FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Empirical and conceptual analyses of perspective-taking: A behaviour-analytic approach

Deirdre Kavanagh

Supervisor: Prof. Dr. Dermot Barnes-Holmes
Co-supervisor: Prof. Dr. Yvonne Barnes-Holmes

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Doctor of Psychology

Academic year 2019–2020

GHENT
UNIVERSITY

Guidance Committee

Prof. Dr. Dermot Barnes-Holmes (Promotor)

Department of Experimental Clinical and Health Psychology, Ghent University

Prof. Dr. Yvonne Barnes-Holmes (Co-Promotor)

Department of Experimental Clinical and Health Psychology, Ghent University

Prof. Dr. Jan DeHouwer

Department of Experimental Clinical and Health Psychology, Ghent University

Prof. Dr. Bart Soenens

Department of Developmental, Personality and Social Psychology, Ghent University

Prof. Dr. Carmen Luciano

Department of Psychology, University of Almeria, Spain

Table of Contents

Acknowledgments

Firstly, thank you to Dermot and Yvonne for the oppotunity to complete my PhD as part of the Odyssus project. Thank-you for your patience and guidance throughout my PhD.

I would like to thank all the people who participated in this research without whom it would not have been possible.

Thank-you to the Odysseus team (past and present) and all the vistors to our lab for the geeky conversations that have made my PhD experience so rich and fufilling.

A special thanks to Martin and Aileen for all the chats and laughter that added so much to my Belgian adventure.

To my family and friends from near and far, thank-you for your never-ending support and belief in me.

Finally, thank-you Mark for keeping me smiling.

XXX

List of Tables

Table 2: Overview of the Multi-Dimensional Multi-Level (MDML) Framework

List of Figures

List of Appendices

Chapter 1

General Introduction [1]

[1] This chapter is based on Kavanagh, D., Barnes-Holmes, Y., & Barnes, Holmes, D. (accepted). The study of perspective-taking: Contributions from mainstream psychology and behavior analysis. *The Psychological Record.*

Perspective-taking has long been considered pivotal for human socialization (Mead; 1934; Piaget, 1948) in terms of enabling an individual to overcome early egocentrism and to adjust their behaviour according to the expectations of others. The ability to take another's perspective is crucial in: competitive settings (Galinsky & Mussweiler, 2001; Galinsky, Maddux, Gilin, & White, 2008); the establishment and maintenance of healthy interpersonal relations (Arriaga & Rusbult, 1998; Hughes, & Leekam, 2004); and strengthening social bonds (Galinksy & Ku, 2004; Vescio, Sechrist, & Paolucci, 2003). Indeed, there is evidence that perspective-taking deficits are associated with significant impairments in social skills (e.g. Tager-Flusberg & Anderson, 1991).

Perspective-taking has been broadly defined as the ability to interpret and predict the thoughts, emotions or behaviours of oneself and of others (Carpendale & Lewis, 2006) in terms of being able to 'assume an alternative perspective' where necessary. While most research on perspective-taking has investigated the ability to assume the perspective of *another* (e.g. Baron-Cohen, Tager-Flusberg, & Cohen, 2000), the metacognitive (self-based) aspect of perspective-taking has also been emphasised (for a review see Bernstein, Hadash, Lichtash, Tanay, Shepherd, & Fresco, 2015). When both self- and other-based aspects of perspective-taking are acknowledged, the key skill is often perceived as the ability to put oneself in the 'mental shoes' of others in terms of imagining how they perceive, think or feel about an object or event (Moll & Meltzoff, 2011a) and to understand that these beliefs/perceptions may be different from one's own (Sigman & Capps, 1997).

The key purpose of this introductory chapter is to provide a comprehensive overview of the literature on perspective-taking from within mainstream psychology and behaviour analysis. The primary focus will be on the behaviour-analytic approach to perspective-taking, which will be divided into what may be described as: (1) traditional behaviour analysis and (2) an area within behaviour analysis that is concerned specifically with human language and cognition, known as relational frame theory (RFT). The review begins by summarising the main ways in which traditional psychological research has studied perspective-taking.

The Mainstream Psychology Approach to Perspective-taking

Perspective-taking abilities have been sub-divided into three domains: *visual or spatial* perspective-taking (Flavell, 1992; Moll & Tomasello, 2006; Tversky & Hard, 2009); *affective or emotional* perspective-taking (Dunn & Hughes, 1998; Wellman, Phillips, & Rodriguez, 2000); and *cognitive* perspective-taking (Flavell, 2004; Gopnik & Slaughter, 1991; Wellman, Cross, & Watson, 2001). Although these domains are often studied independently, all three appear to be involved, to varying degrees, in the ability to take the perspective of another.

Visual perspective-taking. Visual perspective-taking is often also referred to as *spatial* perspective-taking, with the two terms used interchangeably. Flavell (1977) proposed two distinct levels in its development. *Level 1* involves appreciating what others see only from that specific viewpoint (e.g. recognizing that you may be able to see something that someone else cannot see). In a typical task assessing Level 1, a child is asked to position an object so that another person cannot see the object, or to determine whether another individual can see an object that the child can see. Competence on tasks of this kind has been observed in children as young as twelve-and-a-half-months (Luo & Baillargoen, 2007; Sodian, Thoermer, & Metz, 2007).

Level 2 of visual perspective-taking involves appreciating that even when two people can see the same object, they may do so from different vantage points. Typical tasks of this level are similar to Piaget's *Three Mountain Problem*, in which a child must choose the photograph that matches what another person sees, rather than what the child sees (Piaget & Inhelder, 1956). Competence on this type of task has been observed in children aged four years and older (Flavell, Everett, Croft, & Flavell, 1981).

