Exploring the Impact of Specific, Minimal, and Response-Focused Instructions on the Implicit Relational Assessment Procedure

Thesis submitted to the Department of Psychology, Faculty of Science, in fulfilment of the requirements for the degree of Masters of Science, National University of Ireland, Maynooth.

Martin Finn, B.A. (Hons.)

September 2015

Research Supervisor: Prof. Dermot Barnes-Holmes

Head of Department: Dr. Andrew Coogan

Table of Contents

Acknowledgements

I owe this thesis to many people I am fortunate to know.

Dermot, I would like to thank you for this opportunity, working with you has been both a complete pleasure, and privilege. Your guidance and your support throughout the process have been invaluable.

Thank you, Yvonne, for opening my eye and setting my sights.

Thank you, Ian, Ciara, and Aoife, for being there with help, advice, and support.

Aileen, Colin, and Deirdre, you have made this year thoroughly enjoyable, and made it bearable when it was difficult. I am glad I was able to share this with you.

I would like to thank both of my parents for their enduring care and love. This would have been impossible without you.

Thanks also to my friends for the comic relief and the well spent evenings.

Figures and Tables

Appendices

- Appendix A: Informed consent form
- Appendix B: Demographics and Rating scale

Abstract

The Implicit Relational Assessment Procedure (IRAP) is increasingly used in applied and clinical settings, and yet many of the procedural variables of the IRAP have not been subjected to a systematic analysis. One such variable is the type of rules that are employed when instructing the IRAP and the effect this might have on resultant performance and obtained from the procedure. In the current thesis, three experiments assess the impact of three different types of rules or instructions on the IRAP. The instructions employed varied in the degree to which they specified parts of the relational network being assessed by the IRAP. The findings from these three studies show that the type of rule that is presented to participants during an IRAP may have a dramatic effect on the strength and direction of the trial-type effects that are produced by the measure. Furthermore, the type of instructions employed appear to interact with the order in which the IRAP blocks are presented (history-consistent versus history-inconsistent). The findings arising from the current thesis indicate that the behavioural dynamics that occur when participants complete an IRAP require extensive and systematic experimental and conceptual analyses, and this work will have an important bearing on research seeking to investigate the predictive validity of the IRAP in applied and clinical settings.

Introduction

History of Behaviour Analysis; Skinner, Sidman and Relational Frame Theory

Behaviour Analysis. Behavior analysis is a comprehensive scientific approach to the study of behaviour with its own set of philosophical assumptions (see Chiesa, 1994). Behavior analysis seeks to provide a monistic account of behaviour in which behaviour is seen to be controlled by the environment. The behavior-analyst does not appeal to mental states or processes in attempting to explain behavioural change. Behaviour analysis also has a strong tradition of functional-contextual assumptions, which are sometimes contrasted with more mechanistic thinking. For example, in a functional approach the mechanistic conception of spatio-temporally contiguous cause-and-effect is replaced with that of functional relations between events across both time and space. This functional approach also replaces the concept of causation with that of description. The behavior analyst describes the functional relations between environmentally situated events or contingencies, and the behavior of interest, and the impact of these contingencies on that behaviour. In short, the behaviour of interest is seen as a dependent variable and alters as a function of changes in the environment, the independent variable (Skinner, 1950).

Within this framework, behaviour analysis adopts a pragmatic approach where the goal of the behavior analyst is to achieve prediction-and-influence over the behaviour of interest by describing and manipulating the environment controlling it. The subject matter of the behavior analyst is the response class, defined as instances of behavior that have equivalent impact upon the environment (Hineline, 1989). One of the key causal modes of behaviour-analysis is that of selection by consequence, in which the consequences of a response impact on the future probability of that response (Chiesa, 1992). The response class is scrutinised both by observation and manipulation of the environmental variables that are functionally related to the class. Arising from the basic philosophical assumptions and goals of the field, the main unit of analysis found within behaviour analysis is the operant, composed of functional relations between antecedent stimuli, responses by the organism, and the consequences of the responses. This approach of describing environmental variables that impact on the response class through inferred functional relations can be described as a direct contingency approach (direct in the sense that the contingencies impact the behavior directly through the inferred functional relations).

Skinner's Verbal Behaviour and Instructional Control. The operant proved to be a useful unit of analysis for investigations into the behavior of non-human animals (Skinner, 1937). The account of behavior in terms of functional relations as proposed by Skinner was intended to be applicable to behavior generally, whether that behavior was emitted by a human or an animal. However, the extension of Skinner's analysis of behavior to human language in "Verbal Behavior" (1957) and the subsequent dearth of basic research programs arising from it, suggested to some behaviour analysts that not all human behavior could be accounted for by the analysis of direct contingencies (Hayes, Barnes-Holmes, and Roche, 2001).

One key issue with the analysis of direct contingencies as it applied to the functional analysis of human behavior was that there were differences in the behaviour generated by humans and by animals when they were placed on similar schedules of reinforcement (Weiner, 1969; Bentall, Lowe, & Beasty, 1985; Lowe, Beasty, & Bentall, 1983). The argument was made that the human capability for language accounted for these species differences (Lowe, 1979). The basic argument was that a pre-current behaviour (i.e. typically conceptualized as verbal) impacted upon responding on the schedule (Catania, Matthews, & Shimoff, 1989). The presence of a pre-current behaviour occurring between antecedent stimuli and the overt response, thus accounted for the observed so-called *insensitivity* of human behaviour to the direct contingencies arranged by schedules of reinforcement. Often, the insensitivity effect observed with human schedule performance was attributed to the impact of *verbal rules* that were generated by human participants as they interacted with the scheduled contingencies (Vaughen, 1989). Insofar as non-humans did not possess the capacity for generating such rules, their behaviour was seen as being directly controlled by, or entirely sensitive, to the reinforcement schedule. It is worth noting that Skinner (1966) himself introduced the concept of instructions or rules to the behaviour-analytic tradition in a seminal paper on human problem solving, and thus the focus on rule-governed behaviour in the context of human schedule performance was not considered particularly problematic.

On balance, the introduction of the concept of rule-governed behaviour, as a means of explaining human insensitivity to reinforcement schedules, served to highlight a problem with a core assumption in behaviour-analytic thinking. Specifically, appeals to direct contingencies alone could not effectively account for human behaviour in general. It could be argued that one response to this conclusion was to embrace the more mechanistic tradition of cognitive psychology that had emerged and come to dominate the field of psychology during the 1960s and beyond. Many behaviour analysts refused to do so, however, and attempted to continue to develop monistic, functional and pragmatic accounts of human behaviour that addressed the issue of instructions, rules, and the so called insensitivity effect. One excellent example of this attempt was articulated by Hineline (1989) who argued that appealing to verbal behaviour and rules as the basis for human insensitivity could involve the complex interplay among multiple operant classes "... if verbal behavior is functionally involved, the behavior-analytic account distinguishes between the behavior of stating rules, the behavior of following rules, and the behavior described by rules, which are distinct operant classes" (p 253). The basic idea was that the interactions among these distinct operant classes in terms of behavior-behavior relations may be seen as altering the sensitivity to the contingencies operating in the environment. However, the account remained behaviour-analytic because critically the behaviors of rule stating, and rule following, are ultimately controlled by contingencies operating in the environment (Barnes, 1989). Thus, an analysis of such environmental control required an account of rule-governed behaviour in operant terms.

The focus on rule-stating and rule-following as separate operant classes allowed behavioural researchers interested in the study of human operant behaviour to develop programmes of research that extended beyond the so-called schedule-insensitivity effect. For example, studies were conducted that attempted to analyze the differential impact of contingency- versus performance-based descriptions on behaviour on schedules of reinforcement (e.g., Catania, Matthews, & Shimoff, 1989). While clearly useful in moving the field forward in terms of developing an independent basic experimental science of human behaviour in the radical behavioural tradition, a conceptual gap or hole in the work began to emerge. Specifically, some behavioural researchers were seeking to establish a clear functional definition of verbal rules and this was proving less than clear cut.

Skinner (1966) described rules as contingency specifying stimuli, in that they specified the occasion, the required response, and the expected consequence. The proposal seemed plausible. A rule must be an antecedent stimulus, and if it is to alter the sensitivity of an organism to the contingencies operating in the immediate environment it must specify the behavior that is to be engaged in, in place of the behavior that would otherwise occur. From a behavior-analytic perspective the specified behavior would be more probable if the reinforcement contingent on effective rule following was also specified. However, the kind of history that would allow for stimuli to function in such a way was not outlined. It was not made clear how a stimulus might specify another stimulus and as such this manner of speaking about rules, while making intuitive sense, was functionally inadequate. Behaviour analysis had not come up with a description of how a rule would come to have its function. The issue of specifying remained unspecified.

Stimulus Equivalence and Relational Frame Theory. The beginnings of an answer to the question of specifying stimuli within behaviour analysis began to emerge with the delineation of the phenomenon of stimulus equivalence (Sidman, 1971, Sidman & Tailby, 1982). Equivalence relations offered an explanation for the way in which behavior comes under the control of contingencies indirectly. Stimuli related to one another by equivalence relations became functionally equivalent, and so, behaviors occurring in the presence of one stimulus would now take place in the presence of the related stimulus, in spite of the fact that this stimulus had not until that point elicited those functions. By creating a description of how behaviors could be controlled by contingencies to which behaviors need not have been directly exposed, the behavior analyst had been granted a glimpse at a solution to the problem posed by the inadequacies of a direct contingency analysis of human behavior.

A more comprehensive solution to the above problem came with Hayes and Hayes' (1989) approach to rule-governed behavior, and their radical rethinking of the phenomenon of stimulus equivalence. This involved conceptualising stimulus equivalence as arbitrarily applicable relational responding (AARR). They conceived of such responding as involving responding to one stimulus in terms of another stimulus. The stimuli related, by means of AARRing, could be related to one another in an almost infinite number of ways, but the manner in which they were related in any given instance was seen as being controlled by contextual cues. According to this view, exposure to an extended history of relevant reinforced exemplars served to establish particular patterns of over-arching or generalized relational responses units, defined as relational frames. For example, a young child would likely be exposed to direct contingencies of reinforcement by the verbal community for pointing to the family dog upon hearing the word dog or the specific dog's name (e.g., "Rover"), and to emit other appropriate naming responses, such as saying "Rover" or "dog" when the family pet was observed, or saying "Rover" when asked, "What is the dog's name?" Across many such exemplars, involving other stimuli and contexts, eventually the operant class of coordinating stimuli in this way would become abstracted, such that direct reinforcement for all of the individual components of naming would no longer be required when a novel stimulus was encountered. Imagine, for example, that the child is shown a picture of an Aardvark, and the written word, and is told its name. Subsequently, the child may say "That's an Aardvark" when presented with a relevant picture or the word without any prompting or direct reinforcement for doing so. In other words, the relational response of coordinating pictorial, spoken stimuli and written words has been established and directly reinforcing a subset of the relating behaviours "spontaneously" generates the complete set.

