxonomy of cognitive abilities, and although much of the research carried out to date has been descriptive, some also has predictive utility (Davidson & Downing, 2000). For instance, the *gf-gc* theory makes predictions about intellectual development across the life span, and the complex systems models highlight that contextual factors can have serious repercussions for intellectual development.

However, as highlighted by Davidson and Downing (2000), if these models have one common weakness, it is that they all require further specification. The biological models do not fully explain the relationship between the brain and behaviour (e.g., do faster conduction velocities cause higher IQ? Are they a consequence of higher IQ, or is there some other factor influencing both variables?). The hierarchical models classify the structure of cognitive abilities, but the descriptive terms lack precision, and they have failed to specify exactly what the *g* factor represents. The complex systems models have broad scope, in that they emphasise the range, malleability and complexity of intelligent behaviours. However, they too require further specification as it is not clear how these complex systems interact or what aspects of the context is important for the optimal development of cognitive abilities.

In contrast to the traditional approaches to intelligence, RFT adopts a "bottom-up" approach. In other words, the RFT approach to human language and cognition, including human intelligence, involves first specifying the basic behavioural or psychological processes involved in language and cognition. Once the basic processes have been defined, they can be used to construct behavioural accounts of specific domains, such as reasoning, intelligence, perspective-taking, and problem-solving.

In the book-length account of RFT, Y. Barnes-Holmes and colleagues (2001) provide an initial interpretation of intelligence from an RFT perspective. The authors highlight that the core psychological process which lies at the heart of the RFT analysis is arbitrary applicable relational responding. The relationship between AARR and verbal ability was mentioned earlier in this Introduction, and there is now a wealth of empirical data supporting this relationship (see Barnes-Holmes et al., 2004 for a review). Most intelligence tests contain subcomponents that assess verbal ability, such as tests of vocabulary, verbal analogies, or comprehension. Y. Barnes-Holmes and colleagues (2001) suggest that individuals who do well on these tasks have developed great proficiency in AARR, and as a result have highly elaborate relational repertoires. However, the authors point out that it is the relational process that is key rather than the verbal content per se. They go on to suggest that

"A task such as learning to spell is far less relationally rich than

> learning word meanings, and thus it is no surprise that spelling correlates less well with overall levels of intellectual behavior than vocabulary, even though both tasks involve verbal material."
> (p. 160)

Intelligence tests also tend to include sub-components that do not contain overtly verbal content. These are often termed non-verbal, perceptual reasoning, or performance-based tasks, and they typically include matrix reasoning problems, figure analogies, or reproducing a design using blocks. Although these tasks are ostensibly non-verbal, they too involve AARR. The completion of these tasks often involves responding to non-arbitrary features of the stimuli, but they also typically require the application of relational frames, as well as sensitivity to subtle forms of contextual control. RFT uses the term pragmatic verbal analysis to describe this type of problem-solving (Hayes, Gillford, Townsend, & Barnes-Holmes, 2001).

If AARR underlies these problem-solving tasks, as it underlies verbal skills, then we would expect that individuals with highly developed verbal abilities would also be reasonably proficient at "non-verbal" problem-solving tasks. Thus, perhaps it is no surprise that in psychometric tests, performances in these areas are often correlated -- they correlate because they tap into the same underlying relational skills. Thus, Barnes-Holmes and colleagues (2001) have suggested that this core ability of relational responding may, in fact, provide a behavioural, bottom-up, analysis of what is traditionally termed the general *g* factor, which is a major feature of the hierarchical models of intelligence.

RFT suggests that in addition to proficiency in AARR, intelligent behaviours also require the ability to relate relational networks with flexibility and relative ease. Specifically, Barnes-Homes and colleagues (2001) suggest that higher level cognition requires the "ability to elaborate entire networks of stimulus relations quickly, to bring them under increasingly subtle forms of contextual control, to transform stimulus functions through entire networks, and to abstract entire features of the natural environment what will support and sustain relational responding" (p. 161). Each of these skill sets is critical to the RFT analysis of intelligence.

It should be noted, of course, that other researchers have employed bottom-up approaches in their analysis of intelligence (see Deary, 1997 for an overview). However, these analyses have focused, to a large extent, on addressing the "simple" processing of sensory information, such as simple auditory and visual discriminations (Deary, 2000). The approach taken by RFT, which attempts to identify the underlying complex relational processes involved in intelligence, is therefore very different from those approaches that focus only on very simple discriminative abilities.

The research presented in the current thesis is based on the view that the approach offered by RFT will compliment those adopted by mainstream cognitive science. It should be possible that each set of analyses would offer something that could be utilised by the other. For instance, the RFT approach may provide the specification that is missing from mainstream approaches. The cognitive approaches can and have influenced behavioural understandings of the different domains and structures of intelligence (Barnes-Holmes, et al. 2001). Thus, it is likely, if both sets of analysis are on the right track, that the top-down approach provided by cognitive science and the bottom-up approach offered by RFT, might meet, and when this occurs, a more thorough and integrated account of human intelligence may be possible.

At the current time, the RFT contribution to the study of intelligence is extremely limited, and researchers have only just begun to test the basic predictions made by Barnes-Holmes et al. (2001). Specifically, two recent studies have sought to determine if proficiency in the complex relational processes outlined above, actually predict performances on traditional measures of intelligence. Contributing to this research programme is the primary objective of the current thesis. In order to provide the specific context for the current research, it will be useful to review the two earlier studies, which sought to investigate the relationship between derived relational responding and intelligence.

In the first of these studies, O’Hora, Pelaez, and Barnes-Holmes (2005) compared participant performance on derived relations tasks relative to their performance on a number of subtests of the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III; Wechsler, 1997). Specifically, participants were presented with a complex relational task in which they had to learn to relate novel stimuli in accordance with similar/different and temporal relations. Results showed that participants who passed a subsequent test for the predicted novel patterns of relational responding produced significantly higher scores on the Vocabulary and Arithmetic subtests (but not on the Digit-Symbol Coding subtest) than the participants who failed the relational test. Furthermore, the data also showed a significant correlation between the number of accurate responses on the relational training phase and the number of correct responses on both the Vocabulary and Arithmetic subtests of the WAIS-III. The pattern of results was predicted given that both the Vocabulary and the Arithmetic subtest would be considered relationally rich, whereas the Digit-symbol Coding subtest, which assesses processing speed, relies less heavily on relational processes.

The second study to investigate the relationship between relational responding and intelligence involved a replication and extension of O'Hora, et al.'s (2005) work. Specifically, this study (O'Hora, Pelaez, Barnes-Holmes, Rae, Robinson, & Chaudhary, 2008) involved presenting participants with a temporal relations task as well as administering each of the subtests that comprise the four indices of the WAIS-III. The study found that successfully completing the temporal relations task was predictive of better performances on the Verbal Comprehension and the Perceptual Organisation Indices, but not on the Working Memory or Processing Speed indices.

In analysing the results, O'Hora and colleagues (2008) argued that the correlations with the Verbal Comprehension index further supported the claim that language and relational responding are closely linked. Furthermore, in accounting for the correlations observed with the Perceptual Organisation subtests, the authors drew on the concept of pragmatic verbal analysis (outlined above). Thus, the items contained in these subtests often contain non-arbitrary stimuli but the completion of the tasks typically involves verbal analysis or the arbitrary application of relational frames.

Both these studies lend support to the position that AARR lies at the heart of complex cognitive abilities. However, proficiency at relational responding is likely to involve additional skill sets. In their interpretation of intelligence, Barnes-Holmes and colleagues (2001) suggested that the behaviours typically referred to as intelligence require speed, flexibility, and subtlety of contextual control over relational responding. The methodologies employed in the two studies reported by O’Hora, et al., (2005, 2008) were not particularly sensitive to these additional dimensions. For example, the relational tasks employed by O’Hora et al. did not impose a speed requirement on participant’s responding and nor did they require participants to demonstrate flexibility in contextual control over their relational performances. This limitation in the research conducted by O’Hora et al. can only be addressed by considering the methodologies employed in studying relational responding, and it is to this issue that we now turn.

Methodologies for Investigating Relational Responding

A number of methodologies have been employed to examine stimulus equivalence and other types of relational responding. Two of the most widely used methodologies include the matching-to-sample (MTS) paradigm and the Relational Evaluation Procedure (REP). We will briefly outline these methodologies together with their advantages and limitations. The rationale for a third methodology, termed the Implicit Relational Assessment procedure (IRAP), will then be introduced.

Early research into the formation of stimulus equivalence classes typically involved the use of MTS procedures (e.g. Sidman & Tailby, 1982). An MTS task typically involves presenting a single sample stimulus along with two comparison stimuli on a computer screen. The sample and comparison stimuli typically comprise of nonsense syllables, but for ease or communication they will referred to using alphanumeric labels. During the training phase of the MTS task, differential reinforcement is provided for choosing stimulus B1 in the presence of A1, and for choosing B2 in the presence of A2. In this way, the following four MTS performances may be established: A1 → B1, A2 → B2, B1→ C1, B2 → C2. After a number of these training trials, participants are typically exposed to an equivalence test. During this phase, probes are presented to test for the emergence of untrained or derived responses. Thus, when presented with C1 as a sample stimulus, and A1 and A2 as comparisons, participants will often select the A1 stimulus, even though this relation has never been explicitly taught. When these derived relational responses emerge they are often described as equivalence responding.

Although MTS procedures are effective for investigating aspects of derived relational performances, they are also characterised by a number of limitations. For instance, the MTS format does not lend itself readily to the investigation of other types of relational frames beyond those of co-ordination or equivalence. Furthermore, some researchers have argued that the MTS procedure is rather cumbersome, in that many verbally sophisticated individuals, who would undoubtedly be capable of deriving relations of co-ordination, fail to demonstrate the predicted derived performances in the laboratory (e.g., Barnes-Holmes, Hayes, Dymond, & O'Hora, 2001).

In response to these and other limitations, the REP was developed (Cullinan, Barnes & Smeets, 1998; Hayes & Barnes, 1997). The REP allows participants to evaluate, or report on, the stimulus relation presented on a given trial. For illustratation purposes, consider an individual who receives training on temporal (*Before* and *After*) relations. On each training trial, two arbitrary stimuli are presented, one after the other (A1 followed by B1), in the middle of the computer screen. Shortly after the presentation of A1 and B1, two three-element comparison stimuli (referred to as statements) appear on the screen, one in the lower left-hand corner and the other in the lower right-hand corner. The statements contain the A1 and B1 stimuli and an arbitrary contextual cue (e.g., XXX or VVV). Differential reinforcement is then provided such that the XXX stimulus is established as functionally equivalent to the relational cue "BEFORE"; and the VVV stimulus is established as functionally equivalent to the cue "AFTER". Thus the participant is required to select one of the two statements and receives corrective feedback for selecting A1 XXX B1 (where A1 was physically presented before B1 on the training trial); and for selecting B1 VVV A1 (where B1 was physically presented following A1 in training). When the BEFORE and AFTER cues have been trained in this way, they can subsequently be tested using new stimuli.

The REP addresses many of the difficulties associated with the MTS paradigm. It allows for the investigation of stimulus relations beyond those of equivalence, and it is a more flexible and participant-friendly procedure. Indeed, the REP was used in the aforementioned research by O'Hora and colleagues (2005, 2008), which focused on the relationship between relational responding and intelligence. Specifically, in these studies, the REP was employed as a methodology to facilitate the derivation of new relations between previously unseen stimuli. Thus, the methodology probed for proficiency in AARR.

As noted above, we are unlikely to fully understand the importance of relational responding in intelligence by studying AARR in isolation. Thus, Barnes-Holmes and colleagues (2001) argued that a more complete understanding of the behavioural processes involved in human cognition, necessitates the consideration of additional variables. Specifically, speed and flexibility, as well as the subtlety of contextual control over relational responding, will require systematic analysis. It would be useful, therefore, to employ a methodology that taps into at least some of these additional processes.

The Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Barnes-Holmes, Power, Hayden, Milne, & Stewart, 2006; McKenna, Barnes-Holmes, Barnes-Holmes, & Stewart, 2007) represents the first step in the development of such a methodology. The IRAP is a computer procedure which requires participants to respond to a series of relational tasks. Unlike the MTS and the REP procedures, which only provide a measure of the accuracy of relational responding, the IRAP provides a measure of both the accuracy *and* the speed at which participants respond to the various relational tasks. In addition, the IRAP requires two patterns of responding, one consistent and the other inconsistent with previously established relations. The general assumption is that participants will respond more rapidly on consistent than on inconsistent trials, and this prediction has been supported across numerous studies (e.g., Barnes-Holmes et al., 2006). As will be argued in the next chapter, response speed on the inconsistent trials may provide a possibly useful measure of relational or cognitive flexibility.

The measures provided by the IRAP mean that it is a potentially rich methodology for a bottom-up analysis of intelligence. Specifically, the three measures of relational responding provided by the IRAP -- accuracy, speed, and flexibility -- overlap with those additional domains deemed important, by Barnes-Holmes and colleagues, for an RFT analysis of human cognition. A full description of the IRAP will be provided in Chapter 2. What follows is a rationale for the current work and an overview of the subsequent chapters that comprise the thesis.

The Current Thesis

The purpose of the current work was to explore the role of relational responding and relational flexibility in human intelligence, from a functional analytic perspective. The IRAP was employed as a methodology because of its sensitivity to the important domains of speed and flexibility (in addition to accuracy). The IRAP tasks employed across the current studies targeted a number of relational frames including co-ordination, distinction, comparison, and analogies as well as temporal and hierarchical frames. Performances on the relational tasks were then correlated with performances on traditional, standardised intelligence tests. The studies aimed to elucidate some of the core processes that may underpin various domains of human intelligence.

The thesis is comprised of four correlational studies. In Study 1 participants completed similar/different and before/after relational tasks, and subsequently completed a brief intelligence test. This study is presented in Chapter 2. Chapter 3 presents the second study of the research programme, which involved presenting the same two relational tasks as in Study 1, and then administering the WAIS-III^{UK} (Wechsler Adult Intelligence Scale – Third UK Edition: Wechsler, 1997), the AH4 (Alice Heim 4; Heim, 1970), and the Cognitive Failures Questionnaire (CFQ; Broadbent, Cooper, & Fitzgerald, 1982). The third study of the current programme is presented in Chapter 4. In this study participants completed more complex relational tasks than those presented in the two previous studies. They subsequently completed the Cognitive Abilities Test-Third Edition (CAT-3; Lohman, Thorndike, & Hagan, 2001). The final study, presented in Chapter 5, involved the presentation once again of relatively basic relational tasks – this time bigger/smaller and hierarchical relations – followed by administration of the WAIS-III^{UK}. Chapter 6 presents a general discussion of various issues that arose from the research programme. In this final chapter, relevant conceptual issues as well as the broader implications of the current work are discussed.

CHAPTER 2

Study 1: The Relationship Between Similar/Different and Before/After Relational Frames and Performances on a Brief Intelligence Test

Two studies have previously tested the RFT prediction that relational responding is an important aspect of human intelligence (O'Hora et. al., 2005; 2008). Both of these studies employed the Relational Evaluation Procedure (REP). Study 1 of the current research programme also sought to test the relationship between relational responding and intelligence. The study focused on the same relations as those targeted by O'Hora et al. (2005, 2008) but the IRAP was used instead of the REP. Given that this was the very first study to employ the IRAP in this particular way, a relatively brief intelligence test was used, but one that has been shown to correlate with the test employed by O'Hora et al. (the WAIS). The basic prediction, therefore, was that performance on the IRAP would correlate with performance on the intelligence test. If such as prediction was confirmed, only then would it be wise to invest additional resources into employing a more complete intelligence test, such as the WAIS.

As highlighted in the Introduction, the IRAP involves presenting participants with a series of relational tasks. In some blocks of tasks, participants are asked to respond in a way that is consistent with previously learned relations (e.g. "is a Shoe similar to a Sandal? = True"), and in other blocks responding in an inconsistent pattern is required (e.g. "is a Shoe similar to a Sandal? = False"). The

IRAP requires that participants respond accurately and rapidly across blocks of trials, and measures of both these variables are recorded. Recording response speed, in addition to accuracy, seems particularly relevant, because speed of responding may provide an indication of the fluidity or fluency with which individuals can respond relationally (e.g. Merbitz, Vieitez, Merbitz, & Pennypacker, 2004).

Another advantage to using the IRAP as a relational task is the fact that it requires two patterns of responding, one consistent and the other inconsistent with previously established relations. Typically individuals will respond more rapidly on consistent than on inconsistent trials (e.g., Barnes-Holmes et al., 2006). Critically, however, performance on the inconsistent trials is unlikely to be a well practiced or firmly established skill (i.e. individuals rarely practice incorrect responding for protracted periods of time). Consequently, response speed on these trials may provide a useful measure of relational or cognitive flexibility. That is, the faster an individual can produce responses that contradict previously well-established verbal relations (by the wider social community), the more flexible the behaviour. Assessing relational flexibility may be particularly advantageous because flexibility is widely regarded to be an important component of human cognitive abilities (e.g., Cattell, 1971; Kyllonen, Lohman, & Woltz, 1984; Premack, 2004).

On balance, the response latency data obtained on inconsistent trials may provide a measure of not just relational flexibility, but also other variables, such as processing speed. If we wish to obtain a relatively “pure” measure of relational flexibility, then it is necessary to control for the effects of these extraneous variables. Insofar as these variables contribute to both types of trials, the difference between the two, known as the difference-score, may thus provide an uncontaminated measure of relational flexibility (i.e., the smaller the difference-score between consistent and inconsistent trials the greater the relational flexibility). It is also worth noting that the IRAP requires participants to switch back-and-forth repeatedly between blocks of consistent and inconsistent trials, and thus the difference-score should reflect the relational flexibility that is produced across the numerous shifts in the IRAP contingencies.

In the current study, participants were presented with two IRAP tasks, one of which assessed relations of coordination or difference (i.e. similar/different) and the other of which assessed temporal (before/after) relations. Participants were subsequently exposed to the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990). Insofar as relational responding and flexibility are defining features of human intelligence, we would predict that participants with higher scores on an intelligence test would respond more quickly on the relational tasks than those with lower scores, particularly on inconsistent relative to consistent trials. Furthermore, we would also predict that smaller difference-scores, indicating greater relational flexibility, would predict higher IQ scores.

Method

Participants

Participants were 62 (15 male, 47 female) undergraduate students from various faculties of the National University of Ireland, Maynooth. None of the participants had previous exposure to the IRAP protocols used in the current study, nor were they familiar with the IQ test that was subsequently administered. The data from one female participant, who reported being dyslexic on a subject profile form (see below), was not included in the final analyses. In addition, to control for any confounding effects that may occur due to a speed /accuracy trade-off, only data from participants who produced at least 80% correct responding on the IRAPs were included in the final analyses. Fifty-five participants met this criterion on one or both of the IRAPs. They ranged in age from 18-55 years (Mode = 18; Mean = 23).

Materials and Stimuli

Materials included the IRAP and a subject profile form. The IRAP was presented on a Dell Personal Computer with Pentium 4 Processor and a standard keyboard and monitor. The software was used to present the instructions, stimuli and to record responses. Two separate IRAP tasks were presented that differed only in terms of the stimuli that were employed. Details of the stimuli used in each of the IRAPs are contained in Table 1.

The K-BIT is a brief, individually administered measure of verbal and nonverbal intelligence suitable for use with individuals aged 4-90 years. It consists of two subtests, a Vocabulary subtest (consisting of Part A, Expressive vocabulary and Part B, Definitions) and a Matrices subtest. The Vocabulary subtest measures verbal, school-related skills by assessing a person’s word knowledge and verbal concept formation. The Matrices subtest measures nonverbal skills and the ability to solve new problems by assessing a person’s ability to perceive relationships and complete analogies. All Matrices items involve pictures or abstract designs rather than words. Additional materials included the K-BIT Individual Test Record Form, and a stopwatch to ensure a maximum of 30s was allowed for each item on the Definitions task. Kaufman and Kaufman (1990) report a correlational coefficient of .75 between K-BIT Composite IQ scores and Full Scale IQ scores on the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981). Furthermore, the Vocabulary subtest of the K-BIT correlates .60 with Verbal IQ on the WAIS-R, while the Matrices subtest correlates .52 with Performance IQ on the WAIS-R. These coefficients, it has been argued, provide strong support for the construct validity of the K-BIT (Kaufman & Kaufman, 1990). Excellent split-half reliability coefficients are also reported for the K-BIT IQ Composite as well as for the subtests. The reliability of the K-BIT IQ Composite, for instance, averages .93 across the entire age span.

Table 1. *Sample and Target Stimuli Used in the Before/After and the Similar/Different IRAPs.*

Before/After IRAP		**Similar/Different IRAP**	
Sample 1	**Sample 2**	**Sample 1**	**Sample 2**
Before	After	Similar	Different
Targets congruent with Sample 1	**Targets congruent with Sample 2**	**Targets congruent with Sample 1**	**Targets congruent with Sample 2**
Spring Summer	Summer Spring	Oven Grill	Oven Tree
Child Adult	Adult Child	Book Journal	Book Dog
Engagement Marriage	Marriage Engagement	Wall Fence	Wall Parrot
Crawl Walk	Walk Crawl	Chair Seat	Chair Lion
Problem Solution	Solution Problem	Table Desk	Table Cat
Effort Reward	Reward Effort	House Cottage	House Star
Treatment Cure	Cure Treatment	Boat Ship	Boat Shelf
Introduction Conclusion	Conclusion Introduction	Shoe Sandal	Shoe Apple

Procedure

The experiment was split into two sessions. In Session 1, participants completed the two IRAP protocols. In Session 2 they were exposed to the K-BIT. The sessions were conducted on separate days but within one working week of each other. Both sessions were conducted in individual cubicles in the experimental psychology laboratory at NUI Maynooth.

Session 1: IRAPs

Upon arrival at the laboratory, participants completed a subject-profile form, which contained a series of questions on demographic information pertaining to each participant (i.e. age, gender, ethnicity, education level/occupation). They were also asked to state how often they use computers (whether daily, weekly, monthly, annually or never), and whether or not they had any learning difficulties.