The large gap in the ages at which competence in Levels 1 and 2 of visual perspective-taking has been observed (12 months versus four years) is generally consistent with other evidence of variations in age-based outcomes for this skill (see Frick, Möhring, & Newcombe, 2014 for a review), and may pertain to the use of different methodologies. For example, using a more child-friendly task than the Three Mountain Problem, Moll and Meltzoff (2011a) reported that children as young as three years old show Level 2 visual perspective-taking.

Affective perspective-taking. Affective perspective-taking is often also referred to as *emotional* perspective-taking, again with the two terms used interchangeably. Affective perspective-taking involves the ability to recognise the emotional state of someone else (particularly when this differs from the emotional state of oneself), and to understand the relationship between different situations and the specific emotions they typically elicit (Cutting & Dunn, 1999). For example, a child capable of affective perspective-taking can recognise that, while they could be happy after winning a race, other children in the same race may feel sad because they lost. Competence in this regard has been observed in children aged between two and three years old (Eisenberg, Spinrad, & Sadovsky, 2006), and is typically associated with the emergence of simple emotion-based words, such as "happy" and "sad" (Dunn, Bretherton, & Munn, 1987). Indeed, affective perspective-taking is believed to be associated with social and emotional functioning (Arsenio, 2003; Izard et al., 2001), especially the ability to show varied and empathic responses to the distress of others (Eisenberg et al., 2006). For example, Wellman et al. (2000) have shown that two to three year olds can determine whether someone has received a desirable or undesirable object based on their emotional reactions to that object.

It is perhaps not surprising that the conceptual and empirical work conducted to date on affective perspective-taking has highlighted the difficulties in separating this ability from related skills, such as empathy (e.g. Farrant, Devine, Maybery, & Fletcher, 2012) and having a theory of mind (e.g. Cutting & Dunn, 1999). Indeed, many affective perspective-taking tasks require relatively sophisticated cognitive and linguistic skills, and thus result in floor effects when presented to young children (Eisenberg et al., 2006). In addition, affective perspective-taking tasks often require an understanding of conflict between one's own emotional response to a situation and the emotional response of another (Harwood & Farrar, 2006), which is hard to distinguish from an understanding of false belief. It is perhaps, in part, for this reason that some authors have proposed an overlap between affective and cognitive perspective-taking as described below (e.g. Harwood & Farrar, 2006). On balance, other authors maintain that these complex

repertoires are better understood as two distinct forms of social cognition (e.g. Cutting & Dunn).

Cognitive perspective-taking. Almost all of the research on cognitive perspective-taking has been conducted under the rubric of *theory of mind (ToM),* with the core skill believed to involve the correct attribution of mental states to oneself and others as a means of explaining and predicting behaviour (Baron-Cohen, et al., 2000; Harrington, Siegert, & McClure, 2005). While the term ToM was first coined by Premack and Woodruff (1978) in research with chimpanzees, the ability to form a theory of mind is believed to be universal only in human adults (see Call & Tomasello, 2008, for a review).

Cognitive perspective-taking and ToM skills are typically divided into two levels (Baron-Cohen, 2001). *First-order* false beliefs refer to assumptions made about another person's beliefs and *second-order* false beliefs refer to another person's assumptions about beliefs held by a third party (Boucher, 2012), with the latter deemed to be the more complex. Consider the first-order false belief scenario presented in the *Sally-Anne Test*, in which a child is asked about a protagonist known to hold a false belief (e.g. that there is a cookie in the cookie jar, rather than the jar containing nuts) about a situation because this belief differs from the child's own true belief (e.g. that the cookie jar contains nuts). Wellman et al. (2001) reported that children aged four years typically pass the Sally-Anne Test, thus showing first-order false belief attribution. Second-order false beliefs are typically assessed through change in location stories and determine a child's understanding that someone can hold a false belief about someone else's belief (e.g. Flobbe, Verbrugge, Hendricks, & Krämer, 2008). Consider the following scenario from the *Unexpected Transfer Task*. The child is presented with the following scenario: 'Mary and Simon are given a chocolate bar to share. Both put the chocolate in the fridge before going out to play. Soon after, Simon returns to the kitchen and takes the chocolate out of the fridge and puts it in his bag. Later, both children are told that they can eat their chocolate bar.' Then the child is asked: "Where does Simon believe that Mary believes that the chocolate is?" The correct response 'Simon believes that Mary believes that the chocolate is in the fridge' indicates the attribution of a

second-order belief (see Sullivan, Zaitchik, & Tager-Flusberg, 1994). Astington, Pelletier, and Homer (2002) demonstrated competence in second-order false belief attribution in children aged six-seven years.

Relatively recent research has focused less on ToM directly in terms of when it is present in a child's repertoire and more on the developmental pre-cursors that are necessary for ToM to emerge in the first place. For example, Brooks and Meltzoff (2015) reported that infant joint attention (i.e. tracking another's eye gaze or finger pointing in order to co-ordinate attending to a stimulus, such that the learner and the instructor have some element of shared experience regarding that stimulus) predicts mental-state term usage (e.g. using phrases such as "she knows" and "he thinks") at two years and competence in ToM at four-and-a-half years (see also Colonnesi, Rieffe, Koops, & Perucchini, 2008; Kristen, Sodian, Thoermer, & Perst, 2011). Declarative pointing (i.e. pointing that functions as a means for a child to achieve joint reference with the addressee) in infancy also predicts ToM in preschool activities (Sodian & Kristen-Antonow, 2015).