Once this pattern of relational responding has been established, the generalized relational response could then be applied without regard to the formal characteristics of the stimuli, but only in the presence of appropriate contextual cues. Contextual cues were thus seen as functioning as discriminative for particular patterns of relational responding. The cues acquired their functions through the types of histories described above. Thus, for example, the phrase "that is a", as in "*That is a* dog" is established across exemplars as a contextual cue for the complete pattern of relational responding (e.g., coordinating the word dog with actual dogs). Once the relational functions of such contextual cues are established in the behavioural repertoire of a young child, the number of stimuli that may enter into such relational response classes becomes almost infinite (Hayes and Hayes, 1989; Hayes et al., 2001).

Hayes and Hayes' (1989) also provided a relatively precise technical definition of AARR or the unit of analysis, the relational frame. Specifically, relational framing possesses three properties; mutual entailment (if A is related to B then B is also related to A), combinatorial mutual entailment (if A is related B and B is related to C, then A is related to C, and C is related to A), and the transformation of functions (the functions of the related stimuli are changed or transformed based upon the types of relations into which those stimuli enter). Imagine, for example, that you are told that "Guff" is a really tasty new brand of beer, and that you will love it, but you are also told that another new brand is called "Geedy", and it is the complete opposite in terms of taste. It is likely that given a choice between the two beers you will chose the former over the latter, in part because the two verbal stimuli, Guff and Geedy have entered into a relational frame of opposition and the functions of Geedy have been transformed based on its relationship to Guff (in this case, you will expect Geedy to have an unpleasant taste). This general approach to human relational responding became known as Relational Frame Theory (RFT), and allowed for a behavior-analytic approach to key elements of language, such as meaning, reference, and understanding (Barnes and Holmes, 1991).

Following the initial exposition of the approach in the late eighties, the nineties saw a period of demonstration research that was designed to test its basic assumptions and core ideas. Some of this early research showed that AARRing as a process can be shown to occur in several distinct patterns. These patterns, referred to as relational frames (e.g. co-ordination, opposition, distinction, comparison, spatial frames, temporal frames, deictic relations, and hierarchical relations) have been demonstrated across a growing number of experimental studies, and some of the research has also reported reliable demonstrations of the property of transformation of functions (e.g. Steele and Hayes, 1991; Dymond and Barnes, 1995). Research has also shown that relational responding can be trained by a variety of procedures other than matching to sample, such as respondent type procedures (Leader, Barnes, and Smeets, 1997, Barnes, Smeets, and Leader, 1996), indicating that the phenomenon is not tied to a particular experimental preparations or modes of instruction, provided the key functional elements are present. There is also empirical evidence to support the argument that exposure to multiple-exemplars during early development is required to establish the specific relational frames (Lipkens, Hayes, & Hayes, 1993; Luciano, Becerra, & Valverde, 2007). As such, the argument that relational frames may be thought of as overarching or generalized relational operants has gained considerable traction. There is even evidence, although still quite limited, that non-humans, or at least sea lions, may show basic forms of relational framing given appropriate multiple-exemplar or operant training (Schusterman and Kastak, 1993).

Relational Networks and Rules from the Perspective of RFT. As an account of human language and cognition, RFT provided what is still considered today to be the basic operant unit involved in verbal behaviour – the relational frame. However, the seminal text on RFT (Hayes, et al. 2001) also used this unit to provide functional-analytic accounts of specific domains of human language and cognition, and rule-governed behaviour was one of these domains. According to RFT a rule or instruction may be considered a network of relational frames typically involving coordination and temporal relations with contextual cues that transforms specific behavioural functions. The simple instruction, "If the light is green then go" involves frames of coordination between the words "light", "green" and "go" and the actual events to which they refer. In this sense, the technical definition of the frame of coordination, outlined above, provides the functional-analytic definition of "specification" that was missing from earlier treatments of rules or instructions. In addition, the words "if" and "then" serve as contextual cues for establishing a temporal relation between the actual light and the act of actually going (i.e., first light then go). And the relational network as a whole serves to transform the functions of the light itself, such that it now controls the act of "going" whenever an individual who has been presented with the rule observes the light being switched on.

Although the foregoing example is a relatively simple one, the basic concept may be elaborated to provide a functional-analytic treatment of increasingly complex rules and instructions. Strictly speaking, however, the concept of the rule or instruction is not a technically precise one in RFT. Rather the concept of the relational network is the technical concept. Nevertheless, the concept of rules or instructions is so pervasive within behaviour analysis that these terms are often used interchangeably with the more technically precise but broader concept of the relational network. In the current thesis, I will adopt this convention.

The Dynamics of Relational Framing: The Need for New Procedures

As noted previously, much of the research in RFT may be conceptualized as demonstration research that was designed to test the theory's basic assumptions and core ideas. One of the defining features of this so-called demonstration research was a dichotomous approach to AARRing itself. In other words, basic laboratory studies in RFT often focused on showing that particular patterns of AARRing were either present or absent. Thus, for example, participants were required to produce perhaps 18 out 20 correct responses on a test for equivalence responding to demonstrate that the relational frame had emerged. In this sense, the relational frame was either present or absent in the participants behavioural repertoire. A critical feature of the concept of operant behaviour generally, however, is that it may vary in relative strength. Thus, for example, the simple operant of lever pressing for food pellets in rats may be conceived of being at relatively high or low strength. One way in which researchers have typically assessed such strength is by measuring how long it takes for the operant to extinguish when the reinforcement contingency (between lever pressing and food pellets) is terminated. In effect, the longer the extinction process takes, the stronger the operant response class was deemed to be.

The important point to note here is that basic RFT research on AARRing did not appear to have an immediately obvious way to assess relative strength using extinction procedures. One key problem is that AARRing by definition involves behaviour that emerges and may persist in the absence of direct reinforcement for particular responses because the contingencies are deemed to be extremely molar in nature. The generalized operants involved in many relational frames are deemed to have relatively long reinforcement histories behind them, going back to very early language learning. Using simple extinction procedures within the context of a 1-hour experimental session, for example, would not provide a realistic measure of the strength of such well-established operants. In addition, individual relational responses often form parts of larger relational networks, and thus attempting to extinguish individual responses may be unsuccessful because they are maintained based on their coherence with the larger network. Granted, some studies on AARRing examined the extent to which it was possible to reorganize patterns of relational responding that had been established within the laboratory (e.g., Healy, Barnes-Holmes, & Smeets, 2000; Pilgrim and Galizio, 1995), and thus could be seen as relevant to the question of relative strength of responding. However, this work also tended to focus on the dichotomous nature of relational frames in that it sought to establish new (re-organized) patterns that were either present or absent by the end of the training and testing procedures.

Within a few years of the publication of the 2001 RFT book, therefore, the need to develop procedures that could, in principle, provide a measure of relational responding that was non-dichotomous became increasingly apparent. Indeed, research in the applied arena, particularly in acceptance and commitment therapy, which was beginning to rely more and more heavily on concepts such as psychological flexibility, cried out for basic research techniques that could address the issue of relational flexibility. Flexibility, by definition, is a relativistic concept and thus questions around how flexible a particular pattern of AARRing might be required a method for assessing relational frames in a non-dichotomous manner. The initial response to this need or gap in technology was the development of what came to be known as the Implicit Relational Assessment Procedure (IRAP).

The IRAP as a Measure of Relational Responding "in Flight"

The initial inspiration for the development of the IRAP was the question, "How can we capture relational frames in flight", which essentially is a question about the relative strength of AARRing in the natural environment. In developing the IRAP two separate methodologies were combined. The first of these was an RFT-based procedure for training and testing multiple stimulus relations, the Relational Evaluation Procedure (REP) and the second was the Implicit Association Test (IAT). The latter had been developed by social-cognition researchers as a method for measuring what they conceptualize as associative strengths in memory (Greenwald, McGhee, & Schwartz, 1998). When the two measures were combined into the IRAP, however, it was conceptualized as a procedure for measuring the strength of natural verbal relations, or in other words AARRing (Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2008). Due to its close connection to the IAT, however, research with the IRAP quickly became dominated by studies focused on so called implicit attitudes and implicit cognition more generally. On the one hand, this strategy was very useful because it provided a means by which to assess the validity of the IRAP as a measure of natural verbal relations (Vahey, Nicholson, & Barnes-Holmes, 2015). On the other hand, it also served as a distraction from a focus on RFT and AARRing *per se*. This sentiment was recently articulated as follows:

> "Certainly, measures such as the IAT provide impressive predictions of socially and clinically relevant behavior. Nonetheless, it is critical to note here that our interest in implicit cognition came out of our interest in understanding the dynamics of arbitrarily applicable relational responding as it actually occurs. To this end, and inspired in part by the IAT, the second author sought to develop a measure that was capable of producing data on the relative strength of individual relational responses. Nevertheless, the IRAP quickly emerged as a measure of implicit cognition, but one that differed from the IAT". (p12, Hussey, Barnes-Holmes, and Barnes-Holmes, 2015).

To put it bluntly, therefore, the IRAP is of interest to RFT researchers only insofar as the measure offers us the ability to assess the relative strength of relational responses, because these responses provide the units of human language and cognition.

The IRAP: Procedural and Analytic Overview

At this point it is worth considering how the IRAP aims to provide a measure of the relative strength, persistence, or probability of relational responses. The IRAP is a computer based task on which an individual responds to a series of screens which contain verbal stimuli (i.e. verbal as defined by RFT). What are termed the labels appear at the top of the screen, stimuli such as "FLOWER" and "INSECT". What are termed the target stimuli appear in the middle of the screen such as "PLEASANT", "GOOD", "UNPLEASANT", and "BAD". Which labels and targets appear on screen at any point in the task is quasi-random, with the constraint that all labels will appear with all targets at some point during each block. On each trial two response options are provided, which specify particular relationships between the label and target stimuli. For example, "Flower" and "Pleasant" might appear on a given trial with the response options "True" and "False", and in this case participants would be required to confirm (pick "True") or deny (pick "False") that flowers are pleasant.

The IRAP operates by requiring orthogonal patterns of responding across successive blocks. For example "FLOWER" and "PLEASANT" would require the response "True" on one block and "False" on the next block. The IRAP operates on the assumption that the more frequently reinforced, and thus more probable, response pattern will be emitted more readily (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). In order to increase the likelihood that the more probable response is emitted, responding on the IRAP is placed under time pressure. Within the verbal community certain relational responses are more likely to be reinforced than punished (e.g. affirming that flowers are pleasant), while others are more likely to be punished than reinforced (e.g. denying that flowers are pleasant). Thus, the more readily emitted pattern of responding is indicative of the natural contingencies operating in the wider verbal community. Broadly speaking, the IRAP is scored by subtracting the mean response latency for one pattern of responding from the mean response latency of the opposite pattern of responding. Any resultant difference is deemed to be reflective of the differential reinforcement for the two patterns of responding in the pre-experimental history of the individual (note that the difference score is typically normalized using an appropriate statistical method, such as Cohen's *d* or a derivative).