After filling out this form participants were directed towards a display page that was open on the computer screen and which contained the instructions necessary for completing the similar/different or the before/after IRAP tasks (see Appendix 1). The display page contained some general instructions and a consent statement. It also contained specific instructions and illustrations, designed to explain how participants should respond to the tasks. The experimenter verbally explained the nature of the task with the aid of these illustrations. For example, for participants who began the experiment with the before/after IRAP, the experimenter referred them to the first illustration on the before/after instruction page. This illustration contained the word '*before*' at the top of the display, the words '*spring summer*' in the centre, and the words '*true*' and '*false*' at the bottom. Participants were instructed to read this as '*spring before summer*' and then to respond appropriately by pressing the designated *true* or *false* key. Participants were also informed that sometimes they would be required to respond to the tasks in a way that agrees with what they believe, and at other times they would be required to respond in a way that disagrees with what they believe.

Consistent IRAP blocks were defined as those that required responses that were deemed to be generally consistent with common verbal practices (e.g., choosing *true* when presented with the target, *spring before summer*). Inconsistent IRAP blocks required responses that were inconsistent with common verbal practices (e.g., choosing *true* when presented with the target, *summer before spring*). Participants were informed that the program would alternate between the two types of blocks, and that the first two blocks of trials were for practice, but thereafter they should respond as quickly and as accurately as possible. Finally, they were instructed that they should keep their left and right index fingers resting on the ‘d’ and ‘k’ keys, respectively, given that these keys corresponded to the true and false response options (see below). The experimenter then left the room and the participants were free to scroll through the instruction page at their own pace. They were required to press a button on the computer screen when they were sure that they fully understood the task. Having pressed this button, another display page was presented which stated that when an error was made a red X would appear below the target stimulus, and that the correct response must be emitted to continue. The next part of the message invited the participant to press the space-bar to proceed with the tasks.

Each IRAP consisted of two blocks of 32 practice trials and six blocks of 32 test trials. Each participant was randomly assigned to one of two conditions, *consistent-relations-first* (CF) or *inconsistent-relations-first* (IF). Participants assigned to the CF condition began both IRAPs with a consistent trial block and thereafter alternated between inconsistent and consistent blocks. Thus the experimental sequence for the CF group was: consistent practice, inconsistent practice, consistent test 1, inconsistent test 1, consistent test 2, inconsistent test 2, consistent test 3, and inconsistent test 3. Participants assigned to the IF group started both IRAPs with an inconsistent block and subsequently alternated between consistent and inconsistent blocks. The order in which the two IRAPs (similar/different and before/after) were presented was counterbalanced across participants.

Each trial on each of the IRAPs involved presenting a sample stimulus and a pair of target stimuli. Two response options, *true* and *false*, were also presented on each trial. The sample stimulus was presented at the top of the computer screen, the target words were presented in the center, and the *true* and *false* response options were presented in the bottom left and right hand corners. The response options switched sides unpredictably across trials. The various combinations of sample stimuli with target stimuli served to generate four different trial-types. Specifically, one trial-type was generated by presenting Sample 1 with congruent targets (e.g. *Before* with *Spring Summer*); another by presenting *Sample 2* with congruent targets (e.g. *After* with *Summer Spring*); a third by presenting Sample 1 with *incongruent* targets (e.g. *Before* with *Summer Spring*); and a fourth by presenting *Sample 2* with incongruent targets (e.g. *After* with *Spring Summer*). For each block of 32 trials, each target stimulus was presented in a quasi-random order, such that each target was presented twice, once in the presence of each sample stimulus (i.e., the four trial-types were presented eight times).

Figure 1 illustrates examples of the four different trial-types on both the before/after (top section) and the similar/different (bottom section) IRAPs as they would appear on the computer screen. All of the stimuli (i.e., the relational term, target stimuli and the response options) appeared simultaneously and remained on the screen until the participant emitted a response (i.e. pressed either the ‘d’ or ‘k’ key). If the participant emitted a correct response (i.e., consistent during consistent blocks, and inconsistent during inconsistent blocks), the screen cleared and a new trial was presented after a 400ms delay. Following an incorrect response (i.e., consistent during inconsistent, and inconsistent during consistent blocks), a red 'X' appeared on the screen, immediately below the target word, and the participant was required to make the correct response to clear the screen and proceed to the 400ms intertrial interval. Although participants were required to self-correct, the IRAP program recorded these trials as incorrect. Following each 32-trial block, the computer presented feedback for that block on the percentage of correct responses and the median response latency. It also informed participants that during the next block of trials the previously correct and incorrect answers would be reversed. When participants had completed the two practice blocks an additional message appeared on the screen informing them that they would now be completing a test block and they were instructed to respond quickly and accurately. Following the final trial of the final block, the screen cleared and a message appeared asking participants to report to the experimenter. The experimenter then returned to the room and loaded the second IRAP program. The procedure for the second IRAP was similar to that of the first (i.e. participants proceeded through instructions, practice blocks, and test blocks in the manner described above). Following completion of the final trial in the final block, the participant was again instructed to report to the experimenter. This marked the end of the first session.

Session 2: IQ test

The K-BIT was administered and scored in accordance with the standardized procedures set out in the K-BIT manual (Kaufman & Kaufman, 1990). For all participants, testing started with the Expressive Vocabulary subtest and proceeded to Definitions and finally to Matrices. In the Definitions subtest, participants had a maximum of 30 seconds to respond to each item. There was no time limit for either the Expressive Vocabulary or Matrices subtests. All participants started each subtest at the designated starting point for adult age groups. For each subtest, testing continued until all items were completed or until the discontinue criterion was met (i.e. until a participant failed every item in a unit). During administration, the individual’s test record sheet was shielded from view and no feedback was given until participants had completed the entire test.

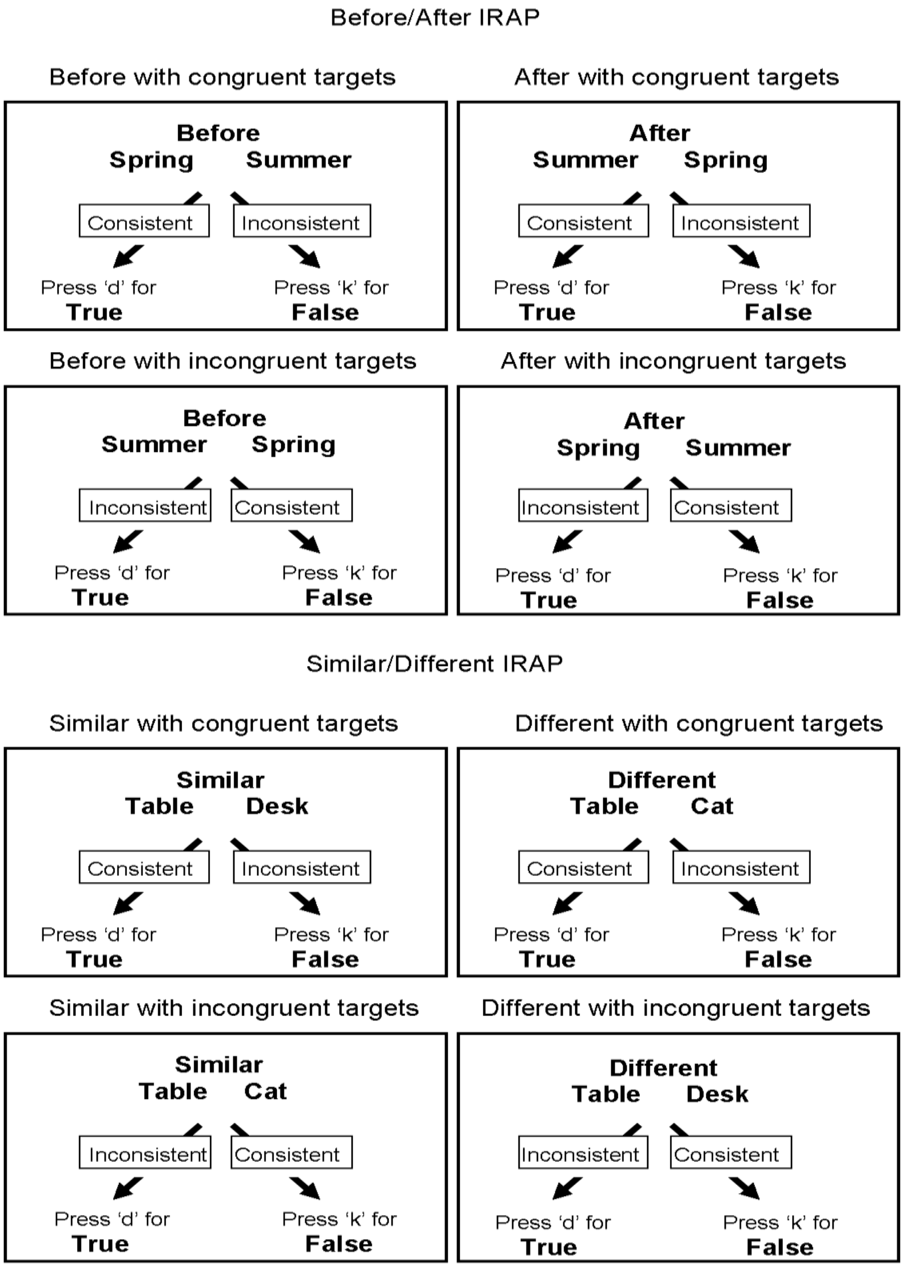

Figure 1. Examples of four trial-types on the before/after IRAP (upper section) and the similar/different IRAP (lower section). Sample stimuli, target words, and response options were presented simultaneously on the screen. Note that the superimposed arrows and text boxes, used here to illustrate consistent and inconsistent responses, did not appear on the screen during the IRAP.

Results

Preliminary Data Screening

Fifty-five participants achieved 80% accuracy on the test blocks for one or both of the IRAPs. Eleven participants met the accuracy criterion on only one IRAP; four on before/after and seven on similar/different. The data from an IRAP exposure that produced a below-criterion performance were removed from subsequent analyses. Given that response accuracy was used to screen out entire data sets falling below 80% correct, it was assumed that accuracy would not discriminate between high and average IQ participants, and preliminary statistical analyses confirmed this prediction (all *ps* > .55). Furthermore, the three latency measures, described subsequently, each failed to correlate with the accuracy scores (all *ps* > .1).

Response latencies on each IRAP trial are recorded from the point of target onset to the first correct response emitted by the participant. To control for extreme outliers, any latencies greater than 10,000ms were removed from the data. Two mean response latencies were then calculated for each participant, one for the consistent and the other for the inconsistent trials.[1] Two overall group mean latencies were then calculated for consistent and inconsistent trials for each IRAP. The data were then checked to determine that the four mean latencies for each participant did not exceed two standard deviations above the corresponding group mean. If a participant produced one or more mean latencies that exceeded two

---

[1] Previous and ongoing research on the IRAP in our laboratory typically involves collapsing the latency data across the test blocks, because block sequence has not been found to interact with the critical IRAP effect, and thus we adopted this practice here.

standard deviations, all of the data for that participant were excluded from further analyses – the data for three participants were removed on this basis. Overall, therefore, 45 and 48 participants contributed data towards the before/after and similar/different IRAP measures, respectively. The response latency and IQ data obtained in the current study were investigated using the Shapiro-Wilk test. The results suggested that there was no violation of the assumption of normality and the data were therefore analyzed using parametric statistics.

Initial IRAP Analyses

Figure 2 shows the overall mean response latencies for consistent and inconsistent blocks on both IRAPs. The consistent blocks on both IRAPs were shorter than on their respective inconsistent blocks (before/after, $M = 2722$, $SE = 70$ versus $M = 3139$, $SE = 89$; and similar/different, $M = 2230$, $SE = 53$ versus $M = 2871$, $SE = 75$). Paired-samples t-tests showed that both differences were significant; before/after ($t[44] = -9.53$, $p < .0001$, $\eta^2 = .67$) and similar/different ($t[47] = -14.22$, $p < .0001$, $\eta^2 = .81$). As predicted, therefore, the current study produced a typical IRAP effect, consistent with that observed in previously reported IRAP studies.

The mean response latencies on the consistent blocks were shorter on the similar/different task ($M = 2230$) relative to the before/after task ($M = 2722$), and the same pattern was observed for the inconsistent blocks (similar/different, $M = 2871$ versus before/after, $M = 3139$). Paired-samples t-tests indicated again that these differences were significant; consistent ($t[40] = 9.99$, $p < .0001$, $\eta^2 = .71$) and inconsistent ($t[40] = 3.21$, $p = .003$, $\eta^2 = .20$). These results indicate that before/after relational responding was more difficult (as reflected in the increased response time) than similar/different responding.

A difference-score was calculated for each participant by subtracting response latencies on consistent trials from latencies on inconsistent trials for each of the IRAPs. The mean difference-score on the similar/different IRAP was greater than that for the before/after IRAP (before/after = 417, and similar/different = 641). A paired-samples t-test comparing the two sets of difference-scores proved to be significant (t[40] = -4.64, $p < .0001$, $\eta^2 = .51$).

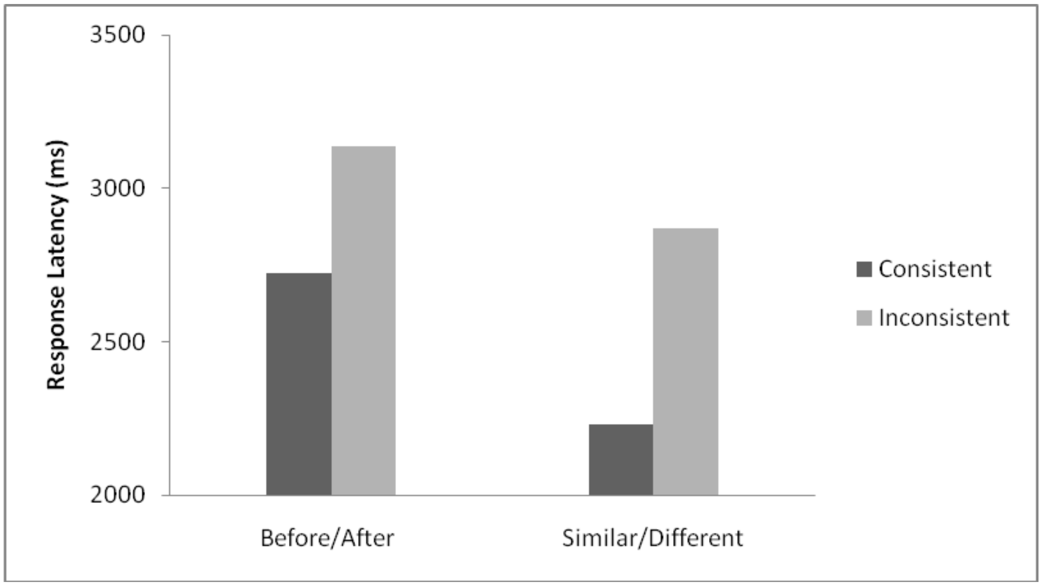

Figure 2. Mean response latencies on consistent and inconsistent blocks for before/after and similar/different IRAPs.

IQ Data

The IQ scores obtained in the current sample ranged from 95-130. The mean IQ score was above average ($N = 52$, $M = 110$), which is unsurprising given that the participants were university students. Furthermore, there was less variation in IQ scores in the current sample ($sd = 8$), than is typically observed in the general population ($SD = 15$; Kaufman & Kaufman, 1990). The IQ score generated by the K-BIT is a composite score, which corresponds to the sum of the standard scores on the Verbal and Matrices subtests. The mean standard score on the Vocabulary subtest was 108 ($sd = 8$) and on the matrices subtest it was 109 ($sd = 9$).

Correlations Between IRAP and IQ

As noted previously, we hypothesied that both the difference-score and response latencies on inconsistent trials may provide a measure of relational flexibility, and such flexibility may constitute an important component of IQ. If this view is correct, negative correlations would be expected between the difference-score and performances on the IQ test. We would also expect more evidence for correlations between response latencies and IQ scores for the inconsistent, relative to the consistent trials. A series of correlational analyses were conducted to test these predictions, the results of which are presented in Table 2.

Table 2. *Correlations Between the IRAP Measures and the Intelligence Test Measures.*

IRAP Measure	IQ	Verbal	Matrices
Response Latency			
Before/after Consistent	-0.27	-0.09	-0.25
Similar/different Consistent	-0.32*	-0.24	-0.24
Before/after Inconsistent	-0.38**	-0.23	-0.32*
Similar/different Inconsistent	-0.35*	-0.43**	-0.16
Difference-score			
Before/after	-0.32*	-0.31*	-0.24
Similar/different	-0.21	-0.44**	0.02

*p < .05. **p < .01

Only one significant correlation was observed between IQ and consistent trials (i.e. the correlation between full-scale IQ and response latency on the similar/different IRAP). In contrast, four significant correlations were observed between the IQ measures and the inconsistent response latencies. Specifically, significant correlations were observed between full-scale IQ and inconsistent trials on both the before/after and similar/different tasks. Furthermore, on similar/different inconsistent trials a significant correlation was observed with standard scores on the Verbal subtest, but the correlation with standard scores on Matrices was non-significant. Interestingly, the opposite pattern was observed for the before/after inconsistent trials (i.e. a significant correlation with Matrices but not with Verbal).

Significant correlations were obtained between the before/after difference-scores and both full-scale IQ and the Verbal subtest, but the correlation with Matrices was non-significant. The difference-scores for the similar/different IRAP also correlated significantly with Verbal, but not with IQ or Matrices. Thus, in contrast to the correlations obtained from the absolute response latency data, neither the similar/different nor the before/after difference-scores correlated significantly with the Matrices subtest, but loaded instead onto the Verbal subtest.

Discussion

The current study showed that individuals who produced higher IQ scores on a standard measure of intelligence tended to respond faster on the relational tasks presented on the IRAP. Importantly, the inconsistent IRAP trials produced a larger number of significant correlations than the consistent tasks, and furthermore only the inconsistent trials produced significant correlations across both similar/different and before/after IRAPs. These results indicate that individuals who performed better on the intelligence test were not only faster at responding relationally, but also demonstrated a greater degree of relational flexibility.

The current study extends the literature on the relationship between relational responding and human cognitive abilities. Most notably, the findings extend the work of O'Hora, et al. (2005) in demonstrating that performances on response-time-based similar/different and temporal relational tasks predict overall IQ as well as subtest scores. Furthermore, unlike O'Hora, et al. (2005, 2008), the current study employed three different response-time-based measures of relational responding (consistent trials, inconsistent trials, and difference-scores), and interestingly different patterns of significant correlations with IQ were observed across the measures. The current findings serve to highlight the importance of using multiple measures of relational responding, rather than relying solely on response accuracy.

The IRAP, as described previously, requires participants to respond as quickly as possible, and thus it might be argued that correlations may emerge only between performance on the IRAP and performance on tests or subtests that require relatively rapid responses (e.g. the Symbol Search or the Digit Symbol-Coding subtest on the WAIS-III). Critically, however, the K-BIT has no such component, and yet one or more of the three IRAP measures correlated with the overall IQ and/or subtest measures (Part B of the Verbal subtest does employ a 30-s cut-off for each question, but this subtest was designed to measure word knowledge and verbal concept formation independent of processing speed; Kaufman & Kaufman, 1990). Furthermore, the current findings may be seen as particularly important given that the K-BIT was designed as a brief IQ test and thus included only the core subcomponents thought to be particularly crucial in the assessment of intelligence. In other words, the observed relationship between the IRAP and K-BIT supports the claim that relational processes, *per se,* are important factors in certain critical aspects of human intellectual ability.

In the current study, response latencies for the similar/different relation were shorter than those for the before/after relation, suggesting that the former relational responding may be at greater strength than the latter responding (at least in the context of the relatively simple and commonly encountered relational stimuli that were employed here). This result is consistent with the argument that the equivalence relation is the most fundamental class of relational responding, and is likely established very early in a child's verbal repertoire (Hayes, 1991). Furthermore, the current data are consistent with the results of a previous study, which showed that response latencies were shorter on tasks that probed for derived sameness/difference relations relative to temporal relations (O'Hora, Roche, Barnes-Holmes, & Smeets, 2002). It is also interesting to note that the difference-score was greater for similar/different relative to the before/after relation, which indicates that reversing the former relation was more difficult than the latter relation. This result would be expected if reversing the similar-different relation does indeed involve responding against the most fundamental and well established verbal relation. Critically, the current results also revealed a relatively strong correlation between the similar/different difference-score and the Verbal subtest. This suggests that the ability to respond rapidly against a highly dominant verbal response (i.e., greater relational flexibility) may be an important feature of verbal intelligence.

The current data revealed that on inconsistent trials the similar/different task loaded onto the Verbal subtest whereas the before/after task loaded onto the Matrices subtest. As argued in the Introduction, however, the difference-score may provide a relatively uncontaminated measure of relational flexibility. Interestingly, the difference-score was found to load heavily onto the Verbal subscale for both the similar/different and the before/after tasks. Thus, the current findings indicate that *flexibility* in relational responding loads almost exclusively onto the verbal as opposed to the performance domain.

It is important to note, however, that although flexibility in relational responding loads onto the Verbal and not the Matrices subscale, this may simply be a function of the specific type of verbal and matrices tasks that were employed in the K-BIT. Perhaps a more extensive intelligence test, such as the WAIS-III, which incorporates a number of verbal and performance subtests, might indicate that relational flexibility loads on to specific performance measures but not others. In fact, O'Hora et al. (2008) found that accuracy on a temporal relations task did indeed correlate with some performance tasks (e.g. Block Design) but not with others (e.g. Matrices and Picture Completion).

As noted earlier, the methodology used by O'Hora and colleagues, only provided a measure of response accuracy. It would be interesting, therefore, to investigate the extent to which the *three* measures of relational responding provided by the IRAP, correlate with performances on the various subtests of the WAIS-III. The purpose of Study 2 was to investigate this issue. Specifically, Study 2, which is described in the following chapter, correlated performances on the same two IRAP tasks (before/after and similar/different) with performances on additional and more extensive measures of cognitive functioning than that employed in the current study.