Cognitive perspective-taking in atypical development. Similar to the research noted above, a small number of studies have examined the impact of congenital deafness on the subsequent development of ToM. For example, Schick, de Villers, de Villers and Hoffmeister (2007) found that hearing-impaired infants, born to hearing-abled parents, show delays in passing false belief tasks at the typical age. Most of the ToM literature has explored other types of developmental deficits in atypical populations (e.g. Broekhof et al., 2015), including: attention deficit hyperactivity disorder (ADHD, Uekermann et al., 2010) and autism spectrum disorder (ASD, Frith, Morton, & Leslie, 1991). In particular, research on the potential deficits in ToM in individuals with ASD has led to the "*Impaired ToM" Hypothesis* (see Baron-Cohen, Leslie, & Frith, 1985). This suggestion is based on evidence that at least some children with ASD fail to shift from their own perspective to the perspective of another (e.g. Baron-Cohen, 1989), and that some older children and even adolescents from this group fail ToM tasks that can be passed by typically-developing children as young as four years (see Peterson, Wellman, & Liu, 2005).

In contrast, there exists a considerable body of evidence that questions the view that ToM is typically deficient in individuals with ASD. For example, Boucher (2012) reviewed a number of studies in which individuals from this group passed both first- and second-order false belief tasks, as well as other types of assessments of ToM, involving metaphor, faux pas and sarcasm (e.g. Scheeren, de Rosnay, Koot, & Begeer, 2013). In support of these outcomes, Happé (1995) suggested that verbally-competent individuals with ASD *can* pass false belief tasks when given ample time because their core deficits may lie more broadly in social affective information processing (see also Tager-Flusberg, 2007).

In the ToM literature, it is widely accepted that ToM skills emerge in tandem with biological maturation (e.g. Grosse Wiesman, Schreiber, Singer, Steinbeis, & Firederici, 2017). Some support for this view comes from evidence that older children and adults generally show greater speed and flexibility in perspective-taking than younger children (Apperly, Back, Samson, & France, 2008; Im-Bolter, Agostino, & Owens-Jaffray, 2016). On balance, competence in ToM in children is susceptible to training and thus may be influenced as much by specific developmental contexts as broad maturational changes (see Hoffman et al., 2016).

In any case, the relationship between perspective-taking and maturation does not appear entirely one-dimensional, because some evidence shows that under certain conditions even adults show slower responding and make more errors on ToM tasks than would generally be expected for their maturity (Epley, Caruso, & Bazerman, 2006). The findings in this regard may be summarised as follows. (1) Attributing beliefs to others is slower than attributing beliefs to oneself (Bradford, Jentzsch, & Gomez, 2015). (2) Switching from self-perspective to other-perspective is slower than switching from other-perspective to self-perspective in the context of false belief (Bradford et al., 2015). (3) More complex tasks produce more errors (Bull, Phillips, & Conway, 2008) and non-mentalistic, reality-based tasks are completed more quickly (Apperly, Riggs, Simpson, Chiavarino, & Samson, 2006). (4) Greater demands on cognitive load generate more errors and longer reaction times (McKinnon & Moscovitch, 2007). (5) Tasks requiring inhibitory control of self-perspective take longer (than non-inhibitory control tasks,

Birch & Bloom, 2007; Keysar, Lin, & Barr, 2003). Taken together, numerous authors have argued that these differential outcomes highlight the role of executive functioning in ToM, and suggest that at least some of the weak ToM performances observed in adults and atypical samples reflect broader deficits in executive functioning, rather than in ToM per se (e.g. German & Hehman, 2006; McKinnon & Moscovitch). Other researchers working under the rubric of relational complexity (RC) theory have proposed that variability in outcomes on perspective-taking and ToM tasks in typical and atypical individuals depends upon the level of relational complexity that is comprised of the number of variables that are to be related in a given task (Andrews, Halford, Bunch, Bowden, & Jones, 2003; Halford, Wilson & Phillips, 1998).

Cognitive perspective-taking in psychological suffering (psychopathology). According to Vaskin et al. (2015), different error patterns in ToM tasks may help to distinguish clinical from non-clinical populations, especially those diagnosed with Borderline Personality Disorder (BPD) or schizophrenia. Specifically, given that disturbances in interpersonal relationships and the misreading of the intentions of others are believed to be core characteristics of BPD, it is perhaps not surprising that some individuals have shown different ToM outcomes, relative to controls (see Németh et al., 2018 for a review). However, the findings in this regard are again inconsistent, with some studies showing that individuals with BPD perform *better* than control participants (e.g. Fertuck et al., 2009) and potentially highlighting 'hypervigilance' to social cues, but others showing that individuals with BPD perform w*orse* than controls (e.g. Sharp et al., 2011), thereby suggesting ToM deficits. Similar to the developmental evidence discussed above, the ToM competencies of this clinical group appear to be influenced by the type of ToM task employed. For example, Preißler, Dziobek, Ritter, Heekeren and Roepke (2010) found that individuals with BPD were more likely to make ToM-based errors on ecologically valid tasks with greater complexity than on simpler tasks. Indeed, Roepke, Vater, Preißler, Heekeren and Dziobek (2013) argued that the more complex tasks may be necessary to tease out ToM deficits in this group.

Several studies have reported that individuals with schizophrenia have also produced different outcomes on ToM tasks, relative to controls (e.g. Savla, Vella,

Armstrong, Penn, & Twamley, 2013). That is, some have found weaker performances in the attribution of thoughts or intentions to others, and particularly in the attribution of emotional states (e.g. Shamay-Tsoory et al., 2007), and others have found correlations between negative symptoms and the attribution of overly simplistic mental states to others (Montag et al., 2011; Pickup & Frith, 2001). Positive symptoms have also been associated with the attribution of overly complex mental states to others (Fretland et al., 2015). However, as noted previously, these performances are also influenced by task-specific features, suggesting that they may not be reflecting ToM deficits per se in this population. For example, on the *Reading the Mind in the Eye Test*, Scherzer, Leveille, Achim, Boisseau and Stip (2012) found no significant difference in ToM performances between individuals with schizophrenia and a group of non-clinical controls. However, performance on the *Hinting Task* was significantly more impaired for individuals with schizophrenia compared to the non-clinical group. Once again, generic deficits in executive function may be at play, in that numerous studies have found a correlation between poor ToM and deficits in inhibition and cognitive flexibility in this sample (for a review see Pickup, 2008).