Brief and Immediate Relational Responses versus Extended and Elaborated Relational Responses

In considering the types of effects that have been obtained on the IRAP, behavioural researchers have often referred to brief and immediate relational responses (BIRRs), which are emitted relatively quickly within a short window of time after the onset of the stimuli presented on any given IRAP trial. In contrast, extended and elaborated relational responses (EERRs) are more complex and are seen as being emitted more slowly and as such occur over a longer period of time (Barnes-Holmes et al., 2010; Hughes, Barnes-Holmes, and Vahey, 2012). The distinction between BIRRs and EERRs was first formalized in the context of the Relational Elaboration and Coherence (REC) model, which was offered as an initial RFT approach to implicit cognition (Barnes-Holmes et al., 2010; Hughes, Barnes-Holmes, & Vahey, 2012). The basic idea behind the model is that the types of effects observed on the IRAP, and indeed other implicit measures, were due to the fact that the task targeted BIRRs rather than EERRs. For example, the fact that the IRAP requires participants to respond relatively quickly on each trial, almost by definition, forces the participant to emit a BIRR. The relative strength of probability of this BIRR is deemed to be a function of the behavioural history of the participant with regard to functionally similar stimuli in that participant's history.

Imagine, for example, a white individual who had resided exclusively in white neighbourhoods, has no non-white friends or family members, and has been exposed to many media images of black people as violent drug dealers and inner-city gang members. When presented with an IRAP that presented pictures of black males carrying guns it is likely, according to the REC model, that BIRRs for confirming that black men are "dangerous" and "criminals" may be more probable than denying such relations. As a result, the participant may respond more rapidly on the IRAP when required to confirm, rather than deny, that a black man carrying a gun is dangerous. In effect, an anti-black racial bias may be revealed by the IRAP. In contrast, such a bias might be absent if the same participant was asked to rate the pictures of the black men from the IRAP with no time constraints for doing so. The lack of racial bias in the latter context is explained in the REC model by appealing to EERRs, which occur given sufficient time for an individual to respond in accordance with a relationally coherent network. In the context of the current example, the participant might fail to report any initial BIRR that involves perceiving the pictures of black males as "dangerous" based on additional relational responding, such as "It is wrong to discriminate on the basis of race" and "I am not a racist", etc. In general, therefore, the REC model attempts to explain the emergence of specific response biases on the IRAP by arguing that the procedure tends to reveal BIRRs rather than EERRs.

The IRAP, Rules, and BIRRs and EERRs

In concluding that the IRAP reveals BIRRs rather than EERRs, the REC model assumes that this applies, with roughly equal force, to all four trial-types. In other words, the IRAP is seen as providing a measure of the strength or probability of four functionally independent BIRRs. Imagine, for example, an IRAP that aimed to assess the response probabilities of four well-established verbal relations pertaining to non-valenced stimuli such as shapes and colours. Across trials, the two label stimuli, "Colour" and "Shape" could be presented with target words consisting of specific colours ("Red", Green", and "Blue") and shapes ("Square", "Circle", and "Triangle"). As such, the IRAP would involve presenting four different trial-types that could be designated as (i) Colour-Colour, (ii) Colour-Shape, (iii) Shape-Colour and (iv) Shape-Shape. One may conceive of these four trial-types as participating in a single relational network as depicted in Figure 1. According to the REC model responding in accordance with this network, such as describing each of the relations involved, may be seen as an EERR because it would require a relatively elaborated and extended relational response pattern. As such, each element within that pattern of relational responding would be functionally dependent on the other elements. For example, saying that “a square is a shape and red is a colour” would cohere with saying that “red is not a shape and a square is not a colour.” In other words, responding in accordance with such relational networks requires, by definition, that coherence within the network is maintained. Maintaining such coherence, particularly as networks become more extended and elaborated, requires time. In contrast, an IRAP is completed under time pressure and thus the functional dependencies or coherence within a given network may not be maintained.

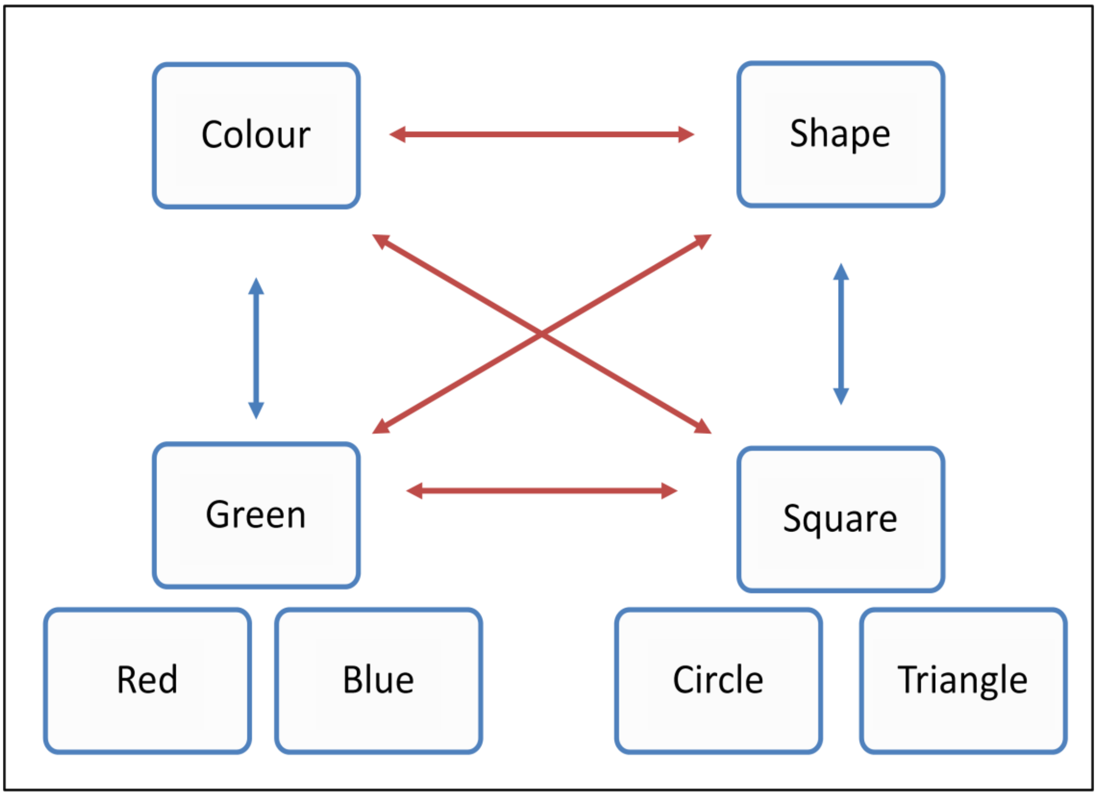

Figure 1. Representation of the relational network that is assessed by a shapes and colours IRAP. Red lines indicate relation of opposition between stimuli, while blue lines indicate relations of co-ordination between stimuli

To appreciate this point, consider again the nature of the IRAP itself. During a "Shapes and Colours" IRAP, participants would be required to respond in a manner that was consistent with their pre-experimental histories during some blocks of trials; (i) Colour-Colour-True; (ii) Colour-Shape-False; (iii) Shape-Colour-False; and (iv) Shape-Shape-True. On other blocks of trials the participants would have to respond in a manner that was inconsistent with those histories; (i) Colour-Colour-False; (ii) Colour-Shape-True; (iii) Shape-Colour-True; and (iv) Shape-Shape-False. All things being equal, the REC model assumes that the IRAP requires participants to emit BIRRs (rather than EERRs) on each trial. Thus, when the four trial-type effects are calculated, by subtracting response latencies for history-consistent from history-inconsistent blocks of trials, functionally independent effects would be observed for each trial-type. For example, when a participant responds, Shape-Shape-True on a given trial, he or she is not required to cohere that response with other elements within the network on that particular trial (because there is, in principle, insufficient time to do so). Or more informally, when an individual responds on any given trial in an IRAP the REC model assumes that it is not possible to "work through" or "rehearse" every other element within that network to ensure overall relational coherence. The current thesis constitutes a first step in testing this basic assumption by providing participants with rules or relational networks of differing levels of complexity on how to respond during exposure to a "shapes and colours" IRAP.

To appreciate the rationale for the research reported herein, consider again the basic REC model assumption that on any given IRAP trial it is not possible, due to time constraints, to check for relational coherence with the network that constitutes all four trial-types. In practice, however, this may not be the case. If a participant is capable of completing a private relational response on an IRAP trial within approximately 800-1000ms, there is an additional 1000-1200ms available for relational activity that extends beyond the targeted response. Crucially, if participants were encouraged to engage in such additional relational activity, perhaps through the provision of specific rules, this may well impact upon the IRAP performance. For example, imagine that participants were instructed on a shapes and colours IRAP to "respond as if colours are colours and shapes are shapes". If participants emitted this rule (privately) on many if not most of the trials before emitting a response it may reduce the BIRR-like properties of such responses. In other words, encouraging participants to rehearse a rule that singles out particular parts of a relational network (or specific trial-types in an IRAP) may in fact generate a pattern of relational responding that does not support the assumption that the IRAP targets BIRRs on each of the four trial-types. The three studies reported in the current thesis aimed to test this suggestion.

Experiment 1

The first study presented in the current thesis aimed to test the extent to which specific versus minimal rules presented before each block of trials in an IRAP impacted upon the relative sizes of the four trial-type effects. Specifically, in Experiment 1 participants were provided with two types of rules before completing each block of a shapes and colours IRAP. Half of the participants were presented with a "detailed" rule that specified part of the "shapes and colours" relational network, targeting in particular two of the trial-types, and half of the participants were presented with a "minimal" rule that did not specify specific parts of the network or specific trial-types. Given the exploratory nature of the study no specific predictions were made Experiment 1, although some differences in the pattern IRAP effects across the two types of rules (detailed versus minimal) were expected.

Method

Participants

62 undergraduate students of psychology at Maynooth University participated in the experiment as part of a required practical element of their degree course. All participants completed an IRAP and a valence questionnaire. The sample comprised of 42 females and 20 males with an age range of 17 to 43 years. The participants were randomly assigned to one of two groups, labelled here as the minimal-rule (MR) group (N=32), and the specific-rule (SR) group (N=30). The distinguishing feature for the MR-group was that they were simply instructed to respond correctly or incorrectly to the stimuli presented by the IRAP but in a manner that was always consistent with the feedback contingencies of the IRAP itself. In contrast, the SR-group was instructed to respond in accordance with a relatively

precise rule but which was also always consistent with the IRAP's feedback contingencies.