CHAPTER 3

Study 2: The Relationship Between Similar/Different and Before/After Relational Frames and Performances on the WAIS-III^{UK}, the AH4, and the Cognitive Failures Questionnaire

The purpose of Study 2 was to test the relationship between relational responding and IQ, using multiple measures of cognitive functioning. Participants completed the same two IRAPs as in Study 1 (i.e. before/after and similar/different). They were then exposed to two cognitive abilities tests: the WAIS-III^{UK} (Wechsler Adult Intelligence Scale – Third UK Edition: Wechsler, 1997) and the AH4 (Alice Heim 4; Heim, 1970); and they also completed the Cognitive Failures Questionnaire (CFQ; Broadbent, Cooper, & Fitzgerald, 1982). The WAIS-III^{UK} was chosen because it is one of the most widely used and comprehensive instruments in the assessment of cognitive functioning. It provides a number of different measures of intelligence, including Verbal Comprehension, Perceptual Organization, Working Memory, and Processing Speed. Thus, the WAIS-III^{UK} provides a rich source of information, useful for determining how relational responding and relational flexibility are implicated in the different domains of intelligence. Furthermore, the inclusion of the WAIS-III^{UK} will allow for a more direct comparison between the current study and the previous research conducted by O’Hora and colleagues (2008).

The AH4 was employed in the current study to provide an additional measure of cognitive functioning. The AH4 is a relatively brief test which consists of two parts. Part I contains verbal and numerical items and Part II contains items with a diagrammatic basis, including figure analogies and superimpositions. The AH4 may have particular utility because it differs in a number of respects form the WAIS-III^{UK}. For example, the verbal subtests on the WAIS-III^{UK} include items that assess an individual's previous learning or knowledge in particular domains. In contrast, the verbal items contained on the AH4 require knowledge of everyday words. Thus, the AH4 does not rely on elaborate verbal content. Instead a heavy emphasis is placed on deductive reasoning and on the ability to follow instructions exactly (Heim, 1970). Tables 3 and 4 provide fictitious examples of the types of items contained in both the WAIS^{UK}-III and the AH4. The differences in the nature of the verbal items presented on the two tests can be identified by examining these tables. Another difference between the two intelligence tests is that a time limit is imposed on both the verbal and non-verbal subtests of the AH4. In contrast, only one of the seven verbal subtests on the WAIS-III^{UK} involves a timed component (the Arithmetic subtest).

The differences between the two intelligence tests may overlap with the distinction between crystallised and fluid intelligence. Most tests of verbal ability (including the verbal subtests on the WAIS-III^{UK}) are described as involving crystallised ability, because they tap into an individual's previous education or cultural experiences. However, given the relatively weak requirement for elaborate verbal knowledge, and the inclusion of a timed component, the AH4 could be described as a test that focuses more on fluid ability (e.g., Parkin and Java, 1999). Thus, although Part I of the AH4 assesses verbal skills, it seems to tap into verbal *processes* rather than particular verbal knowledge. In other words, the AH4 could be described as assessing verbal processes, whereas the verbal subtests on the WAIS-III^{UK} could be described as assessing, at least to some extent, the *outcome* of those processes. It will be interesting to observe whether the IRAP produces differential correlations on these different measures of verbal ability.

As noted in Chapter 1, the three measures of relational responding provided by the IRAP are likely to be important for facilitating a more complete RFT investigation of intelligence. However, it is likely that performances on the IRAP are influenced by variables other than speed, accuracy and flexibility. Specifically, successfully completing an IRAP may tap into additional processes such as sustained attention and freedom from distraction. It might be useful therefore to investigate the impact of these additional processes. The CFQ was employed in the current study for this reason. The CFQ is a measure of self-reported deficits in the completion of simple everyday tasks. The questionnaire is thought to provide a measure of absent mindedness or of slips of attention. Thus, if performance on the IRAP is strongly influenced by the ability to sustain attention then we might expect significant correlations between the questionnaire responses and the measures provided by the IRAP.

In the current study the same two IRAP tasks were presented as in Study 1. All stimuli, including sample stimuli, target stimuli, and response options were the same. However, the IRAPs differed in one respect. Specifically, in Study 1, only one set of practice blocks (one consistent and one inconsistent block) was presented. If participants did not reach the 80% accuracy criterion on either the consistent or inconsistent block, then their data was disregarded. In order to prevent high attrition rates, the current IRAP programme presented up to three pairs of practice blocks. Thus, if participants failed to reach the 80% accuracy criterion on the first pair of practice blocks, they completed up two more pairs until this criterion was met.

The current study represents a partial replication of both Study 1 and the work of O’Hora et al. (2005, 2008), and thus the pattern of correlations that emerged across previous studies allowed us to make some general predictions. Based on the findings of both Study 1 and the study by O’Hora et al. (2008), it was expected that performances on the IRAPs would correlate with the full scale IQ score. Significant correlations were also expected for some of the Perceptual Organisation and Verbal Comprehension subtests, since correlations with these indices were observed in the O’Hora et al. (2008) study. It was also predicted that the consistent and inconsistent response latency measures would correlate with the processing speed subtests, since each of these tasks involve the requirement to respond quickly and accurately. Similarly, one would also expect the response latency measures to correlate with the AH4, since it too involves a timed component. In contrast, it was predicted that the correlations with the working memory subtests would be non-significant, as they did not correlate significantly with the relational tasks in O’Hora et al. studies (2005, 2008), and also because working memory tasks are not specifically designed to tap into relational processes.

Table 3. Examples of the Type of Items Included in Each of the 13 WAIS-III^{UK} Subtests (all specific examples are fictitious, but representative of actual items; adapted from Flynn, 2007).

Verbal Subtests	
Vocabulary	What does "debilitating" mean?
Similarities	In what way are "dogs" and "rabbits" alike?
Information	On what continent is Argentina?
Comprehension	Why are streets usually numbered in order?
Arithmetic	If four toys cost 6 euro, how much do seven cost?
Digit Span	Repeat a series of numbers read aloud
Letter-Number Seq.	Re-order and repeat a series of numbers and letters
Performance Subtests	
Picture Completion	Indicate the missing part from an incomplete picture
Block Design	Use blocks to reproduce a two-colour design
Matrix Reasoning	Select the abstract design that best completes a pattern matrix
Picture Arrangement	Re-order a set of scrambled picture cards to tell a story
Digit-Symbol Coding	Use a key to quickly match symbols with numbers
Symbol Search	Quickly indicate whether one of two symbols is present in a series of symbols

Table 4. Examples of Items Included in Parts I and II of the AH4 (all items are fictitious but representative).

Part I: Verbal & Numerical	• 3, 6, 9, 12... What number comes next? • *Happy* means the opposite of... 1. difficult, 2. sad, 3. excited, 4. weary, 5. hopeless • *Knife* is to *cut* as *spade* is to... 1. tear, 2. fork, 3. bandage, 4. dig, 5. lift • If 18 is more than two times six, write down the figure 3, unless 15 is less than 12, in which case, write 7.
Part II: Diagrammatic	• Identify the relevant features in a series of abstract patterns, and select the missing pattern • Mentally superimpose one design on top of another, and identify the resulting design • Select the shape that is identical to the target • Complete a diagrammatic analogy • Mentally subtract a part from a whole shape, and identify the resulting shape

Method

Participants

Participants were 53 undergraduate students from various faculties of the National University of Ireland, Maynooth. None of the participants had previous exposure to the IRAP protocols used in the current study, nor were they familiar with the cognitive abilities tests that were subsequently administered. The study was conducted across two separate sessions. The IRAPs were presented in session 1. The cognitive abilities tests and the questionnaire were presented in session 2. Six participants failed to return for the second session, and their data are not included in the final analyses. In addition, to control for the confounding effects that may occur due to a speed /accuracy trade-off, only the data from participants that maintained the ≥ 80% correct responding criterion, across the six IRAP test blocks, were included in the final analyses. Forty-three participants (16 male; 27 female) of the 47 participants who completed the study, met this criterion on the similar/different IRAP. Forty-four participants (17 male and 27 female) met the criterion on the before/after IRAP. Participants ranged in age from 17-52 years (mode = 18; *M* = 22).

Materials and Stimuli

Materials included the IRAP software, a participant profile form, the WAIS-III^{UK}, the AH4, and the CFQ. The profile form was identical to that employed in Study 1. The IRAP tasks were also identical to those in Study 1 except for one difference, which will be discussed in the *Procedure*.

The WAIS-III^{UK}

The WAIS-III^{UK} is a comprehensive clinical instrument for assessing the cognitive ability of adults aged 16-89 years. It is comprised of 14 subtests, each of which assesses a particular area of intellectual functioning. The subtests are divided into two Scales: the Verbal Scale and the Performance Scale. The Verbal Scale is comprised of seven subtests, which are presented in question-and-answer format. These are: Vocabulary, Similarities, Information, Comprehension, Arithmetic, Digit Span, and Letter Number Sequencing. The Performance Scale is comprised of seven subtests that require the manipulation or recognition of material in pictorial or three-dimensional form. These are: Picture Completion, Block Design, Matrix Reasoning, Picture Arrangement, Digit-Symbol Coding, Symbol

Search, and Object Assembly. As noted earlier, Table 3 provides an example of the type of questions or stimuli that are presented in each of these subtests. The subtests can be combined in various ways to provide a Verbal IQ score, a Performance IQ score and a Full-Scale IQ score. They can also be combined to provide four index scores: Verbal Comprehension, Perceptual Organisation, Working Memory, and Processing Speed. Figure 3 illustrates how the subtests are combined to produce these scores.

The WAIS-III^{UK} is similar to the WAIS-III, which was standardised in the United States on 2,450 adult subjects. The difference between the two versions is that some items on the WAIS-III^{UK} have been adapted for a UK population (for example, the Information subtest contains a question about a British prime minister rather than a US president). For the purpose of the current study the WAIS- III^{UK} was deemed appropriate, given the geographical proximity, and the cultural similarities between Ireland and the UK. The UK standardisation project took the form of a validity and comparability study between the American norms and the scores of a representative sample (N = 332) of the UK population. The differences between the two samples were not considered large enough to materially affect the application of the US norms in the United Kingdom. The WAIS-III reports excellent test-retest stability coefficients averaging between 0.7 and 0.9 across the age groups. Furthermore, excellent correlations with other intelligence measures are reported. For example, the correlation between the Stanford Binet–IV composite score and the WAIS-III Full-Scale IQ score is 0.88.

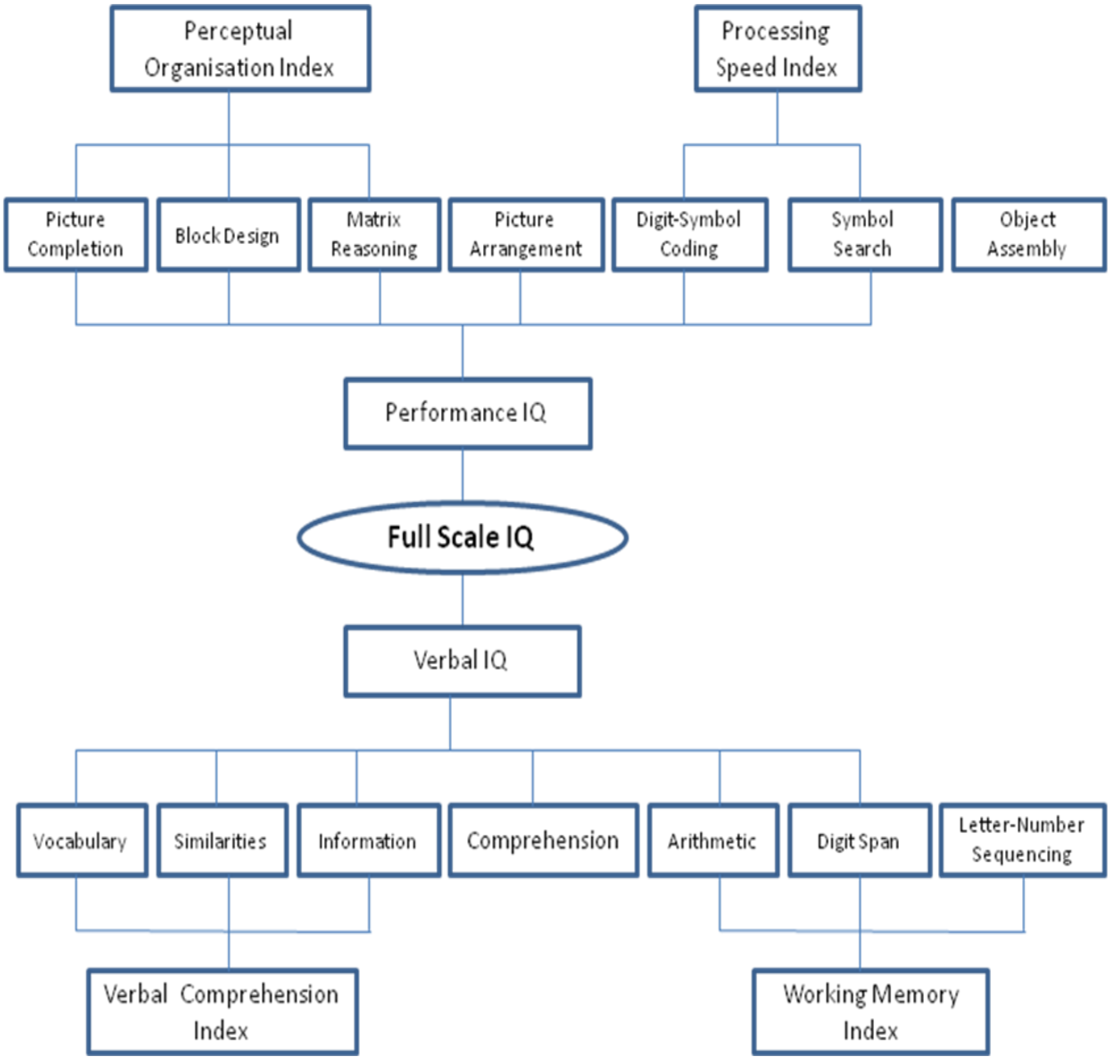

Figure 3. Schematic representation of the organization of the WAIS-III^{UK} subtests. Note that the Comprehension and Picture Arrangement subtests load onto Verbal IQ and Performance IQ respectively, but do not load onto any of the Index scales. Letter-Number Sequencing loads onto the Working Memory Index, but not onto Verbal IQ. The Object Assembly subtest is supplementary, in that it contributes to neither the IQ scores nor the index scales.

The AH4

The AH4 was designed as a group test of general intelligence for use with a cross-section of the adult population. Part I consists of 65 questions which have a verbal and numerical basis. Six types of principle are involved: verbal opposites, numerical series, verbal analogies, simple arithmetical computations, and synonyms. Part II consists of 65 questions which have a diagrammatic basis and which are exemplified by five types of principles: analogies, sames, subtractions, series and superimpositions (Heim, 1970). Each part also contains a set of sample questions, the main purpose of which is to familiarise examinees with the nature of the task. The time limit for each part is 10 minutes, exclusive of the preliminary examples. A separate test score is calculated for Parts I and II, and these can be added together to produce a score of overall test performance. The AH4 records an excellent test-retest co-efficient of 0.9. Correlations with other intelligence tests range from 0.45-0.76. The AH4 was standardised with a cross section of the adult population, but given that the test is 40 years old, these norms are outdated. However, in the context of the current study the outdated norms were irrelevant, as we were not seeking to compare participants with a normative sample.

The CFQ

The CFQ is a self-report questionnaire, containing 25 items. Participants are asked how often they make mistakes on a 5-point Likert scale, from 0 (never) to 4 (very often). The CFQ is scored by adding up the ratings for the 25 items. The highest possible total is 100, with a higher score indicating a higher incidence of cognitive failures. The CFQ has high internal validity (alpha=0.91) and is stable over long periods of time, with a test-retest reliability rate of 0.82.

Procedure

The study was split into two sessions. In Session 1, participants completed the two IRAP protocols. In Session 2 the WAIS-III^{UK}, and the AH4 were administered, and participants filled out the CFQ. The sessions were conducted on separate days but within two working weeks of each other. Both sessions were conducted in the Department of Psychology at NUI Maynooth. Session 1 took place in individual cubicles in the experimental laboratory. Session 2 was conducted in a bright, more spacious room within the department. Session 1 lasted for approximately one hour, and Session 2 took approximately two hours to complete.

Session 1: IRAPs

The IRAPs presented during the current study were identical to those in Study 1 except that participants were presented with up to three pairs of practice blocks. Participants were required to reach a standard of ≥ 80% correct responses across a set of consistent and inconsistent practice blocks, before they entered the test phase. In addition, a median response time of ≤ 4,500 ms was also set as a criterion for entering the test phase. This response time criterion was selected on the basis of the overall mean IRAP performance in Study 1. Specifically, the figure of 4,500 was calculated by identifying the average response latency for inconsistent blocks on the before/after IRAP and the similar/different IRAP, and then adding one and a half times the standard deviation to that figure. These criteria were used to ensure that participants were complying with the IRAP instructions, and to reduce attrition rates relative to the previous study. If participants failed to achieve the two criteria for either of the two practice blocks, the required standard and the standard of responding they had achieved were presented on the screen. Participants were allowed three attempts (a total of six practice blocks) to achieve the practice criteria, and if they failed to do so, they were thanked and debriefed and their data were discarded. Participants who did achieve the practice criteria proceeded to the six test blocks.

Session 2: Intelligence tests

Session 2 involved first completing the WAIS- III^{UK}, then the CFQ, and finally, the AH4.[2] The session was conducted in a bright, well-ventilated room that was free from outside interruptions. Only the participant and the examinee were present in the room during the session, and they sat opposite each other at a table. At the start of the session, the examinee informed participants about the nature of the session, and the type of tasks that he/she would be exposed to. Participants were informed that they may find some items quite easy, and that others may be more difficult, but that they should give their best effort on all tasks.

---

[2] When the current study was being planned it was initially felt that completing the two intelligence tests and the CFQ in a single session may result in participant fatigue. Therefore, at the very start of the study we asked participants to complete the WAIS in one session and the AH4 and CFQ in another session. However, some participants (four in total) did not return to complete the AH4 and the CFQ. Thus, we decided to administer all three measures in a single session, but to encourage participants to take short breaks between tests if required. This procedure did not appear to adversely affect participant concentration or motivation.

The WAIS- III^{UK} was administered and scored in accordance with the standardized procedures set out in the WAIS- III^{UK} manual. All subtests were administered with the exception of object assembly. This subtest was omitted for two reasons. First, it does not contribute to any of the IQ or index scores, and second, we wanted to reduce the length of Session 2, which was already quite protracted. The subtests were administered in the order recommended in the WAIS-III^{UK} manual. This sequence is presented in Table 5. Following the administration of the thirteenth subtest, participants were invited to take a short break, if they so wished, and subsequently participants were presented with the CFQ, and asked to fill it out in accordance with the instruction at the top of the page.

When participants had completed the CFQ, they were presented with the AH4 question booklet, an answer sheet, and a pencil. The AH4 was administered in accordance with the instructions in the manual, with the exception that participants were not asked to fill in the profile page on the AH4 answer sheet (an alphanumeric code was used by the examinee to protect the participant’s identity). Participants were directed to Part I of the question booklet and asked to read the instructions at the top of the sheet, and answer the sample questions. Participants were then asked if they had any queries in relation to these sample items. Once these were addressed, the stopwatch was set, and participants completed Part I of the test booklet. After ten minutes had elapsed, participants were asked to stop working, and then they were invited to complete the sample and test questions in Part II (instructions for Part II were the same as for Part I). After a further ten minutes had elapsed participants were again instructed to stop working. Subsequently, participants received verbal feedback on their performance on the WAIS-III^{UK}, and were fully debriefed and thanked for taking part in the study.

Table 5. Administration Sequence of the WAIS-III^{UK} Subtests. Note that in the current study the final subtest (object assembly) was not administered

Testing Sequence	Subtest
1	Picture Completion
2	Vocabulary
3	Digit-Symbol Coding
4	Similarities
5	Block Design
6	Arithmetic
7	Matrix Reasoning
8	Digit- Span
9	Information
10	Picture Arrangement
11	Comprehension
12	Symbol Search
13	letter-Number Sequencing
14	Object Assembly

Results

Preliminary Data Screening

Forty-four participants achieved 80% accuracy on the test blocks for the before/after IRAP, and 43 participant achieved this criterion on the similar/different IRAP. The data from an IRAP exposure that produced a below-criterion performance were removed from subsequent analyses. The preliminary data screening was similar to that employed in Study 1. Specifically, response latencies greater than 10,000 ms were removed from the data, and then two mean response latencies were calculated for each participant, one for the consistent and the other for the inconsistent trials. Subsequently, two overall group mean latencies were calculated for consistent and inconsistent trials for each IRAP.

The data were then checked to determine that the four mean latencies for each participant did not exceed two standard deviations above the corresponding group mean. If a participant produced one or more mean latencies that exceeded two standard deviations on a particular IRAP, then the data for that IRAP were excluded from further analyses – the data for five IRAP performances were removed on this basis. Overall, 42 and 40 participants contributed data towards the before/after and similar/different IRAP measures, respectively. The response latency and IQ data obtained in the current study were investigated using the Kolmogorov-Smirnov test. The results suggested that there was no violation of the assumption of normality and the data were therefore analyzed using parametric statistics.

Initial IRAP Analyses

The pattern of IRAP results observed in the current study was similar to that obtained in Study 1. As shown in Figure 4, the overall mean response latencies for the consistent blocks on both IRAPs were shorter than on their respective inconsistent blocks (before/after, consistent, (*M* = 2698, *SE* = 87) versus inconsistent, (*M* = 3050, *SE* = 112); and similar/different, consistent, (*M* = 2248, *SE* = 79) versus inconsistent, (*M* = 2779, *SE* = 124). Paired-samples t-tests showed that both differences were significant; before/after ($t[41] = -6.56$, $p < .0001$, $\eta^2 = .25$)

and similar/different (*t*[39] = -8.15, p <.0001, η^2 = .3). Thus, as predicted, the current study produced typical IRAP effects.