The mainstream psychology approach to the potential 'processes' involved in perspective-taking. The reviews above of the developmental and clinical literatures on perspective-taking and ToM highlight the difficulty of ruling out the potential role of broader cognitive concepts and capabilities, of which perspective-taking may be a component. As a result, the field of perspective-taking appears to have struggled to formulate an operational definition of what these skills involve (Davis et al., 2004). Davis et al. summaried the core steps involved in perspective-taking as follows: (1) imagining what the observer themselves would do in that position; (2) imagining similar circumstances from the observer's own past; (3) imagining what the other person would do based on the observer's knowledge of the other person's history; and (4) following general social rules about what others might typically do in that type of scenario.

In a similar vein, Epley and Caruso (2009) proposed three possible steps in perspective-taking that highlight three points at which an individual may falter. (1) The first step involves successfully activating the ability to perspective-take. That is, a

person may fail to identify an instance where another's perspective should be considered and this may result from factors such as the absence of sufficient effort or training (e.g. Idson et al., 2004). (2) The second step involves adjusting one's own egocentric perspective to accommodate that of another, with failure to do so rendering judgements biased in the direction of the initial self-perspective (for reviews, see Epley, 2004; Keysar & Barr, 2002). Indeed, Davis et al. (2004) found that explicit instructions to adopt another's perspective increase the accessibility of self-related thoughts. Furthermore, egocentric biases increase when individuals are asked to respond quickly, but decrease when financial incentives accompany correct responding (Epley, Keysar, Van Boven, & Gilovich, 2004). (3) The third step in perspective-taking depends upon the information about others that is provided, which ranges from idiosyncratic knowledge about specific individuals to general information and even stereotypes (Ames, 2004).

Summary. In mainstream psychology, the broad concept of perspective-taking has been sub-divided into visual perspective-taking, emotional perspective-taking and cognitive perspective-taking (usually referred to as ToM). Visual perspective-taking has been observed as early as 12 months. Affective perspective-taking emerges later, usually around age two, and cognitive perspective-taking (ToM) comes even later around age four, with false belief understanding observed around age six. However, there have been wide variations in the ages at which these skills are typically first observed, with outcomes apparently sensitive to both methodological variations and broader individual differences, especially in executive functioning. There is evidence of weaker perspective-taking performances in samples with a diagnosis of ASD, BPD and schizophrenia, but these too appear to be influenced by task-specific features and broader executive functioning abilities. Indeed, the reviews above of the developmental and clinical literatures highlight the difficulty of ruling out the potential role of broader cognitive concepts and capabilities in assessments of perspective-taking. These difficulties perhaps account for why mainstream psychology has not yet reached consensus on the processes involved in perspective-taking. The current chapter continues by summarising the ways in which traditional behaviour analysis has studied

perspective-taking, and, as will become clear, the basic (behavioural) processes in perspective-taking are only now beginning to attract attention in that field.

The Traditional Behaviour-analytic Approach to Perspective-taking

In the field of behaviour analysis, the ability to respond to one's own responding is pivotal to an understanding of the concept of 'self'. The earliest behavioural writings on self include those by Skinner (e.g. 1974), who proposed that self-knowledge develops through shaping by the knowledge of others and by social contingencies that reinforce the discrimination of one's own behaviour. By asking questions such as "How are you feeling?", other members of the verbal community, in effect, shape an individual's ability to discriminate their own behaviour. Across such exemplars, an individual is believed to become more self-aware, thus acquiring better prediction and control over their own behaviour (see Skinner). However, it is perhaps surprising that the empirical support typically cited for Skinner's interpretation of self-awareness has come from studies with non-humans, such as Lattal's (1975) demonstration that the behaviour of pigeons may be brought under the control of their own previous patterns of responding.

Few in the behaviour-analytic community have explicitly attempted to connect the Skinnerian concept of self with the skills involved in perspective-taking, even though they appear intuitively linked. However, some researchers have attempted to interpret ToM tasks and performances using behavioural concepts. For example, Spradlin and Brady (2008) interpreted ToM performances in terms of Sidman's (e.g. 1971) equivalence relations. Similarly, Okuda and Inoue (2000) and Schlinger (2009) attempted to interpret these performances in terms of operant stimulus control. And, DeBernardis, Hayes and Fryling (2014) offered an interbehavioural interpretation that emphasises the importance of analysing the complex interbehavioural history between the perceiver and the target other.

Most of the behavioural work on perspective-taking has focused on the remediation of deficits in these abilities, based on the broad assumption that perspective-taking repertoires may be established or remediated through the acquisition of an appropriate learning history. Two studies have investigated the use of video modelling

interventions to establish perspective-taking in children with ASD. Specifically, LeBlanc et al. (2003) presented three children with a variation of the Sally-Anne Test. Whilst all three initially failed, the video modelling served to produce highly competent performances in all cases, although generalisation to untrained tasks was recorded for only two of the children. Using a similar procedure again with three children, but including a number of training exemplars, Charlop-Christy and Daneshvar (2003) produced similar outcomes, but were more successful at promoting generalisation to novel stimuli and novel responses. In a related study also using a variation of the Sally-Anne Test, and including prompts for training purposes, Gómez-Becerra, Martín, Chávez-Brown and Greer (2007) presented the task to five children with ASD, five with Down Syndrome (DS) and five who were typically-developing, all aged between four and six years. Three of the typically-developing children passed without prompts, with the remaining two requiring prompts to pass, as did all five children with DS. The five children with ASD all failed even with prompts. Further analyses of the data indicated that only those children with ASD or DS who had language deficits produced weak performances before or after prompting.