Materials

IRAP

The IRAP was presented on standard personal computers. The IRAP software was used to present the instructions, the stimuli, and to record responses. Each trial presented one of two labels; "Colour" or "Shape." The label stimulus was presented along with one of six target stimuli. The target stimuli were all words, three denoting colours; "Red," "Green," and "Blue"; the other three target stimuli denoting shapes; "Triangle," "Circle," and "Square." The rules provided to the MR-group were; (1) "Respond correctly to the stimuli," and (2) "Respond incorrectly to the stimuli." The rules provided to the SR-groups were; (1) "Colours are colours and shapes are shapes," and (2) "Shapes are colours and colours are shapes."

Questionnaire

A questionnaire comprising of demographic information (age, gender) and a rating scale for the stimuli presented in the IRAP was completed by all participants. The rating scale comprised of a seven point scale ranging from very negative to very positive. Each of the labels and targets used in the IRAP were rated.

Procedure

Details of the group-based session. The study was completed in two group-based sessions in the computer laboratory in the Department of Psychology at Maynooth University. The laboratory contained 33 desktop computers with the IRAP program installed on each machine. The experiment was run across two one-hour sessions, with the first session scheduled between 2-3pm and the second session scheduled for 3-4pm. Participants attending the 2-3 session received the minimal rules and participants attending the 3-4 session received the specific rules. Upon entering the computer laboratory, participants were asked to select a computer to sit at and to wait for further instruction. When all participants for that session were seated at a computer, the researcher commenced the experiment.

The researcher welcomed the participants and informed them that he would present a block of the IRAP program via the overhead projector so that everyone would have the opportunity to see what they would be asked to do during the experiment. He then proceeded to present a block of trials from the IRAP while describing what responses were required and the feedback contingencies that would apply. Thus, for example, a trial might have been presented in which the word “Shape” appeared as a label and the word “Circle” appeared as a target. The researcher pointed to these two stimuli and then stated that in this case responding “True” rather than “False” would be required, and to do this pressing the “D” key (rather than “K”) was appropriate. The researcher then noted that when this response was emitted the stimuli disappeared from the screen and were replaced almost immediately with the stimuli for another trial. The researcher continued to present a block of trials in this way providing examples of responding correctly and incorrectly within that block, noting for example, that an incorrect response produced a red X in the middle of the screen and the program only continued to the next trial when the correct response was emitted.

When the researcher had worked through a single block of trials on the overhead projector, he invited the participants to click the mouse for their particular PC and follow the instructions that were presented on screen. Participants were also instructed to turn off mobile phones and any other electronic devices that might distract them from the experiment and to refrain from talking until everyone had completed the experimental session. A number of volunteer graduate students remained in the room throughout the session to ensure that all participants complied with this instruction.

The IRAP.

On each trial of the IRAP, four words appeared on screen; a label at the top centre of the screen ("Colour", or "Shape"); a target at the centre of the screen ("Red", "Green", "Blue", "Circle", "Square", or "Triangle"), and the two response options "True" and "False" at the bottom left and right of the screen, respectively. Participants responded on each trial using either the "D" key for the response option on the left or the "K" key for the response option on the right. The locations of the response options (the words, "True" and "False") alternated from trial to trial in a quasi-random order such that they did not remain in the same left-right locations for more than three successive trials. Examples of each type of trial to which participants were exposed are shown in Figure 2.

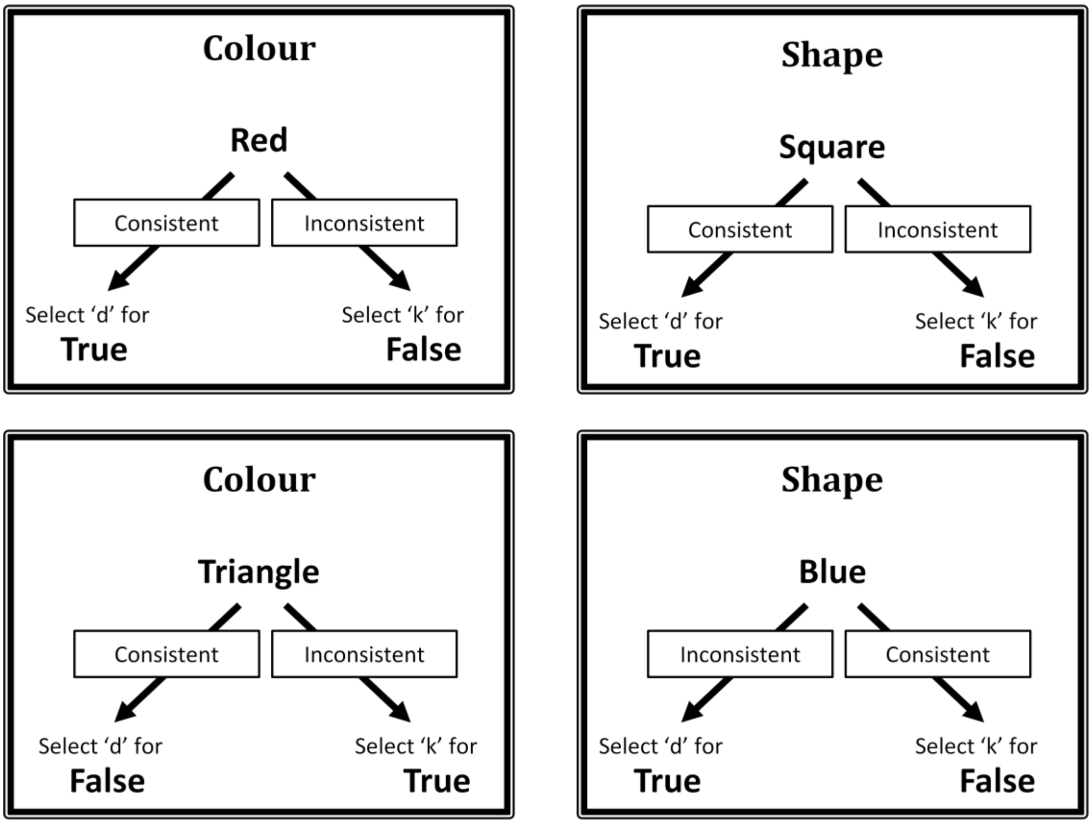

Figure 2. Diagrammatic representation of the four IRAP trial-types presented to participants.

When participants selected the response option that was deemed correct within that block of trials the label, target and response option stimuli were removed immediately from the screen for an inter-trial interval of 400 ms, after which the next trial was presented (i.e., a label, target and two response options appeared simultaneously). When participants selected the response option that was deemed incorrect for that block of trials the stimuli remained on screen and a red X appeared beneath the target stimulus. The participants were required to select the correct response option, and only then did the program proceed directly to the 400 ms inter-trial interval (followed immediately by the next trial). This pattern of trial presentations, with corrective feedback, continued until the entire block of 24 trials was presented. The IRAP program presented the trials in a quasi-random order within each block with the constraint that each label was presented twice with each target stimulus across the 24 trials. Consistent with the majority of previously published IRAP studies, the trials presented within each block may be described as consisting of four different trial-types. In the current study, the four different combinations of label and target stimuli may be denoted as (i) Colour-Colour, (ii) Colour-Shape, (iii) Shape-Colour, and (iv) Shape-Shape (see Figure 2).

When participants completed a block of trials, the IRAP program provided them with feedback on their performance during that block. The feedback consisted of a message informing them how accurately they had responded in terms of percentage correct and how quickly they had responded in terms of median reaction time. The latter was calculated from stimulus onset to the first correct response across all 24 trials within the block. Participants were required to achieve a minimum accuracy of 80 percent correct and a maximum median latency of no more than 2000 milliseconds. The IRAP program was set to allow participants up to 4 pairs of practice blocks to achieve these criteria. Only when participants achieved these criteria across both Blocks 1 and 2, or Blocks 3 and 4, or Blocks 5 and 6, or Blocks 7 and 8 were they permitted by the IRAP program to continue to the critical test blocks.

A fixed set of six test blocks were presented with no accuracy or latency criteria in order for participants to progress from one block to the next. However, percentage correct and median latency were presented at the end of each block to encourage participants to maintain the accuracy and latency levels they had achieved during the practice blocks (see below). In addition to the accuracy and latency feedback the IRAP program provided participants with tailored feedback instructions depending on their level of performance in the previous block. Participants who failed to achieve the both accuracy and latency criteria were presented with the following text on screen between blocks “Learn to accurately follow the rule before attempting to respond quickly”. If the accuracy criterion had been achieved but not the latency criterion, the program presented the instruction “Continue responding as accurately as you can. You'll naturally go quickly when your responses are accurate.” If both latency and accuracy criteria had been met the program presented the following instruction “Continue responding both as accurately and quickly as you can.” These instructions were presented following both practice and test blocks.

Two types of feedback contingencies were applied across the practice and test blocks of the IRAP, denoted here as consistent versus inconsistent with the natural contingencies operating in wider verbal community. The contingencies deemed consistent required participants to choose “True” on *Colour-Colour* and *Shape-Shape* trial-types and to choose “False” on *Colour-Shape* and *Shape-Colour* trial-types. The contingencies deemed inconsistent required participants to respond in an orthogonal pattern, choosing “False” on *Colour-Colour* and *Shape-Shape* trial-types and choosing “True” on *Colour-Shape* and *Shape-Colour* trial-types. The IRAP program typically applies the feedback contingencies in one of two patterns. For one pattern, the first block and all subsequent odd numbered blocks employ the consistent feedback; the second and all subsequent even numbered blocks employ the inconsistent feedback. For the second pattern, the first block and all odd numbered blocks employ the inconsistent feedback and the second and all even numbered block employ the consistent feedback. The use of these two patterns of feedback contingencies was counterbalanced across the participants in the current study. In other words, half of the participants commenced with consistent feedback and then alternated from inconsistent to consistent across all subsequent blocks; the other half commenced with inconsistent feedback and then alternated from consistent to inconsistent across blocks thereafter.

As noted previously, participants were divided into two groups, the MR and SR groups. For both groups two rules were presented, and these were determined by the feedback contingencies that were employed in the immediately following block of trials. For the MR group, each block of trials that employed the consistent feedback presented the rule: “Respond correctly to the stimuli” and each block that employed the inconsistent feedback presented the rule “Respond incorrectly to the stimuli”. In other words, the rules informed participants how to respond during the next block of trials in a way that would avoid the red X, without specifying exactly how to do so. For the SR group, each block of trials that employed the consistent feedback presented the rule: “Respond as if colours are colours and shapes are shapes” and each block that employed the inconsistent feedback presented the rule “Respond as if shapes are colours and colours are shapes”. In this case, therefore, the rules informed participants how to respond during the next block of trials by specifying exactly what pattern of responding to each trial-type was required to avoid the red X.