Consistent with Study 1, the current results indicate that before/after relational responding requires more time (and is therefore more difficult) than similar/different responding. Paired-samples t-tests indicated again that these differences were significant for IRAP performance on both consistent (*t*[37] = 6.8, p <.0001, η^2 = .27) and inconsistent blocks of trials (*t*[37] = 3.2, p = .003, η^2 = .22).

A difference-score was then calculated for each participant by subtracting response latencies on consistent trials from latencies on inconsistent trials for each of the IRAPs. The mean difference-score on the similar/different IRAP was greater than that for the before/after IRAP (before/after = 353, and similar/different = 531). A paired-samples t-test comparing the two sets of difference-scores proved to be significant (*t*[38] = -3.1, p = .003, η^2 = .14).

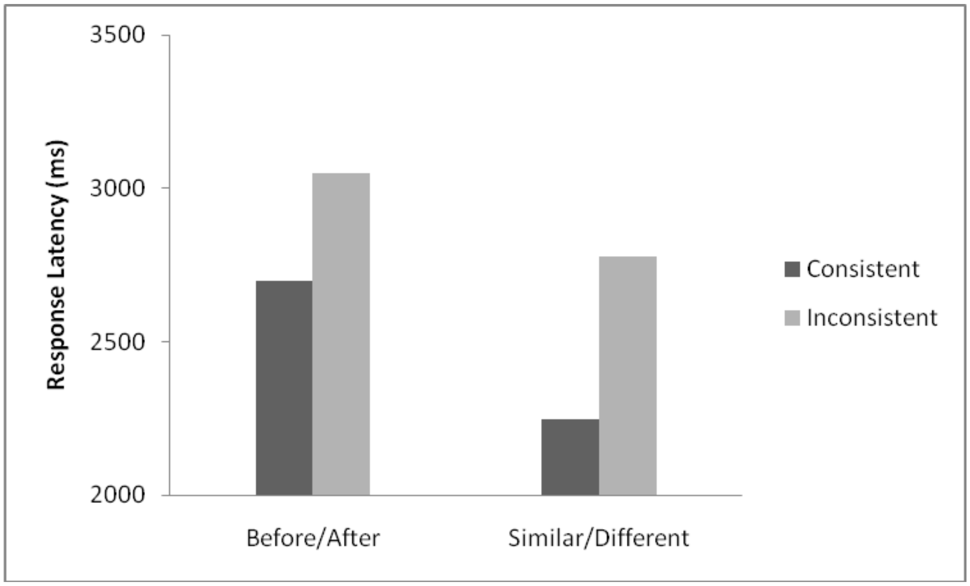

Figure 4. Mean response latencies for consistent and inconsistent block for before/after and similar/different IRAPs

IQ and CFQ Data

The descriptive statistics obtained for performance on the intelligence tests and the CFQ are presented in Table 6. The data highlight that on the WAIS-III^{UK}, the mean IQ scores and index scores were above average, and showed less variation than those observed in the standardised population (*M* = 100, *SD* = 15; Wechsler, 1997). The mean scores on the individual subtests were also above average, but the variation in the data was roughly similar to that expected in the general population (standardised mean = 10, SD = 3). These data are unsurprising given that participants in the current study were university undergraduates.

The data in Table 6 also presents the raw scores obtained by participants on the AH4 (the AH4 manual does not provide for the conversion of raw scores to standardised scores). These data suggest that participants correctly completed more items on Part II of the test, relative to Part I. As would be expected, performances on the two intelligence measures were positively correlated: correlation coefficient for Full-Scale IQ with the overall score on the AH4 = 0.58 ($p < .0001$); verbal IQ with part I of the AH4 = 0.43 ($p = 0.004$); and performance IQ with Part II of the AH4 = 0.63. ($p < 0.0001$).

Table 6. Descriptive Statistics of Performance on the WAIS-III^{UK}, the AH4, and the CFQ.

	N	Mean	SD	Range
WAIS-III^{UK}				
IQ Scores				
Full Scale IQ	44	118	10	102-137
Index Scores				
Verbal Comprehension Index	44	121	11	100-138
Perceptual Organisation Index	44	112	11	86-138
Working Memory Index	44	113	11	94-136
Processing Speed Index	44	115	12	93-148
Subtest Standard Scores				
Similarities	44	14	2	10-18
Vocabulary	44	14	3	8-19
Information	44	13	2	11-17
Comprehension	44	13	3	8-19
Picture Arrangement	44	11	3	3-16
Picture Completion	44	10	3	6-15
Block Design	44	13	3	7-18
Matrix Reasoning	44	13	2	9-17
Arithmetic	44	12	2	9-16
Digit Span	44	12	2	9-16
Letter-Number Sequencing	44	12	2	8-17
Digit-Symbol Coding	44	13	3	8-19
Symbol Search	44	13	2	7-18
AH4 (raw scores)				
Overall	40	100	12	66-117
Part I	40	46	6	32-57
Part II	40	55	8	30-65
CFQ	41	42	11	17-73

Correlations Between the IRAPs and the Cognitive Measures

The instruments employed in the current study provide numerous measures of performance, and therefore a large number of correlational analyses may be conducted. Typically, when conducting multiple tests of significance, a correctional procedure is used to control for the occurrence of Type II errors (e.g., the Bonferroni procedure; Abdi, 2007). However, in the current study we wanted to compare the pattern of results with those that emerged in previous studies and would emerge in subsequent research reported in the current thesis. Consequently, it could be argued that applying a stringent correction procedure may increase the risk of Type I errors, and lead to important similarities across the current, previous and subsequent studies being overlooked. Thus, it is possible that some of the significant but non-corrected correlations, which will be highlighted subsequently, may reflect Type II errors. On balance, however, it is important to note that the current study represents a partial replication of both Study 1 and the work of O'Hora et al. (2005, 2008), and thus the pattern of correlations that emerged across previous studies allowed us to make some general predictions prior to analysing the current data (these are highlighted in the Introduction to this study). Thus, although a correctional procedure was not employed, the current analytic approach is relatively conservative.

The correlations obtained between the IRAP performances and both the intelligence measures and the CFQ are presented in Table 7. As shown, all of the correlations with the WAIS-III^{UK} and the AH4 are negative. Thus, higher scores on the intelligence measures are associated with shorter response latencies, and smaller difference-scores. In the following section we will discuss the correlations obtained under a number of headings.

Correlations Between the Before/After IRAP and the WAIS-III^{UK}: Significant correlations were observed for both Full-Scale IQ with each of the three before/after IRAP measures (consistent, inconsistent, and the difference score). Significant correlations were also observed for the Perceptual Organisation index with each of the three IRAP measures. An analysis of the individual subtests that load onto this index revealed that the Block Design subtest correlated significantly with each of the three IRAP measures, and was therefore the main contributor to the correlation with the Perceptual Organisation index. All of the correlations with both the Verbal Comprehension index and the Working Memory index were non-significant.

As predicted, the consistent and inconsistent response latency measures produced significant correlations with the Processing Speed subtests (Digit-Symbol Coding, and Symbol Search), and also, therefore, with the Processing Speed index. Importantly, significant correlations were not observed between the Processing Speed measures and the difference-scores. This again supports the suggestion that the difference-score controls for extraneous variables, and may thus provide a relatively “pure” and uncontaminated measure of relational flexibility.

Correlations Between the Similar/Different IRAP and the WAIS-III^{UK}: The correlations between the similar/different IRAP and the IQ measures are presented at the right-hand side of Table 7. As highlighted, full scale IQ measures correlated significantly with all of the IRAP measures. Significant correlations were observed for the Verbal Comprehension index with both the inconsistent trials and the difference-score. In contrast, no significant correlations were observed for the Perceptual Organisation index. This pattern of results is the opposite of that observed for the before/after IRAP (i.e., the before/after IRAP produced differential correlations with Perceptual Organisation, not Verbal Comprehension). An analysis of the individual subtests that load on to the Verbal Comprehension index revealed significant effects for Vocabulary subtest with both the inconsistent trials and the difference-score. Significant correlations were also observed between the Comprehension subtest and both the consistent and inconsistent trials (the correlation between Comprehension and the difference-score fell just outside the significant range; $p = .06$). On the Perceptual Organisation subtests, significant correlations were observed between Block Design and both the consistent and inconsistent trials, and between Picture Arrangement and the inconsistent trials and the difference-score.

The pattern of correlations that emerged between the similar/different IRAP and both the Processing Speed and Working Memory indices were similar to those observed on the before/after IRAP. Specifically, the similar/different consistent and inconsistent trials, but not the difference-score, correlated significantly with the Processing Speed subtests, as well as the Processing Speed index. Furthermore, as predicted no significant correlations were observed between the similar/different IRAP measures and the Working Memory measures.

Correlations between the IRAPs and the AH4: Both the similar/different and before/after tasks were significant predictors of performance on the AH4.

Given that the AH4 involves a timed component, we predicted significant correlation between it and the response latency measures. Interestingly, however, significant correlations were also observed with the difference-scores. The before/after difference-score produced highly significant correlations with each of the AH4 measures, while the similar/different difference-score correlated significantly with two of the AH4 measures (the correlation with Part II was just outside the significant range; *p* = 0.06).

Correlations between the IRAPs and the CFQ: Positive correlations were observed between the CFQ and the consistent and inconsistent IRAP measures. This indicates that higher scores on the CFQ are associated with longer response latencies. These correlations reached statistical significance for the before/after, but not for similar/different consistent and inconsistent blocks. Importantly, the correlations between the CFQ and the before/after and similar/different difference-scores were both weak and non-significant. This again supports the argument presented in Chapter 2 that the difference-score controls for extraneous variables (such as slips of attention).

In sum, performance on both the before/after and similar/different tasks correlated significantly with Full-Scale IQ on the WAIS-III and with the AH4. The similar/different task correlated significantly with the Verbal Comprehension index but not the Perceptual Organisation index, although significant correlations were observed with some of the performance subtests. On the before/after IRAP the opposite pattern was evident: the before/after task loaded more heavily onto the Perceptual Organisation index. The difference-scores for both IRAPs correlated significantly with the overall measures of intelligence, but not with the CFQ or the Processing Speed measures. None of the IRAP measures correlated significantly with the Working Memory subtests.

Table 7. Correlations Between the Before/After (BA) and Similar/Different (SD) IRAP Measures with the Intelligence Measures

	BA Consistent	BA Inconsistent	BA Difference Score	SD Consistent	SD Inconsistent	SD Difference Score
WAIS-III^{UK}						
Full Scale IQ	-0.46**	-0.52**	-0.34*	-0.41**	-0.47**	-0.4**
Index Scores						
Verbal Comp.	-0.18	-0.27	-0.26	-0.25	-0.34*	-0.34*
Perceptual Org.	-0.39**	-0.48**	-0.36*	-0.23	-0.26	-0.22
Working Memory	-0.15	-0.18	-0.12	-0.1	-0.13	-0.13
Processing Speed	-0.53**	-0.48**	-0.19	-0.5**	-0.41**	-0.18
Subtest Standard Scores						
Similarities	-0.19	-0.25	-0.2	-0.18	-0.25	-0.24
Vocabulary	-0.19	-0.28	-0.27	-0.28	-0.4**	-0.41**
Information	-0.04	-0.07	-0.13	-0.08	0.11	-0.11
Comprehension	-0.28	-0.28	-0.13	-0.34*	-0.37*	-0.3
Picture Arrangement	-0.16	-0.25	-0.25	-0.22	-0.33*	-0.36*
Picture Completion	-0.18	-0.22	-0.18	-0.07	-0.12	-0.14
Block Design	-0.52**	-0.58**	-0.37**	-0.38**	-0.35*	-0.21
Matrix Reasoning	-0.22	-0.3	-0.26	-0.08	-0.15	-0.18
Arithmetic	-0.09	-0.14	-0.15	-0.03	0.03	-0.02
Digit Span	-0.1	-0.09	-0.03	-0.07	-0.07	-0.04
Letter-Number Seq.	-0.17	-0.2	-0.14	-0.15	-0.26	-0.29
Digit-Symbol Coding	-0.45**	-0.35*	-0.007	-0.41**	-0.34*	-0.14
Symbol Search	-0.49**	-0.52**	-0.28	-0.48**	-0.4**	-0.18
AH4 (raw scores)						
Overall	-0.5**	-0.68**	-0.57**	-0.56**	-0.54**	-0.36*
Part I	-0.43**	-0.55**	-0.42**	-0.39*	-0.42**	-0.34*
Part II	-0.46**	-0.66**	-0.59**	-0.6**	-0.54**	-0.32
CFQ	0.36*	0.33*	0.1	0.3	0.16	-.05

*p < .05. **p < .01

Discussion

The results of the current study highlighted again the important role of relational processes in human intelligence. Specifically, the results revealed that participants who responded faster on the relational tasks produced higher scores on the intelligence measures. The data also highlighted the importance of the difference-scores as relatively “pure” or uncontaminated measures of relational flexibility. That is, the difference-scores appeared to control for both speed of processing (as measured by the Processing Speed index), and momentary lapses in attention (as measured by the CFQ). Critically, the difference-scores for each IRAP were significant predictors of performance on both the WAIS-III^{UK} and the AH4. These results therefore highlight the importance of relational flexibility in human intelligence.

An important finding of the current study is that the before/after and similar/different tasks appeared to discriminate between different types of intellectual ability. The before/after IRAP loaded more heavily onto the Perceptual Organisation index, relative to the Verbal Comprehension index, whereas for the similar/different task the opposite pattern was apparent (heavier loadings on Verbal Comprehension index than on the Perceptual Organisation index). These results are broadly consistent with those obtained in Study 1, which showed that the similar/different task loaded onto the Verbal subtest of the K-BIT, whereas the before/after task loaded more heavily onto the Matrices subtest.

The finding that the similar/different task loaded onto verbal intelligence is perhaps unsurprising. These tasks incorporate the relational frames of co-ordination and distinction. As noted in Chapter 1, RFT emphasises that the frame of co-ordination provides the functional basis for naming and semantic relations more generally (Barnes-Holmes, et al., 2001), and thus the correlation with Verbal Comprehension, and Vocabulary in particular, supports this view. Furthermore, the finding that the before/after task loads onto Perceptual Organisation index, and onto the Block Design task in particular, is also consistent with an RFT account. The Block Design task involves reproducing a two-dimensional design using blocks. Although ostensibly non-verbal in nature, this type of task involves the application of relational frames, and therefore taps into what is referred to as pragmatic verbal analysis (see Chapter 1). Indeed O'Hora et al. (2008) also reported a significant correlation with the Block Design subtest, and they highlighted how the completion of this task depends on this type of verbal behaviour:

> "Pragmatic verbal analysis is a term used in RFT to describe the occurrence of arbitrarily applicable relational responding under the control of nonarbitrary physical-world relations (like those in the temporal relations task). In the Block Design task, participants are required to orient up to nine blocks to create a design provided by the experimenter. The comparisons made by the participant between the stimuli in their orientation and the required orientation, which require physical distinctions to be identified and modified toward a goal state, are thus a form

> of pragmatic verbal analysis. The significant moderately strong correlation obtained between the relational task and the Block Design subtest supports this line of reasoning.” (pg., 578)

Given that both the current study and the previous work by O’Hora et al. (2008) involved presenting participants with before/after relational tasks, it is possible to compare the pattern of results that emerged for these tasks across the two studies. As highlighted above, the before/after tasks in both studies produced significant correlations with the Perceptual Organisation index, and with the Block Design task. Furthermore, O’Hora et al. found that accuracy on the before/after task correlated significantly with three verbal subtests (Vocabulary, Similarities, and Information). In contrast, the before/after task in the current study produced no significant correlations with the verbal subtests (although some correlations approached significance).

One factor that may have influenced the differential correlations is the difference in sample size across the two studies. Relative to the current study, O’Hora et al. (2008) employed a larger sample (N = 81), and this may have conferred greater statistical power. Furthermore, the range in intelligence differed quite dramatically between the two studies (IQ scores of 84 to151 were recorded in the O’Hora et al. study, versus 102-137 in the current study). Perhaps, this too helps to account for the different outcomes across the two studies. However, it also seems likely that some of the differences may be accounted for in terms of the procedures that were employed to present the relational tasks. In the O’Hora et al.

study the before/after task was presented using the Relational Evaluation Procedure (REP). As highlighted in Chapter 1, this procedure involves training participants to relate previously meaningless abstract stimuli, and provides a measure of the accuracy of relational responding, with no requirement to respond under time pressure. The IRAP, in contrast, requires participants to respond quickly, accurately, and with flexibility to pre-established relations. The use of these different methodologies may account for the fact that correlations with the Verbal Comprehension subtests were observed in the O'Hora et al. study, but not in the current study. For example, perhaps the ability to derive new relations among abstract stimuli (which is required by the REP) taps into verbal intelligence (as traditionally defined), whereas fast and flexible responding, that is required by the IRAP, places a relatively heavier demand on the more fluid aspects of intelligence. Indeed, in describing fluid intelligence, cognitive scientists refer to the ability to adapt quickly to novel and changing environments (e.g., Horn & Cattell, 1966). This ability is captured by the requirement on the IRAP to switch repeatedly and rapidly between two opposing response patterns.

The suggestion that the IRAP taps into fluid aspects of intelligence may be further supported by the particular patterns of correlations obtained in the current study. Specifically, both the before/after and the similar/different tasks produced significant correlations with the AH4. In fact, only one of the nine correlations between the IRAPs and the AH4 was outside the range of significance (i.e., the correlation between the similar/different difference-score with Part II of the AH4, $p = 0.06$). As highlighted in the Introduction, the AH4 can be considered a measure of fluid ability (e.g., Parkin and Java, 1999), which differs from the WAIS-III^{UK}, particularly in relation to verbal tasks (i.e., the verbal items on the WAIS-III^{UK} assess knowledge in particular domains, whereas the AH4 places greater emphasis on relational processes). Thus, it could be argued that performance on the IRAPs correlated the AH4 because both tasks tap into the *process* of relational responding. On balance, the IRAP appears to taps into certain aspects of fluid ability and not others. Working memory, for instance, correlates significantly with measures of fluid ability such as reasoning (e.g., Kyllonen, 1996), but the current study demonstrated that working memory is not taped into by the relational tasks on the IRAP.

In order to determine which features of human intelligence are reflected in an IRAP performance, it is important to recognise that the nature of stimuli employed in the procedure may well be important in this regard. Both Studies 1 and 2 of the current thesis involved presenting before/after and similar/different relational tasks using relatively simple stimulus relations. Thus it remains to be seen if IRAP performances continue to correlate with intelligence measures when additional and more complex relations are presented within the IRAP. The subsequent studies were designed to investigate this issue.

CHAPTER 4

Study 3: Relating-Relations and performances on the CAT3 and the AH4

The previous studies focused on relatively basic relational frames, such as similar and different. These frames are considered to provide the basic functional units of human language and cognition, and thus the correlations that were obtained with IQ provide some support for the RFT conceptualization of human cognition. At this point in the research programme it was deemed important to employ more complex relations in the IRAP than were used in previous research. Intuitively, one might expect that if the IRAP tasks overlap to some extent with actual test items found on IQ tests, stronger correlations between the IRAP and IQ might be obtained. On balance, it is important to remember that the IRAP requires relatively rapid levels of continuous responding and the critical measure of relational flexibility is the difference-score. It is possible, therefore, that the intensely time-based IRAP only provides a sensitive measure of intelligence when targeting basic relational frames, rather than relatively complex relational networks. In other words, more complex tasks introduce possible sources of noise (e.g., inappropriate stimulus control) into the measure that may serve to undermine its sensitivity.

In the current study three different IRAP tasks were presented, each of which incorporated more complex stimuli than those employed in the previous studies. The stimuli included, Verbal Analogies (analogy IRAP), basic arithmetic (math IRAP), and perceptual analogies (perceptual IRAP). Details of the sample and target stimuli employed in each of the three IRAPs are presented in Table 8. Consider first, the analogies IRAP. Each of the sample word pairs and target word pairs incorporate relations of either co-ordination or opposition. In order to complete each IRAP trial, the participant must identify the relation that exists between both the sample word pair (e.g., a relation of co-ordination exists between Friend: Pal) and the target word pair (e.g., a relation of opposition exists between Buy: Sell), and then determine if a frame of co-ordination or of opposition exists between the two (see Stewart, Barnes-Holmes, Hayes, & Lipkens, 2001, for a more detailed analysis of relations of analogy).

The math and perceptual IRAP contained non-verbal stimuli, but like the analogies IRAP they too involve identifying relations between relations. In the math IRAP, participants must work out the solution to the target relation (e.g., 15 – 7), and then determine whether the solution is equivalent to the sample stimulus (e.g., < 10). In the perceptual IRAP, the sample and target stimuli consist of a series of abstract symbols. One group of stimuli consist of symbols (or sets of symbols) of equal quantity, whereas the second group consist of stimuli that increase (double) in quantity. Completing the perceptual IRAP trials involves identifying the relation that exists between the target symbols and identifying whether this is equivalent or different to the sample.

In order to determine whether performances on these more complex IRAP tasks would correlate with intelligence, two cognitive abilities tests were employed in the current study. These were the AH4 (Heim, 1970) and the *Cognitive Abilities Test – Third Edition* (CAT3; Lohman, Thorndike, & Hagan, 2001). The AH4 was employed because it provides a useful fluid measure of intelligence (see Chapter 3), and also because it will allow for a relatively direct comparison to be made between the previous and current studies. The CAT3 is designed to assess an individual‘s ability to reason and manipulate different types of symbolic information. It is comprised of three test batteries: a Verbal Battery, a Quantitative Battery, and a Non-verbal Battery. Each of the test batteries contains a further three subtests. Given the time required to administer the entire test (up to 3 hours), only one subtest was administered from each of the three test batteries in the current study.