Developmental behavioural researchers have devoted considerable attention to the likely pre-cursors to perspective-taking (Novak, 1996, 1998; Novak & Pelaez, 2004), with a particular focus on joint attention and social referencing (Moll & Meltzoff, 2011b; Slaughter & McConnell, 2003). *Joint attention,* as the name implies, involves tracking another's eye gaze or finger pointing in order to co-ordinate attending to a stimulus, such that the learner and the instructor have some element of shared experience regarding that stimulus (see also Lowenkron, 1998, for a relevant functional-analytic description of joint control). According to Dube, MacDonald, Mansfield, Holcomb and Ahearn (2004), this joint stimulus orientation results from a relevant history of consequences, including solidarity play. Indeed, there is evidence that joint attention can be established when found to be deficient or absent. For example, MacDonald et al. (2006) investigated joint attention responding and initiation in 21 typically-developing children and 26 children with ASD, aged two to four years. When required to respond to joint attention involving gestures, both groups performed well,

with some superiority observed with advanced age. However, when required to *initiate* joint attention, the children with ASD, especially the younger ones, showed considerably weaker performances than the typically-developing children. In a related remediation study of joint attention and its initiation, Whalen and Schreibman (2003) trained five children with ASD aged four years, using components of Discrete Trial Training (DTT) and Pivotal Response Training (PRT). Baseline performances indicated considerable impairments relative to typically-developing peers, especially in joint attention initiation. However, training facilitated significant improvements in all children on joint attention, and in four of the children on joint attention initiation, including generaliation to novel settings and novel adults. In a similar study, Gould, Tarbox, O'Hora, Noone and Bergstrom (2011) evaluated the use of multiple exemplars of Conditional Discrimination Training (CDT). During the baseline, all three children failed to demonstrate joint attention, but the interventions facilitated rapid acquisition of the target performances and generalisation to untrained stimuli for all participants. However, generalisation in the natural environment was much more limited. Hahs (2015) replicated this study, but failed to find generalisation to untrained stimuli.

Social referencing involves orienting to another person's expression and then responding to a stimulus on the basis of that expression (Peláez-Nogueras & Gewirtz, 1997), with the expression thus functioning as a setting event (Peláez, 2009). For example, if a child discriminates a fearful expression on their mother's face as the child reaches towards a dog, the child may be less likely to touch the dog given this expression. Social referencing, therefore, enables learners to predict the potential reinforcement of stimuli or events without the need for direct contact with the stimulus. Peláez, Virues-Ortega and Gewirtz (2012) investigated social referencing in which maternal facial expressions signaled either positive or negative consequences of the reaching behavior of 11 four- and five-month old babies. Whilst all of the infants failed the baseline, subsequent interventions established the mothers' joyful expressions as discriminative for infant reaching, and fearful expressions as discriminative for not reacting, for all children.

Summary. In traditional behaviour analysis, the ability to respond to one's own responding (i.e. acquire self-awareness) as an essential pre-cursor to perspective-taking, is shaped through a history of interacting with other members of the social/verbal community, although empirical support for this interpretation of self-awareness has tended to come from research with non-humans. Indeed, little behaviour-analytic research has explicitly attempted to connect the concept of self (or self-awareness) with perspective-taking skills in humans. While some researchers have offered behavioural interpretations of ToM performances in terms of equivalence relations and operant stimulus control, most research has focused on remediating deficits in perspective-taking through the provision of appropriate learning histories. The outcomes show that video modelling interventions have enabled children with ASD to pass the Sally-Anne Test, although generalisation is more robust when training exemplars are included. Incorporating training prompts into the Sally-Anne Test has also produced positive outcomes, but prompts appear to be necessary or less effective for participants with language deficits. Developmental behavioural researchers have devoted considerable attention to joint attention and social referencing as pre-cursors to perspective-taking. There is evidence that joint attention responding and initiation can be established in children with ASD, although the latter is more likely to be deficient and more difficult to establish, and generaliation is not always observed. Interventions to establish social referencing have demonstrated positive outcomes in four- and five-month old babies. In general, as one might expect, there has been a far greater emphasis in behavioural psychology on prediction-and-influence of perspective-taking, which in a broad sense further complicates the literature in this area. That is, research from the mainstream literature has tended to produce relatively inconsistent results in attempting to tie deficits in perspective-taking to specific ages, populations and syndromes. The fact that behaviour-analytic research also shows that perspective-taking deficits are relatively amenable to change when targeted by behavioural interventions makes it even more difficult to draw firm conclusions about the very concept of perspective-taking itself. In turning to an area of behaviour analysis that has concerned itself specifically with human language and cognition, RFT, we will find an account of perspective-taking that

is unsurprisingly rooted in human language. As will become clear, this focus has recently generated a more process-oriented account of perspective-taking.

Relational Frame Theory: A Language-focused Behaviour-analytic Approach to Self and Perspective-taking

Some behavioural researchers working under the rubric RFT (see Hayes, Barnes-Holmes, & Roche, 2001 for a book-length treatment) have proposed that self-discrimination involves *verbal* processes that distinguish it functionally from the non-verbal self-discrimination observed with non-humans. In other words, self-awareness requires a human to be "not simply behaving with regard to his behavior, but . . . also behaving verbally with regard to his behavior" (Hayes & Wilson, 1993, p. 297). According to RFT, this type of verbal self-discrimination and perspective-taking comprise repertoires of *derived relational responding*, which is the basis of language itself. This section of the chapter summarises the core concepts of RFT in order to provide the basis for the theory's approach to perspective-taking as derived relational responding.