Upon completion of the IRAP, all participants proceeded to the questionnaire. Thereafter participants were thanked for their time, debriefed, and dismissed.

Results and Discussion

The primary datum of the IRAP is the response latency defined as the length of time in milliseconds from stimulus presentation to a correct response on a particular trial. If participants maintained the accuracy and latency performance criteria across all six test blocks the data from all blocks were included in the analyses. If, however, a participant failed to maintain the criteria on one or both blocks within a given test-block pair (blocks 1 and 2; blocks 3 & 4; blocks 5 & 6), the data from that pair of test blocks were removed from the analyses. If a participant failed to maintain the criteria on two or more blocks from different test block pairs all of the data from that participant were removed. This practice was similar to that adopted by Nicholson and Barnes-Holmes (2012) in order to avoid excessively high attrition rates. Application of these criteria resulted in the exclusion of 20 of the 62 participants form the analysis, which is relatively high but may be accounted for, at least in part, by the fact that the IRAP was conducted in a group setting.

Each participant who completed the current IRAP produced 24 response latencies for each test block. Due to a technical fault in the program's recording of the first trial in each block the latency scores for the first trial were removed from all analyses. The remaining 23 latencies in each test block were converted into the *D*-IRAP scores. For participants who maintained the accuracy and latency criteria across all three pairs of test blocks, the *D*-IRAP scores were calculated as follows:

1.) If 10% of a participant's response latencies were less than 300 ms all of their data were removed (no participant had their data removed on this basis);
2.) All latencies over 10,000 ms were removed;

3.) Twelve standard deviations for the response latencies, calculated for each trial-type, were obtained across the three pairs of test blocks (i.e., blocks 1 & 2; blocks 3 & 4, and Blocks 5 & 6);

4.) Twenty four mean latencies were calculated, one for each trial-type in each block.

5.) A difference score was calculated for each trial-type, in each test block pair, by subtracting the mean latency in the consistent block from the mean latency in the inconsistent block, thus producing 12 difference scores.

6.) The difference score for each trial-type in each test block pair was divided by the standard deviation for that trial-type from the corresponding test blocks, resulting in 12 *D*-IRAP scores – one for each trial-type in each pair of test blocks.

7.) Four *D*-IRAP scores were calculated, one for each trial-type, by averaging scores across the three pairs of test blocks.

The same general method for calculating *D*-IRAP scores was also applied to the data from participants who had data from a pair of test blocks removed except the algorithm was adjusted accordingly (e.g., 8 standard deviations were calculated in step 3 and 16 mean latencies were calculated in step 4).

The foregoing calculations yielded four mean *D*-IRAP scores for each participant, one for each trial-type. A preliminary 2x2x4 mixed repeated measures analysis of variance (ANOVA) was conducted to determine if the sequence in which the IRAP blocks were presented (i.e., consistent-first versus inconsistent-first) impacted significantly upon the *D*-IRAP effects across the four trial-types and two instruction conditions. The main effect for block order and its interactions with the other two variables were all non-significant (*ps* > .06), and thus this procedural variable was removed from all subsequent analyses.

The overall mean *D*-IRAP scores calculated across participants are presented in Figure 3, divided according to the type of instructions they received (SR versus MR). All *D*-IRAP effects were greater than zero, which indicates that both groups responded more quickly during history-consistent than history-inconsistent blocks for each of the trial-types. In effect, participants tended to respond "True" more quickly than "False" when presented with the label "Colour" and the name of a colour, and when presented with the label "Shape" and the name of a shape; conversely, participants tended to respond "False" more quickly than "True" when presented with the label "Colour" and the name of a shape, and when presented with the label "Shape" and the name of a colour.

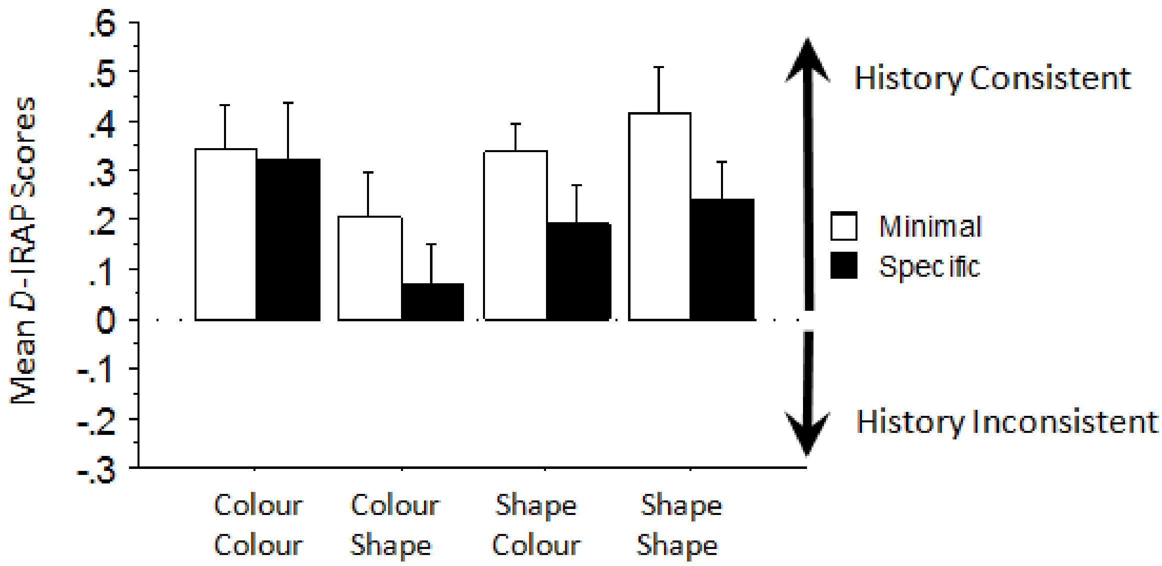

Figure 3. Mean D-Scores, with standard error bars, for each trial-type for each condition in Experiment 1. Positive score indicates effects in a history consistent direction.

In comparing the mean *D*-IRAP scores between the SR and MR conditions the *Colour-Colour* trial-type yielded very similar effect sizes, whereas the remaining three trial-type produced larger effects in the MR relative to the SR condition. In general, the differences between the *Colour-Colour* trial-type and the other three trial-types were less pronounced for the MR relative to the SR condition. The *D*-IRAP scores were subjected to a 2(*rule-type*) X 4(*trial-type*) mixed repeated measures ANOVA, which failed to yield any significant main or interaction effects ($ps > .06$). Although the effect for rule-type was non-significant, the alpha value was relatively low with an effect size that fell between low and moderate (Cohen, 1988), $F(1,40) = 2.79$, $p = .10$, partial eta squared = .07.

Four one-sample *t*-tests indicated that the *D*-IRAP scores for the MR group each differed significantly from zero; *Colour-Colour*, $t(1,20) = 3.56$, $p = .001$; *Colour-Shape*, $t(1,20) = 2.43$, $p = .25$; *Shape-Colour*, $t(1,20) = 5.88$, $p < .0001$; *Shape-Shape*, $t(1,20) = 4.49$, $p = .0002$. An additional four one-sample *t*-tests (for the SR-group) yielded significant effects for three of the four trial-types; *Colour-Colour*, $t(1,20) = 2.77$, $p < .05$; *Shape-Colour*, $t(1,20) = 2.29$, $p <.05$; and *Shape-Shape*, $t(1,20) = 3.04$, $p < .01$ (remaining $p > .42$).

Overall, therefore, the descriptive statistics suggested that the MR condition yielded stronger *D*-IRAP effects that differed less among the trial-types than was the case for the SR condition. The inferential statistics did not provide firm support for this description of the results, although the one-sample *t*-tests were suggestive (i.e., four significant effects for the MR condition versus only three for the SR condition).

As noted above, only one of the three trial-types produced *D*-IRAP effects that were highly similar across the two rule conditions (the *Colour-Colour* trial-type). In the introduction it was speculated that providing detailed instructions may serve to produce "genuine" BIRRing on perhaps one or two trial-types, but more EERR-like responses on the remaining trial-types. Insofar as this was the case for the *Colour-Colour* trial-type for the SR condition in the current experiment, then it may be informative to conduct an inferential statistical analysis with this particular trial-type removed from the data set. In other words, the analysis would be conducted on the three trial-types that may reflect functionally distinct response patterns (i.e., BIRRs versus EERRs). A 2x3 mixed repeated measures ANOVA was therefore conducted, which examined the differences between the two rule conditions across each of the three trial-types (*Colour-Shape*, *Shape-Colour*, and *Shape-Shape*), and this yielded a significant main effect for rule-condition, $F(1,38) = 4.22$, $p = .046$, partial eta squared = .10. In effect, providing minimal rules before each block appeared to produce a significantly stronger IRAP effect than when detailed rules were provided.

Of course, this latter set of analyses are rather post-hoc and could be seen as "cherry picking" the data that supports the current theoretical arguments. While recognising that there may well be some substance to the criticism, it is also important to acknowledge a trend in the data that supports the conclusion that the two rule conditions impacted somewhat differently on the IRAP effects observed in this first study. It is also worth bearing in mind that the study also suffered from a number of technical and procedural problems and thus the clarity of the effects obtained may have been undermined somewhat by these factors. In view of these various issues it was decided to replicate the current research while correcting for a number of issues or problems that it was felt might have contaminated the *D*-IRAP effects obtained therein.

Experiment 2

As noted previously, Experiment 1 suffered from a technical problem, which required that the response latency for the first trial in each block be removed from the data analyses. The impact of losing these data points remains unknown but it is possible that doing so impacted, if only slightly, on the pattern of results obtained. Replicating the current study with an IRAP program that did not contain the “bug” thus seemed wise. In addition, there were two procedural issues associated with Experiment 1 that required attention. First, Experiment 1 was conducted in a group setting with approximately 30 participants completing the IRAP in a large computer laboratory. The standard practice for running IRAP studies, at least at the Maynooth site, is for individual participants to complete the procedure in small sound-attenuated booths on a one-to-one basis with the researcher. Experiment 2 thus adopted this “standard” practice of running participants individually. Second, the two types of instructions provided to participants in the SR condition confounded two variables. That is, the consistent rule was “Colours are colours and shapes are shapes”, whereas the inconsistent rule was “Shapes are colours and colours are shapes”. In effect, the first word (“Colours” versus “Shapes”), and whether the rule was deemed consistent or inconsistent with common verbal practices, were manipulated across the two rules. Typically, when specific rules have been presented in IRAP studies, only the latter variable has been manipulated. Recent research indicates when specific rules are presented at the beginning of each block in an IRAP, they may have a significant impact on the resulting IRAP effects (O’Shea et al., 2015). Thus, in the next Experiment the two specific rules always involved presenting the same word (“Colours”) at the beginning.