The CAT3 was considered a useful measure in a number of respects. First, the authors have stated that in designing the CAT3 they have "attempted to emphasize relational thinking – the perceiving of relationships among elements in a variety of media and settings" (Lohman, et. al, 2001: pg. 1). Thus, the test has been designed so that the items contained in each of the subtests are relatively simple, clear and familiar. What is required to complete the items is not knowledge in particular domains, but the identification of various relationships, and flexibility in manipulating those relationships. Thus, there would appear to be an obvious overlap between the processes that are tapped into by the CAT3 and those involved in successfully completing the IRAP tasks.

Another benefit to using the CAT3 is that, like the AH4, it is a group-based measure of cognitive ability. Employing two group-based measures of intelligence meant that we could conduct the study with a group of participants rather than individually. This allowed us to gather a relatively large sample of participants and conduct the study in a relatively short time-frame.

In sum, the current study employed three IRAP tasks, which were more complex than those in the previous studies (i.e., they each involved relating relations). Two group-based cognitive tests were subsequently employed. Given that the current IRAPs were notably different from those previously employed; no specific predictions were made prior to analyzing the data.

Table 8. Sample and Target Stimuli Used in the Analogy, Math, and Perceptual IRAP tasks.

Analogy IRAP		**Math IRAP**		**Perceptual IRAP**	
Sample 1	**Sample 2**	**Sample 1**	**Sample 2**	**Sample 1**	**Sample 2**
Friend : Pal	Top : Bottom	< 10	> 10	< < <	< << <<<<
Targets congruent with Sample 1	**Targets congruent with Sample 2**	**Targets congruent with Sample 1**	**Targets congruent with Sample 2**	**Targets congruent with Sample 1**	**Targets congruent with Sample 2**
Gift : Present	Buy : Sell	15 – 7	26 - 9	% % %	% %% %%%%
Shop : Store	Loose : Tight	4 + 2 + 6	5 + 4 + 2	+++ +++ +++	^ ^^ ^^^^
Fast : Quick	Up : Down	2 + 6 + 1	3 + 2 + 6	**** **** ****	▯ ▯▯ ▯▯▯▯
Loud : Noisy	Tall : Short	3 X 3	6 X 2	/// /// ///	** **** ********
Easy : Simple	Wide : Narrow	21 - 12	24 - 12	^^ ^^ ^^	((((((((((((((
Triumph:Success	East : West	24 / 4	48 / 4	▯▯ ▯▯ ▯▯	// //// ////////

Method

Participants

Participants were 67 undergraduate students at the National University of Ireland, Maynooth, who were all studying psychology at honours degree level.

Students participated in the current study as part of a practical demonstration for which they were required to write a practical report to receive course credits. None of the participants had previous exposure to the IRAP protocols used in the current study. The study was conducted across two group-based sessions. The three IRAPs were presented in session 1 and the cognitive abilities tests were presented in session 2. Four participants failed to return for the second session, and their data are not included in the final analyses. A further two participants failed to complete the CAT3 subtests and their data for these measures are missing from the final data set. The final data set therefore consisted of 63 participants (18 males, 45 females), 61 of whom completed the entire study. Participants ranged in age from 18-54 years (mode = 19; *M* = 21).

Materials and Stimuli

Materials included the IRAP software, a participant profile form, the AH4, and the CAT3. The IRAP tasks differed in certain respects from those employed in Studies 1 and 2, and will be discussed further in the *Procedure*. Three separate IRAP tasks were employed that differed only in terms of the stimuli that were used. Details of these stimuli are presented in Table 8. The participant profile form was identical to that employed in Studies 1 and 2, and the AH4 was described in Chapter 3.

The CAT3 is a group-based test which assesses the ability to reason and manipulate symbolic information. The CAT3 contains eight different levels, labelled A to H, which are suitable for different age-groups. Level H was selected for administration in the current study because it is targeted for individuals from 15 years through to adulthood. The CAT3 contains three test batteries – a Verbal Battery, a Quantitative Battery and a Non-verbal Battery. Each of the test batteries contains three subtests. For the purposes of the current study, one subtest form each of the test batteries was selected for administration. These were the Verbal Analogies subtest (Verbal Battery), the Number Analogies subtest (Quantitative Battery), and the Figure Classification subtest (Non-verbal Battery). Details of the type of items contained in each of these subtests are presented in Table 9. Each of the subtests is timed -- nine minutes is allocated for the completion of the Verbal Analogies subtest, 12 minutes for the Number Analogies subtest, and 10 minutes for the Figure Classification subtest. The CAT3 reports very high levels of reliability. Internal consistency estimates range from .90-.94 (Deary, Strand, Smith & Fernandes, 2007)

Table 9. Examples of the Type of Items Included in the Verbal Analogies, Number Analogies and Figure Classification subtests of the CAT3. Specific exemplars are fictitious but representative of actual items

Verbal Analogies	*Choose the missing word to complete the analogy:* *Parrot → Bird : Saw →* *A. Hammer; B. Tool; C. Sharp; D. Blade; E. Hawk*
Number Analogies	*Work out how the first two number pairs are linked together, and then choose the number that completes the third pair* *[27 → 24] [13→10] [8 →?] A.5; B. 2; C. 10; D. 9; E. 0*
Figure Classification	This subtest involved identifying a similar feature across three given figures and then selecting one of five additional figures that goes with the other three

Procedure

The study was conducted across two sessions. In session 1, participants completed the three IRAP tasks, and in session 2 they completed the AH4 and the three subtests from the CAT3. Both sessions were conducted in the Department of Psychology at NUI Maynooth -- session 1 was conducted in a computer laboratory and session 2 in a large room used for teaching. Both sessions were group-based and were conducted on separate days of the same week. Neither of the sessions lasted for more than two hours on any given day.

Session 1: IRAPs

The procedure for conducting the current study was similar to that across Studies 1 and 2 (i.e. participants filled out the profile form, received verbal instructions, and then proceeded through the written instructions, practice blocks, and test blocks, as described in Chapter 2). However, there were some important differences between the current and the previous studies. First, three IRAP tasks were presented. The order of presentation of these tasks was counterbalanced across participants, but all participants started each IRAP on a consistent block (and then alternated between inconsistent and consistent in the manner described in Chapter 2). Second, each of the IRAPs employed 12 target stimuli, six of which were congruent with sample 1, and six of which were congruent with sample 2 (see Table 8). This meant that each IRAP block consisted of a total of 24 trials (the IRAPs in Studies 1 and 2 contained 8 pairs of target stimuli, and thus consisted of 32 trials per block). This procedure shortened the length of each IRAP task, and allowed us to incorporate three IRAPs, rather than two. Third, in each of the IRAP tasks a median response latency of ≤ 3,500 ms was selected as a criterion for entering the test phase. In addition, if a participant failed to respond within this time limit on any given IRAP trial, then the words "Too Slow" appeared in red letters directly beneath the target stimulus. The "Too Slow" feedback disappeared once the participant pressed a key on the keyboard (if they pressed the wrong key, a red 'X' appeared, and if they selected the correct key then a new trial appeared after a 400ms intertrial interval). This feedback was included in the current study based on findings from previous IRAP research within our laboratory, which suggested that participants tended to respond far more slowly in group-based settings. When participants respond relatively slowly on the IRAP the difference in response latencies between consistent and inconsistent blocks of trials becomes unreliable (see Barnes-Holmes, Murphy, Barnes-Holmes, & Stewart, 2010), and thus maintaining relatively rapid responding in a group setting was deemed critical.

A final difference between the current and previous studies was that the current session was conducted with a group of participants, rather than individually. The session took place across a two-hour period (2-4 pm). Half of the participants were asked to arrive in the first hour (i.e., 2pm) and the other half were asked to arrive in the second hour (i.e., 3 pm). Five experimenters were present throughout the session to facilitate the loading of the IRAP tasks and address any questions that participants may have had. Prior to commencing the session, the first IRAP task was loaded, with the instruction display page open. When participants entered they were asked to fill out the subject profile form, and then verbal instructions were given to the group in relation to the procedure for completing the tasks. Participants were then asked to read the instructions presented on the display page and then start the tasks. They were also instructed to raise their hand when they finished the task and an experimenter would come to load the next IRAP. When a participant had completed the three IRAP tasks, they were thanked, and told that that they could leave. The procedure for those who arrived in the second hour was the same as that described above.

Session 2: Cognitive Abilities Tests

The cognitive abilities tests were administered in a group-based session. All participants were seated at a desk, in a bright well ventilated room. Three experimenters were present throughout the session. The AH4 was administered first. All participants were given a test booklet and an answer sheet. The test was administered in accordance with the administration manual (Heim, 1970) and as described in Chapter 3. When the 10 minutes for completion of Part II had elapsed, the question books and answer sheets were collected.

Following this, the CAT3 booklets were distributed to each participant. Participants completed the Verbal Analogies subtest first, and then the Number Analogies subtest, and finally the Figure Classification subtest. At the start of each subtest, participants were directed to the relevant page of the booklet by the experimenter. Each subtest contains written instructions and includes an example with an answer marked. The experimenter read the instructions and explained the sample question, in line with the guidelines for administration, presented in the

CAT3 manual. The question booklet also contains practice questions for examinees to attempt prior to starting each subtest. Participants were instructed to respond to the questions by circling the answer that they thought was correct. Participants were given time to answer the sample questions, and then the experimenter instructed them to start the relevant subtest. The time limits for completion of each subtest were in accordance with those recommended in the administration manual (see *Material and Stimuli* Section). When the time for completion of a subtest had elapsed the experimenter instructed the participants to stop working. Participants were then directed to the instructions for the subsequent subtest. When the time for completion of the third subtest had elapsed, the participants were thanked and the booklets were collected. That marked the end of session 2. Participants were fully debriefed at a follow-up session which was held the following week.

Results

Preliminary Data Screening

The preliminary data screening was similar to that employed across Studies 1 and 2. Only the data for those participants, who maintained the $\geq$ 80% correct responding criterion across test blocks, are included in the current analysis. Thirty-seven participants met this criterion on the analogies IRAP, 23 on the math IRAP, and 47 on the perceptual IRAP. As in Studies 1 and 2, the response latency data were checked for outliers (defined as latency data that exceeded two standard deviations of the corresponding group mean). None of the data exceeded this criterion, and thus no further data were removed from the analyses. The results of normality tests (Kolmogorov-Smirnov) suggested that the data was roughly normally distributed, and thus parametric statistics were employed[3].

Initial IRAP Analyses

The overall mean response latencies for each of the IRAPs are presented in Figure 5. As illustrated, the consistent blocks were shorter than their respective inconsistent blocks across all three IRAPs: analogy (consistent, *M* = 2083, *SE* = 45 versus inconsistent, *M* = 2245, *SE* = 49); math (consistent, *M* = 2309, *SE* = 51 versus inconsistent, *M* = 2515, *SE* = 62); and perceptual (consistent *M* = 1540, *SE* = 35 versus inconsistent, *M* = 1762, *SE* = 45). Paired-samples t-tests showed that these differences were statistically significant; analogy ($t[36] = -4.1$, $p = .0002$, $\eta^2 = .32$), math ($t[22] = -5.8$, $p <.0001$, $\eta^2 = .6$), and perceptual ($t[46] = -10.01$, $p <.0001$, $\eta^2 = .68$) . Thus, as predicted, the current study produced typical IRAP effects.

The mean response latencies on the consistent blocks were shortest on the perceptual IRAP task (*M* = 1540) relative to the analogies task (*M* = 2084), and the math task (*M* = 2309). The same pattern was observed for the inconsistent blocks: inconsistent response latencies on the perceptual IRAP were shortest (*M* = 1762) relative to the analogy IRAP (*M* = 2245), and the math IRAP (*M* = 2515) IRAPs. These results suggest that the perceptual IRAP was easiest to complete (because it required the least amount of time), and the math IRAP was the most difficult. Consistent with Studies 1 and 2, a difference-score was calculated for each IRAP.

---

[3] The Kolmogorov-Smirnov statistic was significant for the inconsistent perceptual IRAP data ($p = .02$) and for the difference-scores obtained on the math IRAP ($p = .004$). Non-parametric statistical analyses were conducted with these data sets, the results of which were similar to those obtained from inferential analyses. Thus, only the inferential analyses are reported.

The analogy IRAP recorded the smallest difference-score (*M* = 166, *SE* = 49) followed by the math IRAP (*M* = 206, *SE* = 62). The perceptual IRAP had the largest difference-score (*M* = 223, *SE* = 22). Given that only twelve participants successfully completed all three IRAPs, inferential statistical comparisons were not conducted on these data.

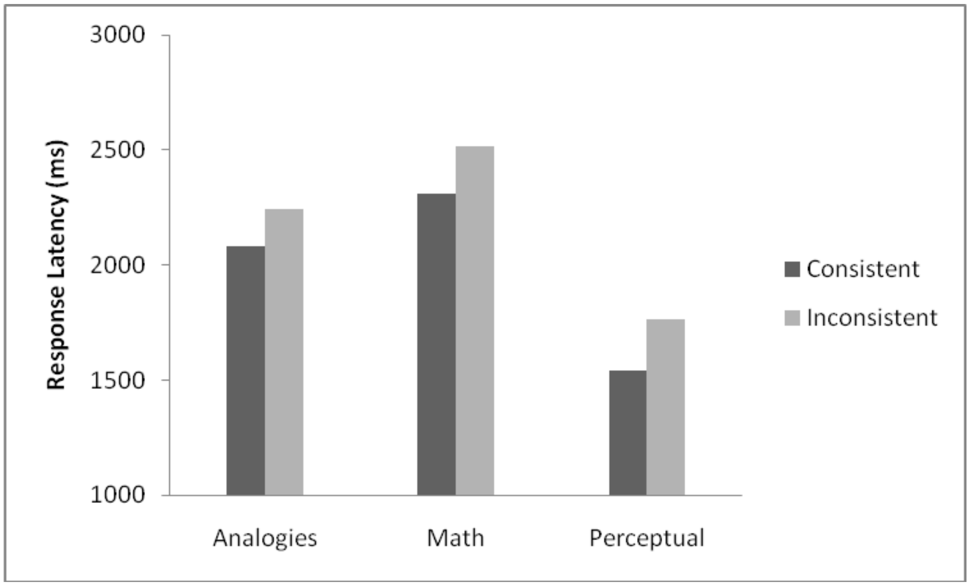

Figure 5. Overall mean response latencies for consistent and inconsistent blocks on the analogies, math, and perceptual IRAPs.

AH4 and CAT3 Data

The data in Table 10 presents the descriptive statistics obtained for performances on the AH4 and the CAT3. The range of scores represents the actual raw scores obtained by participants (standardised scores are not provided in either of the test manuals). Thus, in order for the data to be meaningfully interpreted, the number of items contained on each subtest (and the percentage correct) is presented in the final column of the table. The data obtained for the CAT3 shows that on average, participants got more items correct on the Verbal Analogies subtest (*M* = 23/30) relative to the Number Analogies (10/20) and Figure Classification (12/24) subtests. On the AH4, the data show that participants produced more correct responses on Part II relative to Part I. This pattern was also observed for the AH4 in Study 2. However, the data also reveal that the current mean scores for the AH4 are lower than those obtained in Study 2, and the range of scores and standard deviations obtained here are also larger than those observed in Study 2.

Table 10. Descriptive Statistics of Performance on the CAT3 and the AH4.

	N	Mean	SD	Range	Number of items in test and percentage correct
CAT3					
Verbal Analogies	61	23	4	13-29	30 (76%)
Number Analogies	61	10	5	2-19	20 (50%)
Figure Classification	61	12	5	1-23	24 (50%)
AH4					
Overall	63	92	14	60-124	130 (71%)
Part I	63	45	7	28-60	65 (69%)
Part II	63	48	9	29-64	65 (74%)

Correlations Between the IRAP Measures and the Cognitive Ability Measures

The correlations obtained between the IRAP measures and the cognitive abilities tests are presented in Table 11. On the analogies IRAP, all correlations are very weak and non-significant. On the math IRAP, some significant correlations were observed for consistent and inconsistent trials, and one significant correlation was obtained for the difference-score. On the perceptual IRAP a number of significant correlations were obtained for the consistent and inconsistent measures. However, the difference-score did not produce any significant correlations.

Overall the pattern of correlations obtained diverges from those reported in Studies 1 and 2. Specifically, unlike the previous studies, the difference-score does not emerge as a significant predictor of intelligence across the IRAPs. In fact, only one correlation between the differences scores and the cognitive abilities measures reached statistical significance (i.e., the correlation between the math difference-score and Part I of the AH4). In addition, the correlations with the consistent and inconsistent measures were only observed for one IRAP with most of the correlations occurring for the AH4. Finally, the analogies IRAP produced no significant correlations with the intelligence measures.

Table 11. Correlations between the IRAP Measures and the Intelligence Measures

	Analogy IRAP			Math IRAP			Perceptual IRAP		
	Con	Incon	Diff. Score	Con	Incon	Diff. Score	Con	Incon	Diff. Score
CAT3									
Verbal Analogies	-0.26	-0.03	0.27	-0.5*	-0.54*	-0.23	-0.37*	-0.35*	-0.12
Number Analogies	-0.1	0.04	0.17	-0.32	-0.22	0.09	-0.15	-0.16	-0.09
Figure Classification	-0.17	-0.06	0.18	-0.28	-0.23	0.0001	-0.28	-0.22	-0.06
AH4									
Overall	-0.06	-0.13	0.05	-0.42	-0.5*	-0.28	-0.51**	-0.47**	-0.17
Part I	-0.18	0.05	0.05	-0.39	-0.6*	-0.51*	-0.42**	-0.42**	-0.21
Part II	-0.08	-0.04	0.06	-0.37	-0.31	0.003	-0.45**	-0.39**	-0.09

*p < .05. **p < .01

Group Comparisons

Given the inconsistent pattern of results that emerged, it was deemed necessary to investigate further the relationship between IRAP performances and scores on the cognitive abilities tests. For the purpose of the current analysis, participants were divided into four groups: Group 1 was composed of participants who failed to complete all three IRAPs and thus were not included in the foregoing analyses (*N* = 13); Group 2 included those who successfully completed one IRAP task (*N* = 16); Group 3 included those who successfully completed two IRAPs (*N* = 20); and Group 4 included participants who successfully completed all three IRAPs (*N* = 12). Failing an IRAP is defined as not reaching the ≥ 80% accuracy criterion to enter the test phase, or not maintaining that criterion across test blocks. The mean test scores and standard deviations obtained by participants in each of the four groups are presented in Table 12. The data show that participants who successfully completed more IRAPs tended to score higher on each of the cognitive abilities measures.

Table 12. Mean Test Scores and Standard Deviations Obtained for Participants in each of the Four Groups

	Group 1		**Group 2**		**Group 3**		**Group 4**	
	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
CAT3								
Verbal Analogies	21	4	22	4	23	4	25	3
Number Analogies	9	5	9	4	9	5	11	5
Figure Classification	9	4	11	5	14	3	15	4
AH4								
Overall	85	13	92	14	92	14	102	12
Part I	41	6	44	7	43	8	50	7
Part II	45	8	48	9	49	8	52	8

Three one-way between-groups ANOVAs were conducted to investigate whether the differences in test scores across the groups were statistically significant. Taking first the test scores on the CAT3, the results indicated that on the Verbal Analogies subtest there was a significant difference in test scores between the groups [F (3, 57) = 2.86, p = .04, η^2 = .15]. Post hoc comparisons using Fisher's PLSD tests indicated a significant difference between Group 1 and Group 4 (p = .005). The differences between the other groups were not found to be statistically significant. On the Number Analogies subtest, there was no significant main effect [F (3, 57) = .62, p = .6], and thus no post hoc comparisons were conducted. On the Figure Classification subtest there was a significant main effect for the four groups [F (3, 57) = 7.1, p = .0004, η^2 = .37]. Post hoc comparisons revealed significant differences between Groups 1 and 3 (p = .0006), Groups 1 and 4 (p = .0003); Groups 2 and 3 (p = .01); and Groups 2 and 4 (p = .006).

Three one-way ANOVAs were also conducted to compare scores obtained on the AH4. There was a significant main effect for the overall score [F (3, 57) = 3.1, p = .03, η^2 = .16]. Post hoc comparisons indicated significant differences between Groups 1 and 4 (p = .003). For Part I there was a significant main effect [F (3, 57) = 3.7, p = .016, η^2 = .19]; and post hoc comparisons revealed significant differences between Groups 1 and 4 (p = .002); Groups 2 and 4 (p = .04); and Groups 3 and 4 (p = .01). On Part II of the AH4 a significant main effect was not observed [F (3, 57) = 3.1, p = .19]; and post hoc tests were not conducted.

In sum, the results of the current study revealed that the IRAP protocols produced typical IRAP effects, which is consistent with the findings of Studies 1 and 2. However, the pattern of results obtained on the correlational analyses was dissimilar to those observed across the previous studies, with only one significant correlation with a difference-score, and one IRAP yielding no correlations at all. The subsequent group comparisons revealed that there was a significant relationship between IRAP performance and scores on the cognitive abilities measures. Specifically, those who successfully completed more of the IRAP tasks tended to score higher on the abilities measures.

Discussion

The current study employed three IRAP protocols, which consisted of more complex relational stimuli than those used in Studies 1 and 2. Intuitively one might expect that such protocols would produce relatively strong correlations with the cognitive abilities measures, but the results showed that this was not the case. Of particular note is the fact that relational flexibility (as measured by the difference-score) was not a significant predictor of cognitive ability, and also that no significant correlations were observed between the analogies IRAP and the abilities measures. However, although the correlational analyses produced a different pattern of results compared to those of the previous studies, the subsequent group comparisons revealed that participants who completed more IRAPs tended to perform better on the cognitive abilities measures. This indicates that IRAP performance (i.e., the number of IRAPs successfully completed) was discriminative for cognitive ability.

There are a number of factors that may account for the different pattern of results that were obtained in the current study relative to the two previous studies. First, as highlighted in the Procedure, feedback was provided throughout the IRAP tasks to remind participants to respond quickly (i.e., if participants took longer than 3000 ms to respond the words “Too Slow” appear on the screen). Given that the study was conducted in a group-based session, it was deemed important to include this feedback, in order to keep participants focused on the task. However, in hindsight, the criterion of responding within the three second time-frame was rather stringent. Participants who failed to meet this *latency* criterion did not proceed to the critical test blocks even though they may have achieved the *accuracy* criterion. The exclusion of participants on this basis may have restricted the range of scores in the dataset, resulting in an inaccurate or unreliable estimate of the relationship between IRAP and IQ performances (see Zimmerman & Williams, 2000).