Patterns of derived relational responding. At its most basic, RFT makes an important distinction between non-arbitrary and arbitrary relational responding. In simple terms, *non-arbitrary* relational responding involves relating one stimulus or event to another on the basis of a shared physical property. For example, you might say that two tennis balls are the same because they are the same shape, size an/or colour, although there may be other small physical properties on which the two balls differ. Non-arbitrary relational responding appears to be directly acquired through contingencies and is highly developed in non-humans (see Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001).

In contrast, arbitrary or derived relational responding is *not* based solely on physical stimulus properties and is more likely to be emergent (i.e. derived) in terms of its acquisition. For example, if you train a verbally-able child that 'Tom is faster than David and David is faster than Ann', with no direct contact with these stimuli, they can then derive, in the absence of reinforcement or prompting, that 'Tom is faster than Ann' and that 'Ann is slower than Tom'. The technical term RFT uses to describe this type

of relational behaviour is *arbitrarily applicable relational responding* (AARR) and it appears, at the present time, to be largely unique to verbally sophisticated humans (Brino, Campos, Galvão, & McIlvane, 2014; but see also Hughes & Barnes-Holmes, 2014). RFT researchers have investigated a number of different patterns of AARR, including responding in accordance with relations of co-ordination, distinction, opposition, comparison, hierarchy and perspective-taking. These are summarised briefly below. In each case, an example of at least one study that has demonstrated the specific pattern of relational responding is provided.

Responding in accordance with the relation of *coordination* appears to be the most basic form of AARR that infants learn (Lipkens, Hayes, & Hayes, 1993) around 18-24 months (Luciano, Gómez-Becerra, & Rodríguez-Valverde, 2007). Consider the example, 'If A is the same as B and B is the same as C, then A and C are most likely the same.' O'Connor, Rafferty, Barnes-Holmes and Barnes-Holmes (2009) successfully employed multiple exemplar training (MET) to establish word-picture and picture-word co-ordination relations in 15 children with ASD, as well as three typically-developing children (see also Carr, Wilkinson, Blackman, & McIlvane, 2000).

Responding in accordance with the relation of *distinction* requires responding to arbitrary differences among stimuli, along a particular dimension, by applying a relational cue such as 'is different from' (Dixon & Zlomke, 2005; Roche & Barnes, 1996; Steele & Hayes, 1991). Consider the example, 'If A is different from B, then B is different from A.' Relations of distinction do not always specify the relevant dimension along which the stimuli differ and, of course, there are many ways in which this can occur. For example, you might tell someone that you are very different from your sister, with no need to say exactly how you differ. Dunne, Foody, Barnes-Holmes, Barnes-Holmes and Murphy (2014) established contextual control for distinction responses with both non-arbitrary and arbitrary relations in two children with ASD.

Responding in accordance with the relation of *opposition* requires the abstraction of a particular dimension along which stimuli can be differentiated at either end of a continuum (Steele & Hayes, 1991). As a result, opposition relations likely involve a higher level of complexity than co-ordination and distinction relations, for

example, because opposition relations *involve* co-ordination and/or distinction relations. For example, 'If A is opposite to B (hence A and B are also different) and B is opposite to C', A and C are most likely the same. Barnes-Holmes, Barnes-Holmes, Smeets, Strand and Friman (2004) successfully employed MET to establish opposition relations in typically-developing children, while Dunne et al. (2014) established these relations in four children with ASD. Dymond, Roche, Forsyth, Whelan and Rhoden (2007) also demonstrated the derived transformation of avoidance functions in adults in accordance with opposition relations (see also Whelan & Barnes-Holmes, 2004).

Responding in accordance with the relation of *comparison* requires responding to one event in terms of quantitative or qualitative relations along a specified dimension with another event. For example, 'If A is bigger than B and B is bigger than C', A is bigger than C and C is smaller than A. Responding on the basis of comparison relations has successfully been established in typically-developing children (Barnes-Holmes et al., 2004; Berens, & Hayes, 2007; Hayes, Stewart, & McElwee, 2016) and in children with ASD (Dunne et al., 2014; Gorham, Barnes-Holmes, Barnes-Holmes, & Berens, 2009). Vitale, Barnes-Holmes, Barnes-Holmes and Campbell (2008) also showed some variation in comparison responding in adults between specified and unspecified relations.

Responding in accordance with *temporal* relations requires responding to the relationship between two events in terms of a specified temporal dimension, by applying a relational cue such as 'before/after' or 'now/then. For example, 'If A occurs after B and B occurs after C', C most likely occurs before A and A occurs after C. RFT-based research demonstrating patterns of temporal relational responding is limited, with only a few studies conducted with adults. Specifically, O'Hora, et al. (2008) found that successful completion of a temporal relations task predicted better performances on the Verbal Comprehension and Perceptual Organization indices of the Wechsler Abbreviated Scale of Intelligence-III (WAIS III) in an undergraduate sample. Similarly, O'Toole and Barnes-Holmes (2009) assessed flexibility in responding to temporal and comparison relations, using the Implicit Relational Assessment Procedure (IRAP). The results indicated that faster responding in accordance with temporal and comparison

relations, as well as greater flexibility in these patterns, predicted higher scores on the Kaufman Brief Intelligence Test (K-BIT). Several studies have also assessed the implications of temporal relations for instructional control (Brassil, Hyland, O'Hora, & Stewart, 2019; O'Hora, Barnes-Holmes, Roche, & Smeets, 2004; Hyland, Smyth, O'Hora, & Leslie, 2014; McGreal, Hyland, O'Hora, & Hogan, 2016).