Method

Participants

67 undergraduate students of psychology at Maynooth University voluntarily participated in the experiment. Participants were not offered any form of remuneration for their participation. All participants completed an IRAP, and a series of questionnaires. The sample comprised of 32 females and 34 males with an age range of 17 to 45 years. Thirty-nine participants were exposed to the MR condition and 28 were exposed to the SR condition.

Materials

IRAP

The IRAP program was similar to the version employed in Experiment 1, except the “bug” that contaminated the first trial in a block was not present and the specific rule for the inconsistent blocks read “Colours are shapes and shapes are colours”. The label and target stimuli that were employed in Experiment 1 were also employed in Experiment 2.

Questionnaires

The demographics questionnaire and rating scales (for shapes and colours) employed in Experiment 1 were also employed in Experiment 2.

Procedure

The procedure was similar to Experiment 1, except that each participant attended the laboratory and completed the IRAP, and the questionnaires, individually in a small sound-attenuated cubicle. The introduction to the IRAP also differed, given that it was not presented in a group setting. At the beginning of the experiment for each participant the researcher initiated the IRAP program so that it presented the first rule. The researcher then presented the participant with a print out of the four-trial-types from the IRAP, and explained that each one was representative of the tasks that would appear on screen during the experiment. Participants were told that they would respond to each trial-type by pressing either the “D” key for the response option on the left, or the “K” key for the response option on the right. The researcher described the pattern of responding that was required based on the instruction that was presented to the participant. For example, if the instruction read “Colours are colours and Shapes are shapes” the researcher selected a representative trial on the print out – for example “Colour” appeared as a label with the word “Red” as a target -- and stated that in this case responding “True” rather than “False” was required. The researcher then described what would happen if participants responded correctly (i.e., the screen would clear and the next trial would be presented) and what would happen if participants responded incorrectly (a red X would appear and a correct response would be required before the program continued to the next trial). When the researcher had worked through the four examples of each trial-type, he invited the participants to press the spacebar and engage with the task in accordance with the instructions.

All procedural features of the IRAP were the same as those described for Experiment 1. Upon completion of the IRAP, all participants completed the questionnaires, and thereafter were thanked for their time, debriefed, and dismissed.

Results and Discussion

The data from the IRAP were prepared for analysis in a similar manner to that employed for Experiment 1, except the data point from the first trial in each block was not excluded (because there was no bug in the program). A preliminary

2x2x4 mixed repeated measures analysis of variance (ANOVA) was conducted to determine if the sequence in which the IRAP blocks were presented (i.e., consistent-first versus inconsistent-first) impacted significantly upon the *D*-IRAP effects across the four trial-types and two instruction conditions. Unlike Experiment 1, block order interacted significantly with the other two main variables, and thus it was not removed from subsequent analyses. The mean *D*-IRAP scores for each trial-type divided according to instruction type and block order are presented in Figure 4.

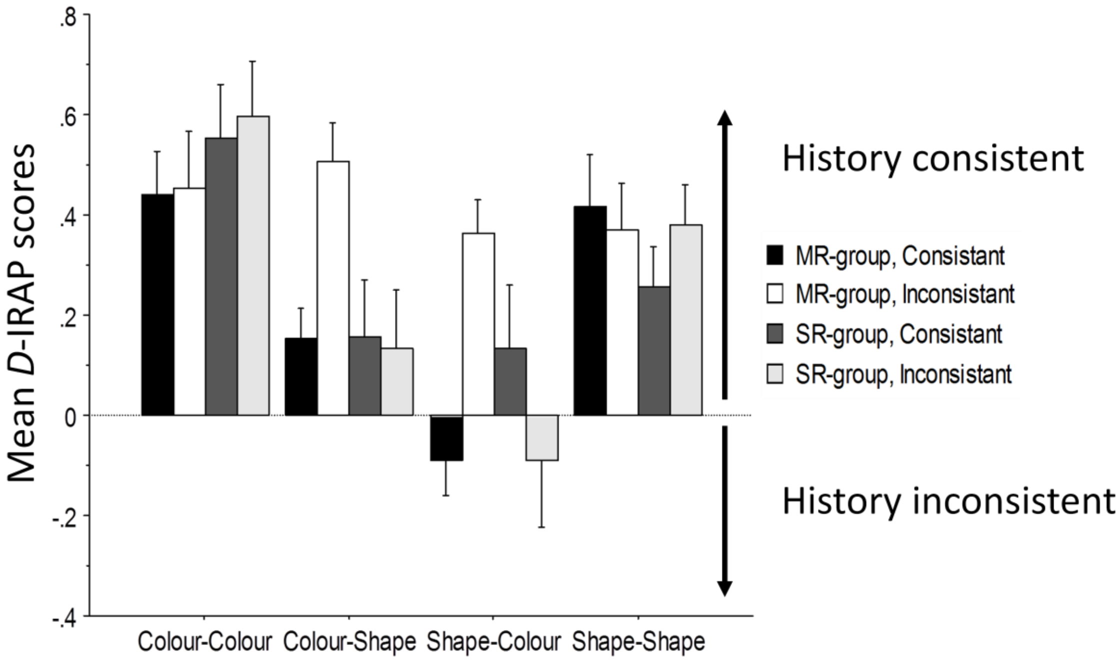

Figure 4. Mean D-Scores, with standard error bars, for each group and block order in Experiment 2. Positive scores indicate effects in a history consistent direction.

The *D*-IRAP effects for the *Colour-Colour* and *Shape-Shape* trial-types were all relatively strong and in a history-consistent direction; the effects for the *Colour-Shape* trial-type were somewhat weaker except for the group who received minimal instructions and commenced the IRAP with an inconsistent block. The *D*-IRAP effects for the *Shape-Colour* trial-type again produced relatively weak effects except for the MR/inconsistent-first group, but on this occasion two of the weak effects were in a history-inconsistent direction (for the MR/consistent-first and SR/inconsistent-first conditions). Overall, therefore, it appears that the type of instructions and the order in which the blocks were presented impacted upon the *D*-IRAP effects recorded across the four trial-types with the exception of the MR/inconsistent-first group.

The results of the 2x2x4 ANOVA revealed a main effect for trial-type, $F(3,45) = 17.47$, $p < .0001$, partial eta squared = .18, but as noted above this was moderated by a three-way interaction, $F(3,45) = 4.87$, $p = .003$, partial eta squared = .08. The nature of this interaction was explored using four follow-up one-way between-participant ANOVAs for each trial-type and four one–way within-participant ANOVAs comparing the *D*-IRAP scores for the four combinations of instruction type and block order. Two of four between-participant ANOVAs proved to be significant, one for the *Colour-Shape* trial-type, $F(3,45) = 4.2$, $p = .01$, eta squared = .22, and the other for the *Shape-Colour* trial-type, $F(3,45) = 5.07$, $p = .004$, eta squared = .25, but the other two did not (*ps* > .6). Three of the four within-participant one-way ANOVAs each proved to be significant; MR/consistent-first, $F(3, 39) = 9.73$, $p < .0001$, partial eta squared = .43; SR/consistent-first, $F(3, 30) = 5.59$, $p = .004$, partial eta squared = .33; SR/inconsistent-first, $F(3, 30) = 10.31$, $p = .0001$, partial eta squared = .51. The ANOVA for the MR/inconsistent-first condition yielded a non-significant effect with a small effect size, $F(3, 36) = .574$, $p = .63$, partial eta squared = .004. The results of 16 one-sample *t*-tests for each of the *D*-IRAP effects are presented in Table 1. The only condition that produced significant effects across all four trial-types was the MR/inconsistent-first condition; the other three conditions all yielded significant effects for the *Colour-Colour* and *Shape-Shape* trial-types. Overall, therefore, the inferential statistics confirmed the conclusion, arising from visual inspection of Figure 4, that instructions and block order impacted upon two of the four IRAP trial-types (*Colour-Shape* and *Shape-Colour*) for three of the four groups.

Table 1.

The t-scores and p-values from the one sample t-tests for each trial-type effect for each group in Experiment 2.

Subject Group	**Trial- type 1**	**Trial-type 2**	**Trial-type 3**	**Trial-type 4**
MR-Con	5.04 (*p* < .0002)	2.72 (*p* < .02)	-1.25 (*p* = .23)	3.96 (*p* < .002)
MR-Incon	4.04 (*p* < .001)	6.58 (*p* < .0001)	5.33 (*p* < .0002)	3.86 (*p* = .002)
SR-Con	5.33 (*p* < .0003)	1.41 (*p* = .18)	1.08 (*p* < .31)	3.09 (*p* < .012)
SR-Incon	5.34 (*p* < .0003)	1.16 (*p* = .27)	-.68 (*p* = .51)	4.74 (*p* <.0008)

At this point in the research programme it became clear that the type of instructions provided to participants when they complete an IRAP may impact quite dramatically on their performances, but this instructional effect was moderated by the order in which the IRAP blocks were presented. The fact that the specific rule conditions both produced relatively strong *D*-IRAP effects on the Colour-Colour and Shape-Shape trial-types is consistent with the argument that the block rules facilitated the BIRR-like properties of these two relational responses during history-consistent blocks. Interestingly, the minimal rule appeared to produce significant history-consistent *D*-IRAP effects across the four trial-types but only when participants commenced the IRAP with a history-inconsistent block of trials. A possible explanation for this somewhat unexpected outcome will be considered below.

Experiment 3

In the previous two experiments two types of rules were employed, specific and minimal. The specific rules clearly focused on particular parts of the relational network and two of the trial-types. Although the minimal rules did not explicitly specify particular parts of the network (or trial-types) it is possible that the instruction to "respond correctly to the stimuli", for example, served to highlight the coherent relations within the network and the trial-types. In effect, the word "correct" coordinated with the pre-experimentally established coherent relations (i.e., Colour-Colour and Shape-Shape). Consequently, participants may have generated their own "specific" rules, which functioned similarly to the two rules presented in the specific-rules condition. The generation of such rules would likely be enhanced when participants commenced with a block of history-consistent trials because the IRAP would require a pattern of responding that cohered with the participants' pre-experimental verbal histories. The foregoing outcome may have been far less likely when the IRAP commenced with the instruction to "respond incorrectly to the stimuli" (and a block of history-inconsistent trials). Of course, it would be possible for participants to self-generate a rule that cohered with that initial IRAP block (similar to that presented within the specific-rule condition -- "Colours are shapes and shapes are colours"). However, the generation of such a rule may be less likely because it fails to cohere with the participants' pre-experimental histories. More informally, participants may have found it easier to generate a rule that cohered with their natural verbal relations than a rule that did not. Thus, there was observed a similar pattern of trial-type effects in the minimal rule condition to that observed in the detailed rule condition (i.e., a mix of significant and non-significant effects), but only when the IRAP commenced with a history-consistent minimal rule (i.e. “respond correctly to the stimuli”).