Second, the pattern of results obtained may have been influenced by the relatively complex stimuli employed in the current study. The IRAP protocols required participants to relate relational networks to other relational networks. Furthermore, the IRAPs, particularly the analogies task, also required knowledge of particular content (e.g., that East and West are opposite), and thus it is possible that these added requirements may have impacted upon the sensitivity of the IRAPs’ critical measures (speed and flexibility). In other words, as the complexity of the task increases, the measures of speed and flexibility of responding become less sensitive, because other controlling variables unrelated to speed, such as general knowledge, come into play. Post-hoc support for this argument is provided by the fact that the analogies IRAP, which relied most heavily on general knowledge, produced no correlations at all. In contrast, the math IRAP, which targeted only a very specific knowledge domain, produced five correlations. And finally, the perceptual IRAP, which was relatively content free, produced eight correlations.

The results of the current study appear to suggest that the IRAP, as a measure of human intelligence, has particular sensitivity to relatively basic relational frames. Thus in the subsequent and final study of the current research programmed, it was deemed appropriate to once again investigate more basic relational frames using the IRAP. However, the subsequent study involved investigating two additional relational frames (hierarchical and comparative relations) which have not previously been systematically examined.

CHAPTER 5

Study 4: Hierarchical and Bigger/Smaller Relational Responding and Performances on the WAIS-III^{UK} and the AH4

The first two studies of the current research programme employed basic relational frames (coordination, distinction and temporal frames) in the IRAP, whereas Study 3 focused on more complex relations (analogies and arithmetic). The purpose of Study 4 was to return to the more basic relational frames but to extend the analysis to two additional frames not previously investigated. Thus, in the current study, comparative (bigger/smaller) and hierarchical frames were targeted in the IRAPs. Consistent with the previous studies, cognitive abilities tests (the WAIS-III^{UK} and the AH4) were also administered.

Comparative and hierarchical frames were targeted in the current study given that they are generally considered to be critical in human cognitive development (Hayes, et al., 2001; see also Jameson & Gentner, 2003; Goswami, 1991). According to Hayes and colleagues (2001), comparative relations involve responding to one event in terms of a quantitative or qualitative relationship with another event. The relation of bigger/smaller was chosen for the current study, but additional subtypes of comparison also exist (e.g., better/worse, more than/less than). Hierarchical frames are identified as sharing the same basic relational pattern as comparative relations but the relationship between members of a category tends to be more specific. Categorising roses, tulips and daffodils as flowers involves hierarchical relational responding, and thus it can be viewed as providing an important type of contextual control over responding on the basis of coordination. Thus, in the example above, three types of flowers are coordinated on the basis of the category label, flower.

The specific stimuli employed in the bigger/smaller and hierarchical IRAPs are contained in Table 13. The target stimuli on both IRAPs differed in one important respect form those employed in the previous IRAPs. In the previous studies the IRAPs presented tasks that may have facilitated participants responding on the basis of only one element rather than the relation between elements. For illustrative purposes, consider the following two trials on the before-after IRAP; Before/Child-Adult (True) and After/Child-Adult (False). With sufficient practice, which the IRAP provides, a participant might simply respond *True* whenever they see *Before* and *Child,* and *False* whenever they see *After* and *Child.* In other words, the participant comes to respond differently to two compound stimuli (*Before-Child* and *After-Child*), rather than the relation between *Child* and *Adult.* In the current study, therefore, the target stimuli were altered to help ensure that participants were responding to the relations that pertain between the target word-pairs. Thus, in both IRAPs one word from each target word-pair was presented again in a second word-pair. For instance, in the bigger/smaller IRAP the word “Horse” was used across two word-pairs; in a smaller-than relation with the word “Elephant”, and in a bigger-than relation with “Dog”. Similarly, in the Hierarchical IRAP the word “Car” represented a hierarchical category when presented with “Toyota”, but represented a member of the hierarchy when presented with “Vehicle”. This measure therefore went some way towards ensuring that participants responded to the relations that pertained between the words on each trial.

The current Study represents a replication and extension of Study 2. An interesting finding from that study was that the similar/different and before/after relational responding appeared to tap into different aspects of ability as measured by the WAIS-III^{UK}. Similar/different relational responding loaded more strongly onto the Verbal Comprehension index, whereas before/after relational responding tapped into the Perceptual Organisation index. Given that no study had explored the relationship between intelligence and the relational frames employed in the current study, no specific predictions were made in this regard. However, similar to before/after relations, comparison and hierarchy are characterised by asymmetry. For example, in the case of before/after, if A occurs before B, then B does *not* occur before A; in the case of comparison, if A is bigger than B then B is *not* bigger than A; and in the case of hierarchy, if A contains B, then B does *not* contain A. In contrast, similar/different relations are symmetrical in this sense (i.e., if A is similar to B then B *is* similar to A; and if A is different to B, then B *is* different to A). On this basis, therefore, one might predict that comparison and hierarchy will overlap functionally, to some extent, with temporal relations. If so, it would be expected that the IRAP measures would correlate with Perceptual Organisation rather than Verbal Comprehension.

Previous studies in the current research programme, as well as those conducted by O’Hora et al., produced weak and non-significant correlations with the Working Memory index. However, given that the target stimuli were altered in the current study to help ensure that participants responded to the relation between word-pairs no predictions were made with regard to working memory (i.e. it is possible that the altered target word-pairs would place a greater load on working memory). However, as in previous studies, it was expected that the response latency data would correlate significantly with the Processing Speed subtests.

Table 13. Sample and Target Stimuli Employed in the Bigger/Smaller and Hierarchical IRAPs

Bigger/Smaller IRAP		**Hierarchical IRAP**	
Sample 1	**Sample 2**	**Sample 1**	**Sample 2**
Bigger	Smaller	Category → Member	Member → Category
Targets congruent with Sample 1	**Targets congruent with Sample 2**	**Targets congruent with Sample 1**	**Targets congruent with Sample 2**
Elephant Horse	Horse Elephant	Vehicle Car	Car Vehicle
Horse Dog	Dog Horse	Car Toyota	Toyota Car
Aeroplane Bus	Bus Aeroplane	Animal Dog	Dog Animal
Bus Car	Car Bus	Dog Terrier	Terrier Dog
Tree Shrub	Shrub Tree	Plant Tree	Tree Plant
Shrub Flower	Flower Shrub	Tree Oak	Oak Tree
Suitcase Handbag	Handbag Suitcase	Fruit Orange	Orange Fruit
Handbag Purse	Purse Handbag	Orange Satsuma	Satsuma Orange

Method

Participants

Participants were 51 undergraduate students from various faculties of the National University of Ireland, Maynooth. None of the participants had previous exposure to the IRAP protocols used in the current study (although some had completed different IRAP tasks through their participation in other research being conducted in the Department). The study was conducted across two separate sessions -- with the IRAPs presented in session 1, and the cognitive abilities tests in session 2. Eleven participants failed to complete the study (i.e., they did not complete the IRAPs and/or failed to return for the second session), and their data are not included in the final analyses. Consistent with the previous studies, only the data from participants who maintained the ≥ 80% correct responding criterion across the six IRAP test blocks were included in the final analyses. All 40 participants (17 male; 23 female) who completed the study met this criterion on both of the IRAPs. Participants ranged in age from 17-44 years (mode = 19; *M* = 20).

Materials and Stimuli

The materials for the current study included the IRAP software. This was identical in all respects to that employed in Study 2, except that the target and sample stimuli differed (see Table 13). Thus, in the current IRAPs the response latency criterion (≤ 4,500 ms), accuracy criterion (≤ 80%), and the number of practice and test blocks were all identical to those in Study 2. Material for the current study also included a participant profile form, the WAIS-III^{UK}, and the AH4. The profile form is described in Study 1, and the WAIS-III^{UK} and AH4 are described in Study 2.

Procedure

The procedure for conducting the current study was identical to that in Study 2. Only the specific instructions for completing the IRAP tasks differed to take account of the particular sample and target words that were employed in the current study.

Results

Preliminary IRAP analysis

Consistent with the previous studies, the response latency data from participants who completed the study were checked for outliers (i.e., response latencies that exceeded two standard deviations of the corresponding group mean). The data from two participants were removed on this basis. The response latency and IQ data from the remaining 38 participants were investigated using the Kolmogorov-Smirnov test. The results showed no violation of the assumption of normality (all *p*'s > .08).

IRAP Data

As shown in Figure 6, the overall mean response latencies for the consistent blocks on both IRAPs were shorter than on their respective inconsistent blocks (hierarchy, $M = 2896$, $SE = 122$ versus $M = 3304$, $SE = 152$; and bigger/smaller, $M = 2584$, $SE = 92$ versus $M = 2932$, $SE = 109$). Paired-samples t-tests showed that both differences were significant; hierarchy ($t[37] = -7.42$, $p < .0001$, $\eta^2 = .6$) and bigger/smaller ($t[37] = -8.35$, $p < .0001$, $\eta^2 = .65$). Thus, the current study produced a typical IRAP effect, consistent with that observed in previous IRAP studies.

The mean response latencies on the consistent blocks were shorter on the bigger/smaller task ($M = 2584$) relative to the hierarchy task ($M = 2896$), and the same pattern was observed for the inconsistent blocks (bigger/smaller, $M = 2932$

versus hierarchy, *M* = 3304). Paired-samples t-tests indicated again that these differences were significant; consistent (*t*[37] = 4.37, *p* <.0001, η^2 = .34) and inconsistent (*t*[37] = 3.54, *p* = .001, η^2 = .25). These results indicate that the hierarchical IRAP tasks required more time than the bigger/smaller tasks.

Consistent with previous studies, a difference-score was calculated for each participant by subtracting response latencies on consistent trials from latencies on inconsistent trials for each of the IRAPs. The mean difference-score on the hierarchical IRAP was greater than that for the bigger/smaller IRAP (hierarchy = 409, and bigger/smaller = 347). However, the difference between the two sets of difference-scores was not statistically significant (*t*[37] = 1.1, *p* = .29).

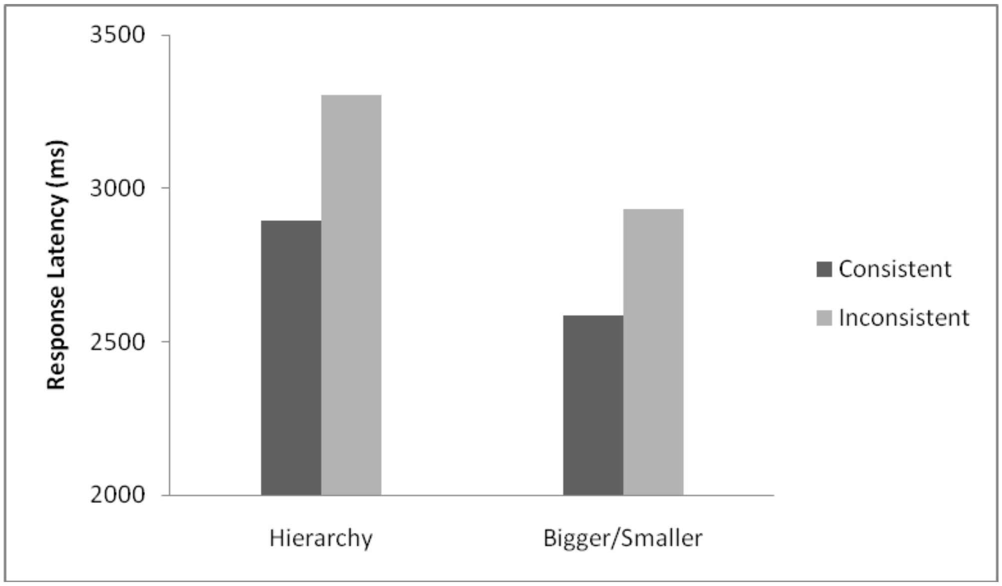

Figure 6. Overall mean response latencies for consistent and inconsistent clocks for the hierarchical and bigger/smaller IRAPs.

IQ Data

The descriptive statistics obtained in the current study are presented in Table 14. As expected (given that participants in the current study were university students) the average scores obtained on the WAIS-III^{UK} were above the general population averages. The data obtained is also somewhat less variable than that associated with the general population. These data are consistent with those obtained in Study 2. Also consistent with Study 2, are the significant correlations that were obtained between performances on the two cognitive abilities tests: correlation coefficient for Full-Scale IQ with the overall score on the AH4 = 0.69 ($p < 0.0001$); Verbal IQ with part I of the AH4 = 0.37 (p =0.02); and Performance IQ with Part II of the AH4 = 0.47. ($p < 0.002$).

Table 14. Descriptive Statistics for Performances on the WAIS-III^{UK} and the AH4

	N	Mean	SD	Range
WAIS-III^{UK}				
IQ Scores				
Full-Scale IQ	38	120	11	99-143
Index Scores				
Verbal Comprehension Index	38	123	12	101-150
Perceptual Organisation Index	38	114	12	88-138
Working Memory Index	38	113	15	86-139
Processing Speed Index	38	112	11	86-137
Subtest Standard Scores				
Similarities	38	14	3	7-19
Vocabulary	38	14	2	10-19
Information	38	14	2	9-19
Comprehension	38	13	2	8-17
Picture Arrangement	38	12	2	9-19
Picture Completion	38	11	3	6-18
Block Design	38	13	3	8-18
Matrix Reasoning	38	13	2	10-17
Arithmetic	38	13	2	10-16
Digit Span	38	12	3	6-19
Letter-Number Sequencing	38	12	3	7-17
Digit-Symbol Coding	38	12	2	6-16
Symbol Search	38	12	2	8-19
AH4 (raw scores)				
Overall	38	104	13	74-124
Part I	38	46	8	29-60
Part II	38	58	7	38-65

Correlations Between the IRAPs and Cognitive Abilities Measures

The correlations that were obtained in the current study are presented in Table 15. Most of the correlations are negative, although some positive but very weak correlations are obtained on each of the IRAPs (e.g., between the picture arrangement subtest and the three hierarchical IRAP measures; and the correlation between the similarities subtest and the bigger/smaller difference-score). The correlations will be discussed under three headings.

Correlations between the hierarchical IRAP and the WAIS-III$^{\text{UK}}$: Each of the hierarchical IRAP measures (consistent and inconsistent measures as well as the difference-score) produced significant correlations with full-scale IQ. Furthermore, the response latency measures (consistent and inconsistent trials) correlated significantly with the Perceptual Organisation index, the Working Memory index and the Processing Speed index. Analysis of the subtests that load onto these index scores revealed that both Processing Speed subtests (Digit-Symbol Coding and Symbol Search) correlated significantly with the response latency measures. Likewise, each of the Working Memory subtests (Arithmetic, Digit Span and Letter-Number Sequencing) also produced significant correlations. Two of the subtests that loaded on to the Perceptual Organisation subtest were also significant – Block Design and Matrix Reasoning. Only one significant correlation was observed with the subtests that load onto the Verbal Comprehension index (i.e., the correlation between the Information subtest and the consistent IRAP measure).

The difference-score correlated significantly with Full-Scale IQ and the Perceptual Organisation index. The Matrix Reasoning subtest in particular contributed to this correlation. Unlike Study 2, significant correlations were also observed with the Processing Speed index (and the Digit-Symbol Coding subtest in particular)

Correlations Between the Bigger/Smaller IRAP and the WAIS-III$^{\text{UK}}$: The response latency measures on the bigger/smaller IRAP correlated significantly with

Full-Scale IQ. They also correlated significantly with each of the four index scores with only one exception (the correlation between the consistent measure and the Verbal Comprehension index fell just outside the range of significance $p = .06$). Analysis of the individual subtests revealed that Information, Matrix Reasoning, and Digit Span all correlated significantly with both of the response latency measures and were the main contributors towards the significant correlations with Verbal Comprehension, Perceptual Organisation and Working Memory respectively. As predicted, the two Processing Speed subtests (and the Processing Speed index) also correlated significantly with the response latency measures.

Only one significant correlation was observed for the difference-score -- the correlation with the Information subtest. All other correlations were weak and non-significant. This pattern of correlations with the difference-score differs from that observed on the hierarchical IRAP as well as those observed across Studies 1 and 2.

Correlations Between the IRAP Measures and the AH4: On the hierarchical IRAP only one of the nine correlations between the IRAP measures and the AH4 was outside the range of statistical significance (the difference-score with Part II, $p = .06$). On the bigger/smaller IRAP the response latency measures correlated with all of the AH4 measures. In contrast, none of the correlations between the AH4 and the difference-score reached statistical significance.

In sum, the two latency scores from both IRAPs produced significant correlations with many of the WAIS-III^{UK} and the AH4 intelligence measures. As expected, response latency correlated significantly with the Processing Speed subtests. The difference-score for the hierarchical IRAP correlated significantly with many of the intelligence measures, including Full-Scale IQ, the Perceptual Organisation index, and the Processing Speed index. In contrast, the bigger/smaller difference-score produced only one significant correlation. Unlike previous studies, both IRAPs produced some significant correlations with the Working Memory subtests.

Table 15. Correlations Between the Hierarchical (HY) and Bigger/Smaller (BS) IRAP Measures with the Intelligence Measures

	HY Consistent	HY Inconsistent	HY Difference Score	BS Consistent	BS Inconsistent	BS Difference Score
WAIS-III^{UK}						
Full-Scale IQ	-0.55**	-0.56**	-0.33*	-0.53**	-0.52**	-0.19
Index Scores						
Verbal Comp.	-0.21	-0.22	-0.15	-0.31	-0.35*	-0.23
Perceptual Org.	-0.46**	-0.53**	-0.44**	-0.35*	-0.34*	-0.12
Working Memory	-0.46**	-0.48**	-0.29	-0.36*	-0.36*	-0.15
Processing Speed	-0.61**	-0.64**	-0.42**	-0.63**	-0.59**	-0.17
Subtest Standard Scores						
Similarities	-0.09	-0.14	-0.18	-0.08	-0.06	0.04
Vocabulary	-0.09	-0.08	-0.21	-0.24	-0.30	-0.26
Information	-0.32*	-0.28	-0.07	-0.44**	-0.52**	-0.40**
Comprehension	-0.31	-0.25	0.006	-0.28	-0.28	-0.11
Picture Arrangement	0.09	0.16	0.24	-0.18	-0.14	0.021
Picture Completion	-0.2	-0.26	-0.26	-0.06	-0.06	-0.007
Block Design	-0.39**	-0.42**	-0.30	-0.31	-0.29	-0.08
Matrix Reasoning	-0.44**	-0.5**	-0.40**	-0.39**	-0.40**	-0.18
Arithmetic	-0.37*	-0.37*	-0.20	-0.25	-0.23	-0.04
Digit Span	-0.47**	-0.49**	-0.31	-0.39**	-0.37*	-0.11
Letter-Number Seq.	-0.34*	-0.36*	-0.22	-0.27	-0.31	-0.22
Digit-Symbol Coding	-0.56**	-0.59**	-0.39**	-0.57**	-0.34*	-0.09
Symbol Search	-0.45**	-0.47**	-0.31	-0.49**	-0.4**	-0.18
AH4 (raw scores)						
Overall	-0.67**	-0.67**	-0.38**	-0.51**	-0.49**	-0.17
Part I	-0.59**	-0.61**	-0.38**	-0.45**	-0.41**	-0.03
Part II	-0.62**	-0.61**	-0.31	-0.48**	-0.49**	-0.27

*p < .05. **p < .01

Discussion

The current study demonstrated that bigger/smaller and hierarchical relational responding, as measured by the IRAP, were significant predictors of performance on intelligence tests. However, the results of the current study are not as uniform as those observed in Studies 1 and 2 of the current research programme. Of particular note is the finding that relational flexibility as measured by the difference-score, predicted intelligence for the hierarchical IRAP but not for the bigger/smaller IRAP (only one subtest – information -- correlated significantly with the bigger/smaller difference-score).

The lack of correlation with the bigger/smaller difference-scores may reflect the fact that some of the participants who took part in the current study had previously completed other IRAP protocols. These other IRAPs included different stimuli than those contained in the current study, but the basic IRAP procedure was identical (i.e., all IRAPs require participants to respond to sample and target stimuli, and to switch back and forth between consistent and inconsistent trials). It is possible therefore that those participants with previous exposure to IRAP protocols may have developed strategies for responding that undermined the targeted relational stimulus control. For instance, one participant reported that on the inconsistent blocks for the bigger-smaller IRAP she simply read the target word pair from right to left rather than vice versa. For illustrative purposes, consider a trial containing the sample word *Bigger*, at the top of the screen, and the target *Elephant-Horse*, in the centre. Participants might typically read this as *Elephant – Bigger –Horse*, and then respond appropriately. On inconsistent trials, the correct response for this trial is *False*. However, if participants read the stimuli from right to left (i.e., *Horse –Bigger- Elephant*) and correctly respond by selecting *False*, then they are effectively providing a consistent response. In this scenario, the difference-score would reflect the length of time required to read from right to left (instead of left to right) rather than relational flexibility per se.

Critically, it is likely that this strategy of reading from right to left, would have been easier on the bigger/smaller IRAP, which contained a single sample word on each trial (i.e., either bigger or smaller) relative to the hierarchical IRAP which included two sample words on each trial. Thus participants would need to read one word pair in one direction and the other word pair in the opposite direction in order to respond consistently on an inconsistent trial. This strategy appears to be particularly difficult to implement and thus seems unlikely, and indeed no participants reported using this approach. In effect, perhaps the hierarchical difference-score produced correlations with the intelligence measures because it captured the critical requirement for flexibility, whereas this was not the case for the bigger/smaller IRAP. On balance, however, it is not possible to determine whether or not this was the case, since detailed and systematic information on the strategies used by participants was not gathered.