Responding in accordance with the relation of *hierarchy* appears to be even more complex and again contains some of the relations described above. For example, 'If B is a member of group A'; A is a class containing B and any other members of A are likely to be similar to B, at least to some extent. For example, apples and oranges are both members of the food group fruit, but they differ in many other ways. Several studies have investigated hierarchical relations in adults (e.g. Gil, Luciano, Ruiz, & Valdivia-Salas, 2012; Gil, Luciano, Ruiz, & Valdivia-Salas, 2014; Griffee & Dougher, 2002; Slattery & Stewart, 2014) and typically-developing children (Mulhern, Stewart, & McElwee, 2017). Some studies have also successfully trained hierarchical relational responding in typically-developing children aged five-six years (Mulhern, Stewart, & McElwee, 2018).

A small number of studies have explored the possible sequence in which the above repertoires of AARR develop naturally, because this may have important developmental and educational implications (e.g. Cassidy, Roche, & Hayes, 2011; Dixon, 2014). For example, Dunne et al. (2014) assessed the repertoires in the following sequence with children with ASD: co-ordination, opposition, distinction and comparison. Their outcomes showed that all 10 children demonstrated co-ordination relations: four demonstrated opposition relations; and two demonstrated distinction, comparison and hierarchical relations, thus suggesting weaker performances as the relations became more complex. In addition, the number of training trials needed during the intervention phase to establish the target repertoires decreased steadily as more repertoires were established, thus implying that the earlier relational skills facilitated the latter.

In a subsequent study, Kent, Galvin, Barnes-Holmes, Murphy and Barnes-Holmes (2017) directly compared two training sequences. Training Sequence A

consisted of teaching co-ordination, distinction, comparison, opposition and then hierarchical relations, while Training Sequence B switched the order of the comparison and opposition relations (i.e. co-ordination, distinction, opposition, comparison and then hierarchy). The results indicated that participants who completed Training Sequence B (comparison before opposition) demonstrated significantly better performances on comparison relations than participants who completed Training Sequence A. This finding suggested that establishing opposition relations may facilitate the emergence of comparison relations.

The RFT approach to perspective-taking. As some of the findings above suggest, once these core patterns of AARR emerge, they likely provide the basis for more complex relational repertoires, such as that involved in perspective-taking. For RFT, perspective-taking is also AARR that becomes abstracted through learning to talk about your perspective in relation to others (McHugh, Barnes-Holmes, & Barnes-Holmes, 2004). Across multiple exemplars, this abstraction generates the constancy that characterises your perspective and once the perspective-taking relations are established, they become an intrinsic feature of almost all of our verbal behaviour (see Hayes et al., 2001). Imagine a very young child who is asked "What did you have for lunch today?" while they eating an evening meal with their family. If the child responded simply by referring to what a sibling is currently having for dinner, they may well be corrected with "No, that's what your brother is eating now, but what did you eat earlier today?" In effect, this kind of on-going refinement of the three deictic relations allows the child to respond appropriately to questions about their own behaviour in relation to others, as it occurs in specific times and specific places (e.g. McHugh et al., 2004). Thus, 'having a perspective' is a continuous experience and an individual is always operating from the same 'self' perspective (Hayes, 1984).

For RFT, the core relations involved in perspective-taking are referred to as *deictic* (Hayes et al., 2001), and include responding from one's own perspective in relation to others, time and place. Specifically, the *interpersonal* relations involve responding to *I* and *you,* the *spatial* relations involve responding to *here* and *there,* and the *temporal* relations involve responding to *now* and *then.* For RFT, the relational

properties of I vsersus you, here versus there, and now versus then become constants, against which environments that are continually changing in terms of time and space can be understood, categorised and communicated about.

The original deictic relations protocol. Most of the empirical research on deictic relational responding has employed various iterations of a developmental protocol originally developed by Barnes-Holmes (2001)[2]. The original extensive 256-trial protocol targeted the three deictic relations (I-YOU, HERE-THERE and NOW-THEN), as well as three levels of relational complexity, referred to as: simple, reversed, and double reversed relations. In an attempt to reflect a typical developmental sequence, the protocol targeted the interpersonal I-YOU relations first, followed by the spatial HERE-THERE relations and finally the temporal NOW-THEN relations. Specifically, *Level 1* first targeted *simple I-YOU relations* (e.g. "I have a red brick and you have a green brick. Which brick do I have? Which brick do you have?"), followed by *reversed I-YOU relations* (e.g. "If I have a red brick and you have a green brick and if I was you and you were me. Which brick would I have? Which brick would you have?").

Level 2 targeted HERE-THERE relations, including I-YOU relations from Level 1. Again, *simple HERE-THERE relations* were assessed first (e.g. "I am sitting here on the blue chair and you are sitting there on the black chair. Where am I sitting? Where are you sitting?"), followed by *reversed HERE-THERE relations* (e.g. "I am sitting here on the blue chair and you are sitting there on the black chair. If here was

[2] Guinther (2017) offered an experimental protocol for modelling perspective-taking that appears to require derived 'mental rotation' on behalf of typically-developing adult participants. More recently, Guinther (2018) extended the model to include false belief, but with mixed results (i.e. only two of four participants demonstrated false belief in the absence of direct training). As an aside, initially Guinther (2017) argued that the protocol developed by Barnes-Holmes (2001) and the derived mental rotation model constituted competing accounts of perspective-taking. More recently, Guinther (2018) appears to recognise that the two approaches are RFT-consistent, but are different in so many ways that it is difficult to argue that one is somehow more precise or better than the other.

there and there was here. Where would you be sitting? Where would I be sitting?"). Level 2 then combined the interpersonal and spatial relations in what was referred to as an *I-YOU/HERE-THERE double reversal* (e.g. "I am sitting here on the blue chair and you are sitting there on the black chair. If I was you and you were me and if here was there and there was here. Where would I be sitting? Where would you be sitting?").