The foregoing explanation is of course quite speculative, but it does suggest that it should be possible to provide instructions that neither specify parts of the network being assessed, nor cue self-generated rules that serve a similar function. In an effort to achieve this objective, instructions were devised that served to emphasize each trial-type equally. That is, participants were instructed to respond as if “True” is consistent and “False” is inconsistent on history-consistent blocks, and to respond as if “True” is inconsistent and “False” is consistent on history-inconsistent blocks. These rules do not therefore specify correct or incorrect responding in terms of the network, which may highlight coherent over incoherent relations, but simply told participants how to respond to the response options on each trial-type. Would these rules produce four significant trial-type effects on an IRAP that commenced with a history-consistent block of trials?

Participants

Nineteen undergraduate students at Maynooth University voluntarily participated in the experiment. Participants were not offered any form of remuneration for their participation. All participants completed an IRAP, and a questionnaire. The sample comprised of 12 females and 7 males with an age range of 20 to 30 years.

Materials

IRAP

The IRAP program was similar to the version employed in Experiment 2, except the rules provided to the participants focused on the functions of the response options rather than the functions of the labels and targets for a particular block. The rule for the consistent block read “Please respond as if true is consistent and false is inconsistent”; for the inconsistent blocks the rule read “Please respond as if true is inconsistent and false is consistent”. The label and target stimuli that were employed in Experiments 1 and 2 were also employed in Experiment 3.

Questionnaires

The demographics questionnaire and rating scales (for shapes and colours) employed in Experiments 1 and 2 were also employed in Experiment 3.

Procedure

All procedural features of the IRAP were the same as those described for Experiment 2, except that all participants commenced with a history-consistent blocks of trials. In addition, the instructions were modified in order to focus on each trial-type equally by emphasizing the function of the response options. That is, before each history-consistent block of trials the instruction read “Please respond as if true is consistent and false is inconsistent”; before each history-inconsistent block of trials the instruction read “Please respond as if true is inconsistent and false is consistent.” Upon completion of the IRAP, all participants completed the questionnaires, and thereafter were thanked for their time, debriefed, and dismissed.

Results and Discussion

The data from the IRAP were prepared for analysis in the same manner to that employed for Experiment 2. The data from 7 participants were excluded from the analysis because they failed to maintain the latency and accuracy criteria for a minimum of two out of the three critical test block pairs. The mean *D*-IRAP scores for each trial-type are presented in Figure 5, which shows that all four effects were relatively strong and in a history-consistent direction; the effects for the Colour-Colour and Shape-Colour trial-types were stronger than the other two.

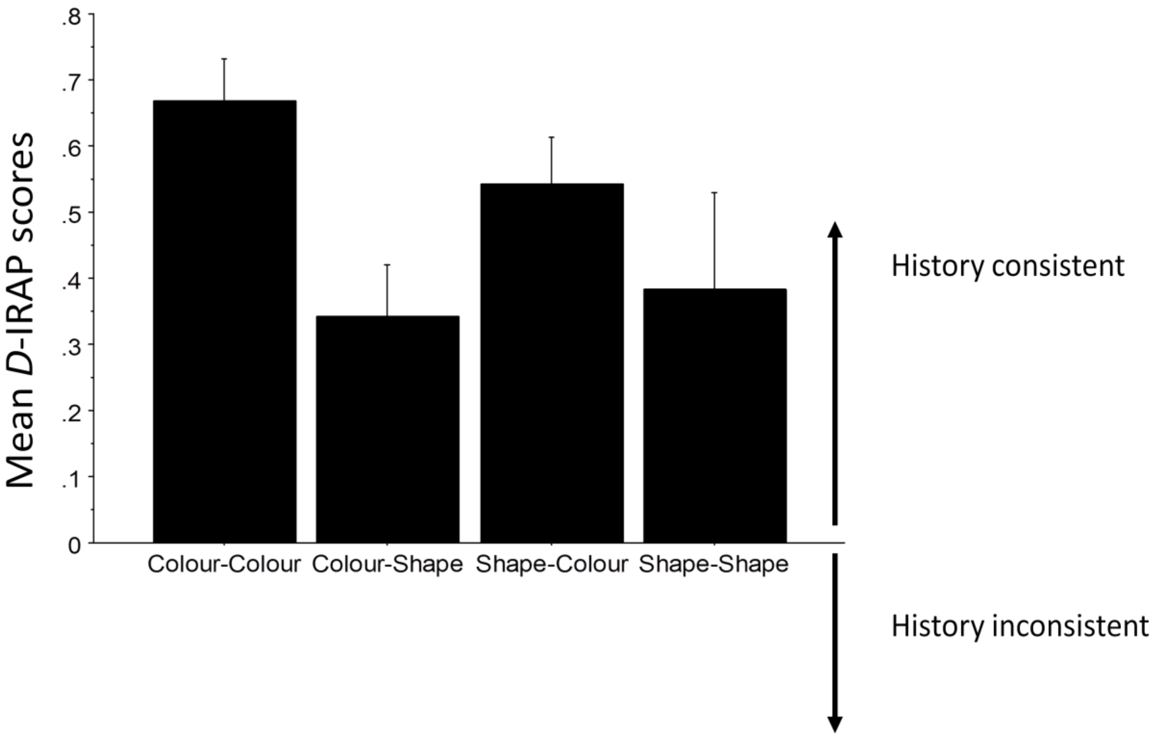

Figure 5. Mean D-Scores, with standard error bars, for each trial-type in Experiment 3. Positive scores indicate effects in a history consistent direction.

The results of a one-way within-group ANOVA revealed a marginally significant effect for trial-type, $F(3,33) = 2.81$, $p = .055$, partial eta squared = .07. Differences between trial-types were investigated by means of paired *t*-tests and these revealed that the *D*-IRAP effects for the *Colour-Colour* trial-type differed significantly from the *Colour-Shape,* $t(11) = 4.14$, p < .01, and *Shape-Shape*, $t(11) = 2.18$, $p = .05$, trial-types (remaining *ps* > .14). Four one-sample *t*-tests confirmed that each of the trial-type effects differed significantly from zero; *Colour-Colour* $t(11) = 10.52$, $p < .0001$; *Colour-Shape* $t(11) = 4.24$, $p = .001$; *Shape-Colour* $t(11) = 7.66$, $p < .0001$; *Shape-Shape* $t(11) = 2.65$, $p = .02$. In summary, therefore, providing response-focused instructions appeared to generate relatively strong (and significant) IRAP effects for each of the four trial-types, even when participants commenced the procedure with a history-consistent block of trials.

General Discussion

Overall Summary

The initial experiment presented in the current thesis sought to establish whether altering the level of specificity of the rules provided to participants undertaking an IRAP would impact on the observed pattern of responding. The data from this first experiment suggested that rules of differing levels of specificity may impact upon the pattern of responding. Although the observed difference between groups only tended toward statistical significance the pattern was sufficiently compelling to warrant further investigation. The interpretation of the data from Experiment 1 was hampered somewhat by procedural and technical issues, namely a bug in the program, the group setting in which the experiment was conducted, and the provision of rules that did not begin with the same word across the two types of IRAP blocks. That is, for history-consistent blocks the rule started with "colours are" whereas for history-inconsistent blocks the rule started with "shapes are". Given the suggestive data, combined with the potential procedural confounds, it was decided that refining Experiment 1 was in order.

Experiment 2 was a replication of Experiment 1, except the procedural and technical issues mentioned above were corrected. The data from this experiment showed that the level of specificity of the rule provided to participants impacted the pattern of responding, but this effect was clearly moderated by IRAP block order. That is, participants in the SR condition produced significantly different effects across trial-types in both block-order conditions (i.e. trial-types 1 and 4 yielded relatively strong effects but trial-types 2 and 3 produced relatively weak effects). A similar pattern of differences was also observed in the MR condition but only for the group who commenced with history-consistent IRAP blocks. The MR group who started with history-inconsistent blocks produced a pattern of IRAP effects each of which were relatively strong and significantly different from zero.

At this point it was speculated that MR participants were likely to generate self-directed rules that were similar to those presented to the SR group if initial contact with the contingencies of the IRAP cohered with their pre-experimental verbal histories. Or more informally, if the IRAP reinforces responding to colours as colours and shapes as shapes, then a participant will likely generate exactly this rule, which is functionally equivalent to the SR rule. As noted in the introduction, if participants tend to rehearse a rule that singles out particular parts of a relational network (or specific trial-types in an IRAP) this may generate a pattern of relational responding that does not support the assumption that the IRAP targets BIRRs on each of the four trial-types, hence producing the uneven pattern of effects observed in the SR condition. In the final Experiment, therefore, instructions were presented that were designed to discourage the production of self-directed rules that specified particular elements of the relational network or particular trial-types of the IRAP.

For Experiment 3 participants were presented with instructions that focused on the response options rather than the coherence or incoherence between specific relations within the network/trial-types. Specifically, participants were instructed to respond to “True” as consistent and “False” as inconsistent during history-consistent blocks and “True” as inconsistent and “False” as consistent during history-inconsistent blocks. All participants commenced the IRAP with a block of history-consistent trials. If the response-option focused instructions discouraged the generation of self-directed rules that specified particular parts of the network, then relatively strong and significant effects should be observed for each trial-type, and indeed this turned out to be the case.

At this point in the research programme, therefore, it had been established that providing detailed rules that specify particular parts of a relational network, and trial-types within the IRAP, may have quite dramatic effects on IRAP performances. Furthermore, it had also been found that even when minimal rules are presented, a participants initial contact with the IRAP contingencies (history-consistent versus inconsistent) may moderate the IRAP effects quite dramatically. The interpretation of this moderating effect (i.e., participants produced self-directed rules) was tested, with a third and final experiment and this provided evidence to support our speculative explanation. A range of conceptual issues arising from the current research will now be consisdered.

A Conceptual Analysis

At the end of the introduction to the current thesis it was suggested that presenting a rule or instruction that specified only particular parts of a relational network may generate IRAP effects that involve relational responding that is less BIRR-like than would be suggested by the REC model. The pattern of IRAP effects observed across the three studies reported in the current thesis provide some support for this argument. In closing, however, it seems important to provide a potential explanation for the full pattern of effects observed in the current research. The main finding that appears to require a clear explanation is the fact that Colour-Colour and Shape-Shape trial-types tended to produce relatively strong, history-consistent effects across all conditions, but the Colour-Shape and Shape-Colour trial-types did not. In effect, providing a detailed rule did not appear to undermine the IRAP effect for all four trial-types, but only for one or two of them. Why might this be the case?