The results of the current study showed that both IRAPs tapped into working memory. Conway and colleagues (2002) highlighted that the function of working memory is to “maintain memory representations in the face of concurrent processing, distraction, and/or under attention shifts” (pg. 164). As noted in the Introduction, on each of the IRAPs a single word from each target word-pair was presented again in a second word-pair (see Table 13) to discourage responding to only one word in each pair. This approach was designed to ensure that participants responded to the relation between the words and were thus required to hold this relationship in memory, rather than a single word, in order to respond correctly. Indeed, Halford and colleagues (2005) highlighted that relational responding requires working memory. It is possible therefore that the observed loading onto working memory may reflect the additional attentional demands that result from the increased relational nature of the tasks in the current study.

It is worth also noting that if altering the nature of the target stimuli was indeed responsible for the observed correlations with working memory, then it would seem that the IRAP is quite sensitive in this regard. That is, small changes in the IRAP protocol seem to have the potential to produce significant changes in relation to the cognitive demands of the task. This may well have implications for the design of future IRAP protocols.

In the current study, the working memory correlations were observed for the response latency measures, but not for the difference-score. This suggests that the difference-score controlled for the attentional demands of the task, and supports the conclusions from Studies 1 and 2 that the difference-score represents a rather "pure" measure of relational flexibility. On balance, however, the hierarchical difference-score correlated with the Processing Speed index and Digit-Symbol Coding in particular. Previous studies in the current research programme showed that the difference-scores tended to correlate with intelligence, but that they controlled for the confounding effects of processing speed. It is not clear why significant correlations were obtained for the hierarchical difference-score and not for those of other IRAPs. Perhaps, it is not possible to disentangle completely the domains of speed and flexibility. Specifically, the difference-score may control for speed of responding that is not directly related to the measure of interest, but speed seems to be an inherent or even defining feature of flexibility. That is, if a divergent / “flexible” response is emitted only rather slowly, perhaps the response cannot truly be defined as flexible. This and related issues will be dealt with further in the general discussion.

The results of Study 2 demonstrated that the similar/different IRAP showed differential correlations with the Verbal Comprehension index and the Vocabulary subtest in particular, whereas before/after relational responding correlated with the Perceptual Organisation index, and in particular the Block Design task. Some differential loadings were also observed for the hierarchical and bigger/smaller IRAPs, as measured in the current study. The hierarchical IRAP produced no significant correlations with the Verbal Comprehension subtests, but loaded instead on the Perceptual Organisation index and onto Matrix Reasoning in particular (the latency measures correlated with Block Design). The bigger/smaller IRAP produced differential correlations with one of the Verbal Comprehension subtests (Information), and like the hierarchical IRAP, it also correlated significantly with the Matrix Reasoning subtest, but only for the latency measures.

CHAPTER 6

General Discussion

The primary aim of the current research programme was to investigate the relationship between relational responding and intelligence. Specifically, the current work sought to determine whether proficiency and flexibility in relational responding predicted performances on standardised intelligence tests. The research programme also sought to extend the literature on Relational Frame Theory, and contribute toward a behaviour-analytic account of complex human cognition. In the following section, the findings that emerged from the research will be reviewed. Subsequently, conceptual and applied issues that have arisen as a result of the current work will be discussed.

Summary of the Findings

The current research programme was comprised of four correlational studies, each of which employed the IRAP as a methodology to explore the relationships between relational responding and cognitive abilities. In Study 1, similar/different and before/after relational tasks were presented and participants were subsequently exposed to the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990). The study highlighted that individuals who performed better on the intelligence test were not only faster at responding relationally, but also demonstrated a greater degree of relational flexibility. The findings were particularly impressive given that the K-BIT is a brief IQ test and only includes those subtests thought to be particularly critical in the assessment of intelligence.

Study 1 also highlighted that similar/different relational responding was easier or took less time than before/after responding. This finding is perhaps unsurprising given that the similar/different relation is established very early in a child's verbal repertoire (e.g., Hayes 1991). However, a related finding emerged, which suggested that reversing the similar/different relation (i.e., providing an inconsistent response) was more difficult than reversing the before/after relation. Furthermore, the study indicated that this ability to respond against a highly dominant verbal response (i.e., flexibility, as measured by the difference-score) was predictive of verbal intelligence rather than performance on the non-verbal (Matrices) subtest.

The purpose of Study 2 was to investigate again similar/different and before/after relational responding. However, in Study 2 participants completed an extensive intelligence test (the WAIS-III^{UK}), as well as a brief test of fluid ability (the AH4), and a questionnaire probing for momentary lapses in attention (the CFQ). Results showed that similar/different responding loaded more heavily onto the Verbal Comprehension index of the WAIS-III^{UK}, whereas before/after relational responding loaded onto the Perceptual Organisation index. These findings were true for both response speed measures (i.e., consistent and inconsistent trials) and the measures of relational flexibility (i.e., difference-scores). The results were broadly consistent with those obtained in Study 1, as well as those reported by O'Hora and colleagues (2008). The correlations observed with the Perceptual Organisation index, and in particular with the Block Design subtest, highlight the important role that relational processes play in solving what are traditionally termed “non-verbal” problems.

In Study 3, three very different relational tasks were presented using the IRAP. These tasks were designed to tap into more complex relational networks, and included verbal analogies, arithmetic and perceptual analogies. Although intuitively it might be expected that these tasks would produce stronger correlations with intelligence, the results revealed the opposite. That is, the study produced *less* significant correlations between the relational tasks and the abilities measures, when compared with the two earlier studies.

The final study of the current research programme focused once again on more basic relational frames. Specifically, in Study 4 hierarchical and bigger/smaller relational tasks were presented on the IRAP and then the WAIS-III^{UK} and the AH4 were administered. Consistent with Study 2, the results revealed some differential loadings onto particular domains of intelligence. Hierarchical responding loaded onto the Perceptual Organisation index. The response latency measures correlated with the Matrix Reasoning and Block Design subtests, and the difference-score produced differential correlations with Matrix Reasoning. Fast and flexible responding on the bigger/smaller task predicted performance on one of the Verbal Comprehension subtests (Information), and the response latency measures also correlated significantly with the Matrix Reasoning subtest.

Study 4 also showed that flexible responding on the bigger/smaller task was not predictive of performance on the intelligence tests (only the Information subtest correlated significantly with the bigger/smaller difference-score). This finding is unusual since in each of the other basic IRAP tasks (similar/different, before/after, and hierarchy) significant correlations were observed between flexible responding and intelligence. A further finding which diverged from the results of Study 2 was that both of the IRAPs in Study 4 produced significant correlations with the Working Memory subtests on the WAIS-III ^{UK}.

Specific Issues Arising from the Findings in Each Study

The Difference-score and Relational Flexibility

As argued throughout the thesis, the difference-score provides a relatively pure measure of relational flexibility. In the context of the current research programme, therefore, the difference-score may represent a more useful measure than those provided by the response latency data. Given that flexibility is often regarded as a critical and defining feature of intelligence (e.g., Cattell, 1971) it is perhaps unsurprising that significant correlations between it and the intelligence measures were obtained across the majority of the current IRAP tasks.

In considering the observed difference-score correlations, some interesting patterns emerged across the studies. In the studies that employed the WAIS-III (Studies 2 and 4) all of the difference-scores were predictive of full-scale IQ, with the exception of the bigger/smaller difference-score. However, different patterns of correlations were observed for the index and subtest scores. These are summarised in Table 17. The similar/different difference-score predicted performance on the

Verbal Comprehension Index and the Vocabulary subtest in particular; the before/after difference-score predicted performance on the Perceptual Organisation index and the Block Design task in particular; the hierarchical IRAP also predicted performance on the Perceptual Organisation index but on this IRAP, a significant correlation was observed with the Matrix Reasoning subtest; on the bigger/smaller IRAP the difference-score did not correlate significantly with any of the index scores, but one significant correlation was observed with the Information subtest.

Although these results appear to confirm that the difference-score predicts specific types of intelligence, Study 3 produced only one significant correlation for the difference-score (Math IRAP with Part I of the AH4). As noted previously, this study involved presenting relatively complex relational tasks (analogies, arithmetic, and perceptual analogies). It is possible, therefore, that the more complex relational stimuli introduced additional sources of variance, such as general knowledge, and as a result these tasks were less sensitive to the IRAP's critical measure of flexibility. In addition, the relatively stringent exclusion criteria (i.e. response speed of $\leq$ 3000 ms) may have restricted the range of scores in the data set, resulting in inaccurate or unreliable correlations with the abilities measures. It is also worth noting that the difference-score was not predictive of full-scale IQ for the bigger/smaller IRAP presented in Study 4. On this IRAP, only the Information subtest registered a significant correlation. As discussed previously, however, one reason for the paucity of significant correlations for the bigger/smaller IRAP may be that some participants were "IRAP-savvy", and had developed strategies to facilitate easier responding on the inconsistent trials. Overall however, the data presented in the current thesis do provide reasonably strong support for the argument that relational flexibility is an important feature of specific forms of human intelligence.

Table 17. Summary of the WAIS-III^{UK} index and subtest scores which correlated significantly with the difference-scores on each of the IRAP tasks presented across Studies 2 and 4. Note: VCI = Verbal Comprehension Index, POI = Perceptual Organisation Index, PSI = Processing Speed Index.

Before/after	Similar/different	Hierarchy	Bigger/smaller
Full Scale IQ	Full Scale IQ	Full Scale IQ	Information
POI	VCI	POI	
Block Design	Vocabulary	PSI	
	Picture Arrange.	Matrix Reasoning	

Different Patterns of Correlations for Each of the Relational Tasks: A Comparison of Studies 2 and 4

The information in Table 17 highlights that differential correlations were observed for the various relational tasks. While many of the findings are unsurprising, and support RFT arguments, the following analyses are post hoc, and simply offered in the service of directing future research in the area.

The findings from Study 2 highlighted that similar/different relational responding produced a correlation between the difference-score and the Vocabulary subtest. As noted in Chapter 3, this finding is predicted by RFT in that the relational frame of similarity or coordination provides the functional basis for naming and semantic relations more generally (Barnes-Holmes, et al., 2001). Thus, the correlation with Verbal Comprehension and Vocabulary in particular, supports this core postulate of RFT.

The before/after task produced correlations between the difference-score and the Perceptual Organisation index, and with Block Design in particular. During the Block Design task participants are shown pictures of a series of two-colour designs and are asked to reproduce the designs using blocks. Ostensibly, the tasks would appear to target *spatial*, rather than temporal relations. Nevertheless, a critical part of the task involves placing each block *sequentially* in order to recreate the whole design. Thus the sequential nature of the task may draw on before/after relational responding. Interestingly, similar correlations between before/after relational responding and Block Design were observed by O'Hora and colleagues (2008).

It should also be noted that the response latency measures (consistent and inconsistent blocks) on both the similar/different IRAP in Study 2 and the hierarchical IRAP in Study 4, produced significant correlations with the Block Design subtest. However, neither of these IRAPs produced significant correlations with Block Design on the IRAP's critical measure of flexibility (the difference-score). On balance, the Block Design task involves a timed component (i.e., extra points are awarded for a speedier completion of the tasks), and thus it is likely that the observed correlations with the response latency measures were largely reflective of processing speed, rather than relational flexibility.

Like the before/after task, the hierarchical task also predicted performance on the Perceptual Organisation index. However, on this relational frame correlations were observed for Matrix Reasoning (across both the response latency measures and the difference-score). As part of the Matrix Reasoning task, participants were required to select one of five abstract designs that best completes a pattern matrix. Successfully completing the task involves identifying the common feature across the two parts of the matrix, and thus the task requires identification of non-arbitrary relations of similarity. For instance, to complete the simple pattern represented in Figure 7, one needs to identify what is similar across the two series of shapes, in this case, the left-to-right sequence; *triangle-circle- square.* In this sense, the task involves relating a relation to another relation (Stewart, & Barnes-Holmes, 2004), and this is also what the hierarchy task required on the IRAP. For example, when the sample *Category – Member* was presented with the target *Vehicle – Car*, this involved determining whether the sample relation was the same as the target relation (i.e., is *vehicle* the category and *car* the member?). Thus it is possible that the hierarchy IRAP correlated with matrix reasoning, not because hierarchy *per se* predicts performance on this subtest, but because both tasks involve relating relations. This conclusion highlights an important conceptual issue concerning the relationship, in purely functional terms, between the IRAP tasks and the intelligence subtests. This issue will be considered in a subsequent section.

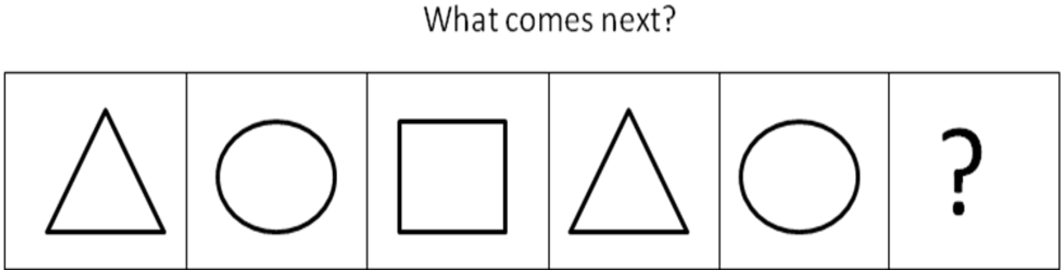

Figure 7. An example of a simple Matrix Reasoning task.

The bigger/smaller task presented in Study 4 produced only one significant correlation with the difference-score, this time with the Information subtest. As highlighted earlier, some participants completing this task may have developed strategies to facilitate easier responding on the inconsistent trials, and this makes it difficult to interpret the one significant result that emerged.

Comparison with Previous Studies of Relational Responding and Intelligence

The current research programme extended the work of O'Hora and colleagues (2005, 2008). In that previous research, before/after relational tasks were presented using the Relational Evaluation Procedure (REP; see Chapter 1). In the 2005 study participants subsequently completed three subtests of the WAIS-III (Vocabulary, Arithmetic, and Digit-Symbol Coding), and in the 2008 study the entire WAIS-III was administered. Thus, a relatively direct comparison can be made between the O'Hora et al.'s (2008) study and Study 2 of the current research programme (since both involved presenting before/after relations followed by full administration of the WAIS-III). As highlighted in the previous section, in both studies significant correlations were obtained between before-after relational responding and the Block Design subtest. In contrast, in O'Hora et al.'s (2008) study responding on the before/after task predicted performances on three of the Verbal Comprehension subtests (Vocabulary, Similarities and Information), but no such relationships with before/after were observed in Study 2 of the current research. Similarly, the results of Study 4 revealed that the hierarchical IRAP also produced no significant correlations with the verbal subtests, and the bigger/smaller IRAP only correlated significantly with one of the verbal subtests (i.e., Information). These findings may suggest that the IRAP, with its requirement for fast and flexible responding, places a greater demand on fluid (associated with performance) rather than crystallised (associated with verbal) aspects of intelligence. Indeed, this suggestion is supported by the significant correlations that were obtained between the relational tasks and the AH4, which as highlighted above, is considered a fluid measure of cognitive ability (Parkin & Java, 1999).

Although the relational tasks (the IRAP and the REP) showed correlations with Block Design and Matrix Reasoning, fewer significant correlations were observed for the two other performance-based subtests (Picture Completion and Picture Arrangement). The Picture Completion subtest involves identifying a missing aspect of a picture, and might not be considered highly relational. Therefore, the lack of significant correlations with this subtest is perhaps unsurprising. In contrast, the Picture Arrangement subtest involves rearranging a series of scrambled pictures into a logical sequence, and thus performances on this subtest would appear to be heavily reliant on the ability to apply temporal (before/after) relations (i.e., the picture cards are arranged such that cards depicting an event that logically happens before or after another event must be placed in the correct order). Despite this, however, no significant correlations were observed for Picture Arrangement in either the before/after task presented in Study 2, nor that presented in O'Hora's (2005, 2008) research. However, it is important to note that concerns have been muted in relation to the reliability and validity of the Picture Arrangement subtest. For example, Lichtenberger and Kaufman (2009) noted that "despite much systematic analysis, the final system for assigning credit for the completion of items appears to be more or less arbitrary" (pg. 27), and it is for this reason that the Picture Arrangement subtest does not load onto the Perceptual Organisation index. Indeed, this is also the reason that Picture Arrangement has been dropped from the recently released edition of the WAIS (i.e., the WAIS-IV; Wechsler, 2008). Thus, the lack of correlation between the IRAP and REP with this subtest could be seen as a strength rather than a weakness, because it was subsequently removed from the WAIS.

Like the Perceptual Organisation index, some of the Verbal Comprehension subtests failed to correlate significantly with the IRAP measures. Specifically, across the IRAP tasks, no significant correlations were observed for the Similarities or the Comprehension subtests. However, as mentioned previously, the IRAP seems particularly sensitive to fluid intelligence, and thus these non-significant results may be unsurprising. Nevertheless, a number of conceptual issues are raised by the patterns of the correlations that were observed across the various IRAP tasks. It is to these conceptual issues that we now turn.

Conceptual Issues

The findings obtained from the current research programme suggest that the IRAP is a useful measure for investigating relational responding. Specifically, the IRAP provides three chronometric measures and therefore compliments and advances previous methodologies, which only provided a measure of response accuracy (e.g., the REP). In addition, the difference-score represents an important advance in that it facilitates an analysis of relational or cognitive flexibility. In so doing, it opens up a complex area for investigation, and one which was largely ignored by behaviour analysts.

The current research programme highlights some issues related to using the IRAP, which might be borne in mind by future researchers. One limitation of the current work is that it is not clear to what extent the observed correlations indicated relationships between generic patterns of relational framing and intelligence, or between particular features of the IRAP task, including the stimuli employed, and the IQ measures. Consider, for instance, the significant correlations that were observed between the hierarchical IRAP and the Matrix Reasoning subtest, highlighted previously. The observed correlations may indicate that the hierarchical relational frame provides the functional basis for performance on matrix reasoning tasks. Alternatively, however, the correlations may be reflective of how the hierarchical IRAP protocols were designed. As reported above, the hierarchical IRAP involved employing two sample word pairs (i.e. *Category-Member*) and a series of target word-pairs (e.g., *Vehicle- Car*), and thus the task required participants to relate the target relation to the sample relation. Perhaps it was the requirement to relate relations, rather than hierarchical responding *per se,* that was responsible for the significant correlations with Matrix Reasoning.

A related point is that the IRAP's measures of speed and flexibility appeared to place greater demand on fluid rather than crystallised intelligence. In contrast, the before/after REP employed by O'Hora, et al. (2005, 2008), which is a non-timed accuracy measure, appeared to tap both fluid and crystallised intelligence (i.e., the latter loaded on to both performance and verbal abilities). Once again, therefore, the nature of the relational task appears to influence the types of intelligence that are tapped into. Recall, also, that even changes within an IRAP task may impact upon the correlations observed. Specifically, in Study 4 significant correlations were observed between the IRAP measures and the Working Memory subtests, but significant correlations with working memory were not observed in Study 2. In Study 4, the target stimuli were altered such that each target word-pair was presented again in a second word-pair, and this appeared to place a heavier demand on working memory. The forgoing discussion highlights the procedural difficulties in this type of research, and the difficulties inherent in designing a task that only taps into the variable(s) of interest (in this case relational responding and relational flexibility).

Despite the inherent methodological difficulties in research of this nature, the findings that emerged are important. Indeed, many of the correlations that were obtained appear to be robust. Specifically, correlations between similar/different relational responding and vocabulary subtests were observed by O’Hora et al. (2005) and across Studies 1 and 2 of the current programme. Similarly, correlations between before/after relational responding and Block Design were observed both by O’Hora et al. (2008) and in Study 2. Given that two different methodologies were used across the studies (i.e., the IRAP in the current research, and the REP in O’Hora et al.’s research), it can be concluded that these correlations were not particularly task-specific. Rather, the correlations appear to be indicative of functional relationships between generic patterns of relational responding (i.e., relational frames) and more complex aspects of human cognition. In other words, the findings suggest that specific types of relational framing are related to particular aspects of human intelligence. This conclusion may have important implications in applied areas, and some relevant issues will be considered in the next section.

In highlighting possible relationships between specific relational frames and traditional measures of intelligence, it is important to note that the purpose here is to build a “bottom-up” account of intellectual functioning. As highlighted in the Introduction, Relational Frame Theory differs from cognitive theories in that it attempts to provide a “simple-to-complex” account of human cognitive abilities. This approach involves attempting to specify the core behavioural processes involved in human cognition, and how these processes are involved in complex human behaviours. Thus, identifying correlations between specific frames and sub-scales of particular IQ tests constitutes only the first step towards constructing a more fully developed functional account. The involvement of specific functional units, such as single relational frames, in defining and analysing human intelligence will surely provide only one part of a complex picture. According to RFT, human cognition involves a great deal more than basic relational framing activity (Hayes et al., 2001). Relating relations, complex relational networks, relating relational networks to other relational networks, and subtle and flexible forms of contextual control over such complex relating are seen as central to human cognition and to what we typically mean by intelligence. Thus, future research in this area will need to target other types of relational activity beyond relational frames per se. Nevertheless, the current work provides a useful starting point.

Broader Implications

The findings of the current research programme suggest a number of implications that may be relevant to the applied educational arena. First, the finding that relational flexibility (as measured by the difference-score) produced a range of significant correlations with IQ is important. It has been argued that flexibility in relational responding, particularly as it comes under increasingly subtle forms of contextual control, may underpin abilities such as creativity and problem-solving, and intelligence more generally (Hayes, 1994; Barnes, Hagerty, & Smeets, 1997; Healy, Barnes-Holmes, & Smeets, 2000). The current findings lend support to this argument. Furthermore, within mainstream cognitive psychology the ability to adopt flexible response strategies has long been regarded as an important feature of human intelligence (e.g., Cattell, 1971; Kyllonen, Lohman, & Woltz, 1984).

If flexibility is indeed a core feature of human intelligence, then it follows that rigidity, the antithesis of flexibility, is likely to be detrimental to intelligence. Indeed there are numerous studies that demonstrate the detrimental effects of cognitive rigidity across many different populations (e.g. Autism, Turner 1999; Attention Deficit/Hyperactivity Disorder, Lovecky, 2004; Schizophrenia, Pishkin & Williams, 1997; and normal adults, Wulfert, Greenway, Frakas, Hayes & Dougher, 1994).