One explanation is that given a history-consistent rule the IRAP involves relating a “same” relation to another “same” relation on trial-types 1 and 4, but relating a “same” relation with a “difference” relation on trial-types 2 and 3. That is, given the rule, “colours are colours and shapes are shapes” the first part of the rule coheres with trial-type 1 and the second part of the rule coheres with trial-type 4. Insofar as relational coherence is frequently reinforced within the wider verbal community, relational responding on these trial-types may occur relatively rapidly (i.e., it is BIRR-like). In contrast, neither parts of the rule cohere with trial-types 2 and 3 (i.e., colour-shape and shape-colour), which may serve to undermine BIRR-like responding. Now consider the history-inconsistent blocks. Trial-types 1 and 4 do not cohere with the rule, which may well reduce BIRR-like responding, relative to history-consistent blocks. Consequently, trial-types 1 and 4 will tend to yield relatively strong trial-types effects in a history-consistent direction. For trial-types 2 and 3, however, although they cohere with the rule for responding on the IRAP they do not cohere with the relational networks established by the wider verbal community (e.g., colours are not in fact shapes). Consequently, the BIRR-like properties of the relational responses produced on these two trial-types may also be undermined during the history-inconsistent blocks of trials. In summary, therefore, the specific rules presented in the current thesis may have supported BIRR-like responding for trial-types 1 and 4 during history-consistent blocks but not during history-inconsistent blocks; in contrast, BIRR-like responding may have been much reduced or entirely absent for trial-types 2 and 3 during both history-consistent and history-inconsistent blocks. Insofar as this was the case one would expect predictable

IRAP effects for trial-types 1 and 4 but unpredictable effects for trial-types 2 and 3, which is what was observed.

In offering the foregoing conceptual explanation for the pattern of results observed in the current thesis it must be recognized that they are rather speculative and important questions remain. For example, one finding that raises questions is the fact that although all four IRAP effect in Experiment 3 were significantly different from zero they did differ significantly from each other. Given that participants were provided with response-focused instructions, rather than rules which specified particular parts of a relational network, why did these differences emerge? One possibility is that even under these instructional conditions some participants' self-generated rules that were broadly similar to those presented during the SR conditions in Experiment 1 and 2. In making this argument it is important to bear in mind that many of the participants who completed Experiment 3 had previously participated in at least one IRAP study, which involved the provision of detailed rules. It is possible, therefore, that this pre-experimental history with the IRAP, and the provision of rules specifically, encouraged some participants to generate specific rules during Experiment 3, thus generating a somewhat uneven pattern of IRAP effects. Perhaps future research could explore the impact of previous IRAP exposures with and without detailed rules on subsequent IRAP performances.

Conclusion

The findings obtained across the three studies reported in the current thesis raise a number of important questions for the use of the IRAP as a measure of so called implicit cognition and as a tool for measuring relational framing "in-flight". Clearly, the types of rules that are presented to participants during an IRAP are not an insignificant variable that have little or no impact on performance. The instructions or rules appear to be quite impactful and indeed interact with other IRAP variables, such as the order in which the blocks are presented (history-consistent versus history-inconsistent) and possibly with prior exposures to other IRAPs. Of course, it is worth noting at this stage that the current research did not attempt to assess the impact of rules or instructions on the predictive validity of the IRAP effects (see Vahey, Nicholson, & Barnes-Holmes, 2015). For example, it has yet to be determined if strong IRAP effects have greater predicative validity than weaker IRAP effects, or indeed vice versa. Nevertheless, the fact that the current research has shown that rules do impact on IRAP performances highlights that future research will need to explore this area in a systematically if the IRAP, and its derivatives, are to be used with increasing precision and confidence over the coming years.

References

Barnes, D. (1989). Behavior-behavior analysis, human schedule performance, and radical behaviorism. *The Psychological Record*, 39, 339-350.

Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record*, 41, 19-32.

Barnes, D., Smeets, P. M., & Leader, G. (1996). New Procedures for establishing emergent matching performances in children and adults: Implications for stimulus equivalence. In T.R Zentall, & P. M. Smeets (Eds.), *Stimulus class formation in humans and* animals (pp. 153-171). Amsterdam: Elsevier.

Barnes-Holmes, D., Hayden, E., Barnes-Holmes, Y., & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study. *The Psychological Record*, 58(4), 497-516.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, 60(3), 527-542.

Bentall, R. P., Lowe, C. F., & Beasty, A. (1985). The role of verbal behavior in human learning: II Developmental differences. *Journal of the experimental Analysis of Behavior*. 43(2), 165-181.

Chiesa, M. (1994). *Radical behaviorism: The philosophy and the science*. Authors Cooperative.

Chiesa, M. (1992). Radical behaviorism and scientific frameworks: From mechanistic to relational accounts. *American psychologist*, 47(11), 1287-1299.

Catania, A. C., Shimoff, E., & Matthews, B. A. (1989). An experimental analysis of rule-governed behavior. In S.C. Hayes (Ed,), *Rule-governed behavior* (pp. 119-150). New York: Plenum Press.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. New York: Academic Press.

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(2), 163-184.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology,* 74(6), 1464-1480.

Hayes, S. C., & Hayes, L. J. (1989). The verbal action of the listener as a basis for rule-governance. In S.C. Hayes (Ed,), *Rule-governed behavior* (pp. 153-190). New York: Plenum Press.

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

Healy, O., Barnes-Holmes, D., & Smeets, P. M. (2000). Derived relational responding as generalized operant behavior. *Journal of the Experimental Analysis of Behavior*, 74(2), 207-227.

Hineline, P. N., & Wanchisen, B. A. (1989). Correlated hypothesizing and the distinction between contingency-shaped and rule-governed behavior. In S.C. Hayes (Ed.), *Rule-Governed Behavior* (pp. 221-268). New York: Plenum Press.

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science*, 1(1), 17-38.

Hussey, I., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). From Relation Frame Theory to implicit attitudes and back again: clarifying the link between RFT and IRAP research. *Current Opinion in Psychology,* 2, 11-15.

Lipkens, R., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of the development of derived relations in an infant. *Journal of Experimental Child Psychology*, 56(2), 201-239.

Lowe, C. F. (1979). Determinants of human operant behavior. *Advances in analysis of behavior*, 1, 159-192.

Lowe, C. F., Beasty, A., & Bentall, R. P. (1983). The role of verbal behavior in human learning: infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 39(1), 157-164.

Luciano, C., Becerra, I. G., & Valverde, M. R. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior*, 87(3), 349-365.

Nicholson, E., & Barnes-Holmes, D. (2012). Developing an implicit measure of disgust propensity and disgust sensitivity: Examining the role of implicit disgust propensity and sensitivity in obsessive-compulsive tendencies. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(3), 922-930.

O'Shea, B., Watson, D. G., & Brown, G. D. (2015). Measuring implicit attitudes: a positive framing bias flaw in the Implicit Relational Assessment Procedure (IRAP). *Psychological Assessment*. Advance online publication. http://dx.doi.org/10.1037/pas0000172

Pilgrim, C., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior*, 63(3), 225–238.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion (Zalophus californianus) is capable of forming equivalence relations. *Psychological Record*, 43(4), 823-839.

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

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of the Experimental Analysis of behavior*, 37(1), 5-22.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological review*, 57(4), 193-216.

Skinner, B. F. (1938). The behavior of organisms: An experimental analysis. New York: Appleton-Century-Crofts.

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

Skinner, B. F. (1966). An operant analysis of problem solving, (1966). In B. Kleinmuntz (Ed.), *Problem Solving: Research. Method, and Theory,* (pp. 225-257) New York: Wiley.

Smeets, P. M., Leader, G., & Barnes, D. (1997). Establishing stimulus classes in adults and children using a respondent-type training procedure: A follow-up study. *The Psychological Record*, 47(2), 285-309.

Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior*, 56(3), 519-555.

Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of behavior therapy and experimental psychiatry*, 48, 59-65.

Vaughan, M. (1989). Rule-governed behavior in behavior analysis. In S.C. Hayes (Ed.), *Rule-governed behavior* (pp. 97-118). New York: Plenum Press.

Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior*, 12(3), 349-373.

Appendices

Appendix A

INFORMED CONSENT FORM FOR STUDENT RESEARCH

In agreeing to participate in this research I understand the following:

This research is being conducted by **Martin Finn**, a postgraduate student at the Department of Psychology, Maynooth University (email; MARTIN.FINN.2011@nuim.ie; address: Postgraduate Room, Department of Psychology, John Hume Building, Maynooth University; contact number: 01 708 6086) The current study is investigating the effects of different kinds of instructions on the Implicit Relational Assessment Procedure.

- If I have any concerns about participation I understand that I may refuse to participate or withdraw at any stage. At the conclusion of my participation, any questions or concerns I have will be fully addressed.
- I have been informed as to the general nature of the study and agree voluntarily to participate.
- I will complete a questionnaire as well as a number of computer based pairing tasks where I will be asked to pair words.
- All data from the study will be treated confidentially. The data from all participants will be anonymised, compiled, analysed. The data will be retained for approximately 5 years before being destroyed. My data will not be identified by name at any stage of the data analysis or in the final report.
- It must be recognized that, in some circumstances, confidentiality of research data and records may be overridden by courts in the event of litigation or in the course of investigation by lawful authority. In such circumstances the University will take all reasonable steps within law to ensure that confidentiality is maintained to the greatest possible extent
- I may withdraw from this study at any time without giving a reason, and may withdraw my data at the conclusion of my participation if I still have concerns.
- I confirm that I am of 18 years of age or older.

Participant's signature ***Researcher's signature***

Participant name (print) ***Date***

Should you be in distress or experiencing any form of mental health complaint we encourage you to contact the NUIM student counseling service. This professional, free, and confidential counseling service can be contacted on 01-7083554 or counselling.nuim.ie to schedule an appointment.

If during your participation in this study you feel the information and guidelines that you were given have been neglected or disregarded in any way, or if you are unhappy about the process, please contact the Secretary of the Maynooth University Ethics Committee at research.ethics@nuim.ie or +353 (0)1 708 6019. Please be assured that your concerns will be dealt with in a sensitive manner.

If you should have any questions about the research being conducted please contact Professor Dermot Barnes-Holmes, SF09, Psychology department, Maynooth University, Dermot.Barnes-Holmes@nuim.ie

Appendix B

Demographics

Participant code

____

Gender

____

Age

In years

____

How many similar experiments have you participated in, where you were asked to pair words or images quickly according to a rule?

____

Rating Scale

How positive or negative do you find the following words?

	Very Negative	Negative	Somewhat Negative	Neutral	Somewhat Positive	Positive	Very Positive
	1	2	3	4	5	6	7
Red	1	2	3	4	5	6	7
Green	1	2	3	4	5	6	7
Blue	1	2	3	4	5	6	7
Circle	1	2	3	4	5	6	7
Square	1	2	3	4	5	6	7
Triangle	1	2	3	4	5	6	7