The issue of rigidity may be particularly problematic in applied settings considering that behavioural interventions have predominantly been used in the treatment of those with autistic spectrum disorders. A tendency toward rigid, stereotypic behaviour is of course a defining feature of autism (DSM-IV-TR, 2000). Furthermore, behaviour-analytic research has demonstrated that sustained reinforcement for particular types of responding can result in less variable repertoires (e.g., Vogel & Annau, 1973). Recall also the results of Studies 1 and 2, which demonstrated that participants found reversing a highly practiced relational response (similar/different) more difficult than reversing a less rehearsed response (before/after). Thus, the current work and previous studies lend support to the idea that promoting relational flexibility may be crucial in educational settings. Indeed, this conclusion is consistent with RFT (e.g., Barnes, Hegarty, & Smeets, 1997), and preliminary findings in applied educational settings. For instance, O’Connor (2004) found that introducing an intervention program designed to promote flexible response repertoires facilitated the acquisition of new skills in children with autism.

One of the findings from the current research programme was that particular types of relational frames predict performance on certain types of cognitive tasks. Specifically, similar/different responding predicted performance on vocabulary tasks, whereas temporal relations predicted performance on the Block Design task. These results suggest the possibility of designing individually tailored intervention programmes to target specific cognitive deficits. Consider for instance, a child who is assessed for cognitive functioning on a standardised intelligence test, and the results reveal particular deficits on the Vocabulary subtests relative to the other subtests. A typical intervention may involve attempting to improve the child’s vocabulary through reading and conversations (e.g., Curtis & Longo, 2001). However, this necessitates training a vast array of content, and the intervention is therefore likely to be quite protracted. The current results suggest there may be utility in targeting the core acquisition of similar/different relational responding, since this is the relation that appears to underpin performance on vocabulary tasks. Similarly, deficits on the Block Design task might indicate the need for targeting before/after or sequential relations. By targeting the core relational processes in this way, it may be possible to tap into the particular areas of developmental need. In other words, training fluency and flexibility in the relevant relational frames may facilitate more efficient acquisition of broader complex skills.

The foregoing discussion highlights the implications of the current work for remedial education. However, the findings may also have relevance for educational practice in general. If flexibility is a core and defining feature of intelligence, then it would make sense to target flexibility in any learning environment. However, flexible thinking skills are often neglected in mainstream educational curricula. In the Irish education system for instance, a dominant preoccupation amongst second-level students is the Leaving Certificate examination, which functions as a competition for entry to third level courses. There is widespread concern however, that the system is deeply flawed in that it rewards rote learning and the memorisation of content, and discourages innovation, critical thinking and creativity (Conway, 2009; McDowell, 2007). Indeed, many commentators have raised concerns about the extent to which school-leavers appear ill-equipped to tackle the demands of third-level education or to solve work-place and even personal problems (e.g., Collins, 2010). While the current research programme was not designed to inform mainstream educational practice, the findings appear to be relevant, if indirectly, to these important issues.

Concluding Remarks

The current research programme has provided some of the first steps necessary for constructing a functional, bottom-up account of human intelligence. In so doing, it has facilitated a greater understanding and greater specification of the processes that underlie complex cognitive performances. The current research also highlighted the utility and the sensitivity of the IRAP as a methodology for investigating relational responding. Although the current research was not designed specifically to inform educational practice, the results that emerged highlighted the potential utility of targeting specific relational frames and relational flexibility across a range of educational contexts. Further research might follow up on this valuable research endeavor.

REFERNCES

REFERENCES

Abdi, H (2007). Bonferroni and Šidák corrections for multiple comparisons. In N. J. Salkind (Ed.). *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage.

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

Barnes-Holmes, D., Barnes-Holmes, Y., & Cullinan, V. (2000). Relational frame theory and Skinner's Verbal Behavior: A possible synthesis. *The Behavior Analyst, 23*, 69-84.

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

Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record, 41*, 19-31.

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, D., Hayes, S. C., Dymond, S., & O'Hora, D. (2001). Multiple stimulus relations and the transformation of stimulus functions. In S. C. Hayes, D., Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 51-71).

New York: Plenum.

Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit Relational Assessment Procedure (IRAP): Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white Irish individuals. *The Psychological Record, 60,* 57-66.

Barnes-Holmes, D., Staunton, C., Barnes-Holmes, Y., Whelan, R., Stewart, I., Commins, S., Walsh, D., Smeets, P. M., & Dymond. (2004). Interfacing relational frame theory with cognitive neuroscience: Semantic priming, the implicit association test, and event related potentials. *International Journal of Psychology and Psychological Therapy, 4*, 215-240.

Barnes-Holmes, Y., & Barnes-Holmes, D., & Cullinan, V. (2001). Education. In S. C. Hayes, D., Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 181-195). New York: Plenum.

Biglan, A., & Hayes, S. C. (1996). Should the behavioral sciences become more pragmatic? The case for functional contextualism in research on human behavior. *Applied and Preventive Psychology: Current Scientific Perspectives, 5*, 47-57.

Broadbent, D. E., Cooper, P. F., Fitzgerald, P., & Parkes, K. R. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology, 21,* 1-16.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*.

Cambridge: Cambidge University Press.

Cattell, R. B. (1963). *Theory of fluid and crysallised intelligence: A critical experiment*. Journal of Educational Psychology, 54, 1-22.

Cattel, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.

Cattell, R. B. (1987) *Intelligence: Its structure growth, and action*. New York: North-Holland.

Ceci, S. J. (1996). *On intelligence: A bioecological treatise on intellectual development*. Cambridge MA: Harvard University Press.

Chiesa, M. (1994). *Radical behaviorism*. Boston, MA: Authors Cooperative.

Conway, A. R. A., Cowen, N., Bunting, M. F., Therriault, D. J. & minkoff, S. R. B. (2002). A latent variable analysis of workin memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30,* 163-183.

Collins, T. (2010, August 16). We're missing the point on how to prepare for college. *The Irish Independent.* Retrieved September 16, 2010, from http://www.independent.ie/opinion/analysis/tom-collins-were-missing-the-point-on-how-to-prepare-for-college-2298580.html

Conway, P. F., Murphy, R., Rath, A., Hall, K. (2009). *Learning to teach and its implications for the continuum of teacher education: a nine-country cross national study*. Dublin: The teaching Council

Cullinan, V., Barnes, D., & Smeets, P.M. (1998). A precursor to the relational evaluation procedures: Analyzing stimulus equivalence. *The Psychological*

Record, 48, 121-145.

Curtis, M. E. & Longo, A.(2001) Teaching Vocabulary to Adolescents to Improve Comprehension, Reading Online: *An electronic journal of the Reading Association*. Retrieved on September 16, 2010, from http://www.readingonline.org/articles/art_index.asp?HREF=/articles/curtis/index.html

Davidson, J. E., & Downing C. I. (2000). Contemporary models of intelligence. In Sternberg, R. J. (Ed.), *Handbook of intelligence* (pp. 34-49). Cambridge: Cambridge University Press.

Deary, I. J. (1997). Intelligence and information processing. In H. Nyborg (Ed.), *The scientific study of human nature: Tribute to Hans Eysenck at eighty* (pp282-310). New York: Pergamon.

Deary, I. J. (2000). Simple information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 267-284). Cambridge: Cambridge University Press.

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

Dougher, M. J., Hamilton, D., Fink, B., & Harrington, J. (2003). *Generalized relational stimuli and the transformation of operant and respondent stimulus functions.* Paper presented at the annual convention of The Association for Behavior Analysis, San Francisco, USA.

DSM-IV-TR (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision.* Washington DC: American Psychiatric

Association

Dymond, S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more-than, and less-than. *Journal of the Experimental Analysis of Behavior, 64*, 163-184.

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

Gardner, H. (1993). *Frames of mind: A theory of multiple intelligences* (10th anniversary edition) New York: Basic Books.

Gardner, H. (1998). Are there additional intelligences? The case for naturalist, spiritual, and existential intelligences. In J. Kane (ed.), *Education, information and transformation* (pp.111-132). Englewood Cliffs, NJ: Prentice Hall.

Gathercole, S.E., Alloway, T.P., Willis, C., & Adams, A.M. (2006). *Working memory in children with reading disabilities . Journal of Experimental Child Psychology, 93*, 265-281.

Gentner, D. (2003). Why we’re so smart. In D. Gentner and S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp.195-235). Cambridge, MA: MIT Press.

Gentner, D., & Loewenstein, J. (2002). Learning: analogical reasoning. *Encyclopedia of Education, Second Edition.* New York: Macmillian.

Gifford, E. V., & Hayes, S. C. (1999). Functional contextualism: A pragmatic philosophy for behavioral science. In W. O'Donohue & R. Kitchener (Eds.), *Handbook of behaviorism* (pp. 285-327). San Diego: Academic Press.

Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. *Child Development, 62,* 1-22.

Gould, S. J. (1981). *The Mismeasure of Man*. New York: W.W. Norton & Co

Hair, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlet, E., Wu, J. C., Paek, J., Browning, H. L. & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence, 12*, 199-217.

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science, 16*, 70-76.

Hayes, S. C. (1991). A Relational Control Theory of Stimulus Equivalence. In L. J., Hayes, & P. N. Chase, (Eds.), (pp. 19-40). *Dialogues on verbal behavior*. Reno, NV: Context Press.

Hayes, S. C. (1993). Analytic goals and the varieties of scientific contextualism. In S. C. Hayes, L. J. Hayes, H. W. Reese, & T. R. Sarbin (Eds.), *Varieties of scientific contextualism* (pp. 11-27). Reno, NV: Context Press.

Hayes, S. C. (1994). Relational frame theory: A functional approach to verbal events. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 9–30). Reno, NV: Context Press.

Hayes, S. C., & Barnes, D. (1997). Analyzing derived stimulus relations requires more than the concept of stimulus class. *Journal of the Experimental Analysis of Behavior, 68*, 225-233.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition.* New York:

Plenum.

Hayes, S. C., & Berens, N. M. (2004). Why Relational Frame Theory alters the relationship between basic and applied behavioral psychology. *International Journal of Psychology and Psychological Psychotherapy, 4*, 341-353.

Hayes, S. C., Blackledge, J. T., & Barnes-Holmes, D. (2001). Language and cognition: Constructing an alternative approach within the behavioral tradition. In S. C. Hayes, D., Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 3-20). New York: Plenum.

Hayes, S. C., Fox, E., Gifford, E., Wilson, K., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 21-49). New York: Plenum.

Hayes, S. C., Gillford, E. V., Townsend, R. C. & Barnes-Holmes, D. (2001). Thinking, problem-solving, and pragmatic verbal analysis. In S. C. Hayes, Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 87-102). New York:

Plenum.

Hayes, S.C., Hayes, L. J., & Reese, H. W. (1988). Finding the philosophical core: A review of Stephen Pepper’s world hypothesis: A study in evidence.

Journal of the experimental analysis of behavior, 50, 97-111.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An experimental approach to behavior change.* New York: Guilford Press.

Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior, 74*, 207-227.

Heim, A. W. (1970). AH4: *A group test of general intelligence*. Berkshire: NFER-NELSON.

Hendrickson, A. E. (1982). The biological basis of intelligence. Part I: Theory. In H. J. Eysenck (Ed.), *A model for intelligence* (pp151-196). New York: Springer Verlag.

Hernstein, R., J. & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57,* 253-270.

Jameson, J., and Gentner, D. (2003). *Mundane comparisons can facilitate relational understanding.* Proceedings of the Twenty-fifth Annual Meeting of the Cognitive Science Society. Retrieved September 16, 2010 from http://groups.psych.northwestern.edu/gentner/papers/JamesonGentner03.pdf

Jensen, A. (1987). *Consensus and controversy*. Philadelphia PA: Routledge

Kaufman, A. S., & Kaufman, N. L. (1990) *The Kaufman Brief Intelligence Test*. Minnesota: American Guidance Service.

Kilgour, A. H. M., Starr, J. M., & Whalley, L. J. (2010). Associations between childhood intelligence (IQ), adult morbidity and mortality. *Maturitas*, *65,* 298-105

Kuhn, T.S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kyllonen, P. C. (1996). Is working memory capacity Spearman's *g*? In I. Dennis& P. Tapsfield (eds.), *Human abilities: their nature and measurement* (pp. 49-75). Mahwah, NJ: Erlbaum

Kyllonen, P.C., Lohman, D. F., & Woltz, D. J. (1984). Componential modeling of alternative strategies for performing spatial tasks. *Journal of Educational Psychology, 76,* 1325-1345.

Lichtenberger, E., O., & Kaufman, A. (2009) *Essentials of the WAIS-IV Assessment.* John Wiley & Son: New Jersey

Lohman, D. F., Thorndike, R. L., & Hagan, E. P., 2001. *Cognitive Abilities Test 3*. London: GL Assessment

Lovecky, D., V. (2004). *Different minds: Gifted children with AD/HD, Aspergers syndrome, and other learning deficits.* London: Jessica Kingsley Publishers.

MacCorquodale, K. (1969). B. F. Skinner's Verbal Behavior: A retrospective appreciation. *Journal of the Experimental Analysis of Behavior, 12,* 831-841.

McDowell, M. (2007, August 19) Failures of the Leaving Cert. *The Sunday*

Business Post. Retrieved on September 16, 2010, from http://archives.tcm.ie/businesspost/2007/08/19/story25937.asp

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf–Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151−179). New York: Guilford.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research, *Intelligence, 37*, 1-10.

McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): Gf–Gc cross-battery assessment.* Boston: Allyn & Bacon.

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D. (2004). A relational frame account of the development of complex cognitive phenomena: Perspective-taking, false belief understanding, and deception. International Journal of Psychology and *Psychological Therapy, 4*, 303-323.

McKenna, I. M., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology and Psychological Therapy, 7*, 253-268.

Merbitz, C., Vieitez, D., Merbitz, N. H., & Pennypacker, H. S. (2001). Precision Teaching: Foundations and classroom applications. In D. J. Moran, & R. W. Mallott (Eds.), *Evidence-based educational methods*, (pp. 47-62). San Diego, CA: Elsevier.

Morris, E. K. (1989). Questioning psychology’s mechanism: A review of Costall and Still’s “Cognitive Psychology in Question”. *The Behavior Analyst, 12,* 59-67.

O’Connor, J. (2004). *Designing procedures to establish repertoires of derived relational responding.* Unpublished doctoral thesis, National University of Ireland, Maynooth, Maynooth, Co. Kildare, Ireland.

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

O’Hora, D., Pel◌́◌́aez, M., Barnes-Holmes, D., Rae, G., Robinson, T., & Chaudhary, T. (2008). Temporal relations and intelligence: Correlating relational performance with performance on the WAIS-III. *The Psychological Record, 58,* 569-584.

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

Parkin, A. J., & Java, R. I. (1999). Deterioration of frontal lobe function in normal aging: Influences of fluid intelligence versus perceptual speed. *Neuropsychology, 13*, 539-545.

Pickering, S. J. (Ed.). (2006). *Working memory and education*. San Diego, CA:

Academic Press.

Pinker, S. (2002). *The Blank slate: The modern denial of human nature*. New York: Penguin.

Pinker, S. (2004). *The language instinct*. NewYork: Harper Collins

Piskin, V., & Williams, W. V. (1977). Cognitive rigidity in information processing of chronic undifferentiated schizophrenics. *Journal of Clinical Psychology, 33*, 625-630.

Premack, D. (1994). Is language the key to human intelligence? *Science, 303,* 318-320.

Reece, H. W. (1968). *The Perception of Stimulus Relations: Discrimination Learning and Transposition.* Academic Press: New York.

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

Sidman, M. (1994). *Stimulus equivalence: A research story*. Boston: Authors Cooperative.

Sidman, M. & Tailby, W. (1982). Conditional discrimination versus matching-to-sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37,* 5-22.

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

Spearman, C. (1904). “General intelligence” objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Steele, D., & Hayes S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, *56*, 519–555.

Stewart, I., Barnes-Holmes, D., Hayes, S. C., & Lipkens, R. (2001). Relations among relations: Analogies, metaphors, and stories. In S. C. Hayes, Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 73-86). New York: Plenum.

Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: Chicago University Press.

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

Vernon, P. E., & Mori, M. (1992). Intelligence, reaction times, and peripheral nerve conduction velocity. *Intelligence, 16*, 273-288.

Vogel, R., & Annau, Z. (1973). An operant discrimination task allowing variability of reinforced response patterning. *Journal of the Experimental Analysis of Behavior, 20*, 1–6.

Wahlsten, D. (1997). The malleability of intelligence is not constrained by heritability. In B. Devlin, S. E. Frienberg, D. P. Resnick, & Roeder, K. (Eds.), *Intelligence, genes and success: Scientists respond to The Bell Curve* (pp. 71-88). New York: Springer-Verlag.

Wechsler, D. (1981) *The Wechsler Adult Intelligence Scale – Revised.* San Antonio: The Psychological Corporation.

Wechsler, D, (1997) *The Wechsler Adult Intelligent Scale – Third Edition*. San Antonio: The Psychological Corporation.

Wechsler, D, (1997) *The Wechsler Adult Intelligent Scale – Third Edition –UK Version.* London: The Psychological Corporation.

Wechsler, D. (2008) *The Wechsler Adult Intelligent Scale – Fourth Edition*. San Antonio: Pearson.

Wulfert, E., Greenway, D. E., Farkas, P., Hayes, S. C., & Dougher, M.. J. (1994). Correlation between self-reported rigidity and rule-governed insensitivity to operant contingencies. *Journal of Applied Behavior Analysis*, *27*, 659–671.

Zimmerman, D.W., &Williams, R. H. (2000). Restriction of range and correlation outlier-prone distributions. *Applied Psychological Measurement*, *24*, 267–280.

APPENDIX

APPENDIX

Instructions Presented on Computer Screen to Participants Prior to Completing the IRAP Tasks

INTRODUCTORY INSTRUCTIONS

Our research investigates cognitive processes that are used in decisions that involve memory. We are seeking to develop and test theories of cognitive processes that occur inside and outside of awareness in the routine use of memory.

Stimuli will be presented on this display screen, and your responses will be entered on the keyboard.

The research assumes that you can read English fluently, and that your vision is normal or corrected to normal. If you do not consider yourself fluent in English, or if your vision is not normal or corrected to normal, and ESPECIALLY IF YOU ARE HAVING SOME DIFFICULTY READING THIS DESCRIPTION, PLEASE ask the experimenter now whether or not you should continue.

Your identity as a subject is confidential. Further, you are free to discontinue participation at any time, without penalty.

In keeping with standard practice, your data may be retained for 5 years or so, during which time only the investigators on this or successor projects will have access to them.

PLEASE NOW READ THE STATEMENT BELOW, WHERE YOU WILL BE ASKED TO RESPOND TO A STANDARD INFORMED CONSENT QUESTION.

CONSENT STATEMENT

I have read the description of the procedure. I understand that the questions I may have about this research will be answered by Professor Barnes-Holmes or one of the other researchers working on this project.

If you consent to participate in the research that has been described on the preceding display pages you should now read the Instructions for the sorting tasks below.

[INSTRUCTION: If you wish to ask any questions first, alert the experimenter now. IF YOU WISH NOT TO PROCEED, you should inform the experimenter].

INSTRUCTIONS FOR THE SORTING TASKS

Shown below are illustrations of the four different types of task that will be presented repeatedly in this part of the experiment. To help you understand the tasks each of the four illustrations is explained immediately underneath. Please examine each illustration and then read carefully the explanation attached to it. Please make sure that you understand each task before continuing with the experiment.

IMPORTANT: From trial to trial the positioning of the response options (True and False) will vary randomly between left and right.

Illustration 1

Before

Spring Summer

Select ‘d’ for **True**	Select ‘k’ for **False**

Explanation for Illustration 1

If you select “True” by pressing the ‘D’ key, you are stating that “Spring Comes Before Summer.”

If you select “False” by pressing the ‘K’ key, you are stating that “Spring Does **Not** Come Before Summer.”

Illustration 2

After

Spring   Summer

Select ‘d’ for **True**	Select ‘k’ for **False**

<u>Explanation for Illustration 2</u>

If you select “True” by pressing the ‘D’ key, you are stating that “Spring Comes After Summer.”

If you select “False” by pressing the ‘K’ key, you are stating that “Spring Does **<u>Not</u>** Come After Summer.”

Illustration 3

Before

Summer   Spring

Select ‘d’ for **True**	Select ‘k’ for **False**

<u>Explanation for Illustration 3</u>

If you select “True” by pressing the ‘D’ key, you are stating that “Summer Comes Before Spring.”

If you select “False” by pressing the ‘K’ key, you are stating that “Summer Does **<u>Not</u>** Come Before Spring.”

Illustration 4

After

Summer Spring

Select ‘d’ for **True**	Select ‘k’ for **False**

<u>Explanation for Illustration 4</u>

If you select “True” by pressing the ‘D’ key, you are stating that “Summer Comes After Spring.”

If you select “False” by pressing the ‘K’ key, you are stating that “Summer Does **<u>Not</u>** Come After Spring.”

NOTE: During the experiment a range of other word pairs apart from “Spring” and “Summer” will also be presented.

REMEMBER: From trial to trial the positioning of the response options (True and False) will vary randomly between left and right.

<u>FINAL INSTRUCTIONS</u>

During the experiment you will be asked to respond as **quickly <u>and</u> accurately** as you can across all trials.

It is very important to understand that sometimes you will be required to respond to the tasks in a way that ***<u>agrees</u>*** with what you believe and at other times you will be required to respond in a way that ***<u>disagrees</u>*** with what you believe. **<u>This is part of the experiment</u>**.

When you make an incorrect response for a task it is signalled by the appearance of a red ‘X’ in the centre of the screen. To remove the red ‘X’ and continue please make the correct response quickly.

If you do not understand something about the foregoing instructions or have any further questions please talk to the researcher before clicking on the red button.