Level 3 focused on the temporal relations and its relationship with interpersonal and spatial relations. Again, *simple NOW-THEN relations* were targeted first (e.g. "Yesterday I was watching television, today I am reading. What am I doing now? What was I doing then?"), followed by *reversed NOW-THEN relations* (e.g. "Yesterday I was watching television, today I am reading. If now was then and then was now. What was I doing then? What would I be doing now?"). It is notable from the examples above that even when presented in simple form, temporal relations do not combine I and YOU, instead only one is presented in any trial. This is because combining interpersonal and temporal relations leaves some relations unspecified. For example, if I tell you that 'I was sleeping yesterday and my sister is working today', you cannot know what my sister was doing yesterday and what I am doing today. Similar to Level 2, Level 3 also assessed *HERE-THERE/NOW-THEN double reversals* (e.g. "Yesterday I was sitting there on the blue chair, today I am sitting here on the black chair. If here was there and there was here and if now was then and then was now; Where would I be sitting then? Where would I be sitting now?").

Empirical research using the deictic relations protocol with typically-developing children. A large proportion of the research using the deictic relations protocol has presented it to typically-developing children (see Montoya-Rodríguez, Molina, & McHugh, 2017, for a review). The results of this body of research may be summarised as follows: (1) The data support the distinctions among the three types of deictic relations (McHugh et al., 2004); (2) The deictic relations vary on a continuum of complexity from simple relations to reversed relations and double reversed relations (McHugh et al.; Heagle & Rehfeldt, 2006); (3) There appears to be a developmental trend in which the interpersonal and simple relations emerge first (McHugh et al.); (4) Once established in typically-developing children, these perspective-taking repertoires

can generalise to both new stimuli and real-world conversational topics (Heagle & Rehfeldt); (5) Perspective-taking repertories can be successfully established in more natural language-like contexts, such as within a children's story (Davlin, Rehfeldt, & Lovett, 2011); and (6) The establishment of the deictic relations may be enhanced when multiple exemplars of established cues for deictic responding are incorporated into the protocol (Montoya-Rodríguez, & Molina Cobos, 2018).

Empirical research using the deictic relations protocol with atypically-developing children. Similar to the mainstream ToM literature, a considerable number of studies on deictic relational responding in children have focused on investigating possible deficits associated with ASD. The findings that have been observed using the deictic relations protocol may be summarised as follows: (1) Overall, children with ASD produce weaker performances than their typically-developing peers (Rehfeldt, Dillen, Ziomek, & Kowalchuk, 2007); (2) The performances of children with ASD can also be differentiated in terms of both relation type and level of complexity, and deficits can be remediated with direct training (Jackson, Mendoza, & Adams, 2014); (3) Deictic relational responding can be established using naturalistic variations of the original protocol, such as children's stories (Gilroy, Lorah, Dodgea, & Fiorello, 2015); and (4) The transformation of stimulus functions through deictic relations has been demonstrated in certain training contexts (Barron, Verkuylen, Belisle, Paliliunas, & Dixon, 2018; Belisle, Dixon, Stanley, Munoz, & Daar, 2016).

The use of the deictic protocol with typical adults. Several studies have used the original protocol to examine deictic relational responding in typical adults. McHugh et al. (2004) reported that adults (18-30 years) produced less errors overall than adolescents and children, and that adolescents made less errors than children. Interestingly however, even adults produced error rates ranging from 25% on reversed I-YOU relations to 50% on reversed NOW-THEN relations. In a replication study in the same paper, in which the protocol was presented in an automated rather than table-top format, very similar performances overall were observed.

The use of the deictic protocol with atypical adults. A number of studies have investigated deictic relational responding in various groups of adults, that may be

referred to as atypical, either in terms of intellectual competence or clinical presentation. Gore, Barnes-Holmes and Murphy (2010) investigated 24 adults with varying levels of intellectual disability and found that performances on the deictic protocol correlated with verbal ability, full-scale IQ and performance IQ. Lovett and Rehfeldt (2014) successfully used MET with three adults with ASD to establish competent performances on the protocol and some level of generalisation to natural social interactions. O'Neill and Weil (2014) presented the protocol to three adults with mild-moderate intellectual disability and with schizophrenia. Baseline results indicated considerable deficits in responding across all three levels of complexity from 17% accuracy on double reversals to 50% on simple relations. After explicit training, significant improvements were observed on all tasks, with accuracy now ranging from 79% on double reversals to 96% on simple relations.

Villatte, Monestès, McHugh, Freixa, i Baqué and Loas (2008) compared performances on the protocol within a non-clinical sample of college students who scored high versus low on social anhedonia. Overall, both groups performed very well, with minor weaknesses observed on the more complex trials. Where superiority in performance *was* observed in the low social anhedonia group, this occurred on reversed I-YOU and reversed HERE-THERE relations, I-YOU/HERE-THERE double reversals and HERE-THERE/NOW-THEN double reversals. In a related study, Vilardaga, Estévez, Levin and Hayes (2012) reported that performance on a modified version of the deictic protocol correlated with social anhedonia, empathy and experiential avoidance.

Villatte, Monestès, McHugh, Freixa, i Baqué and Loas (2010a) investigated performances on the protocol with 15 adults with and without schizophrenia. The sample with schizophrenia produced significantly more errors on all reversed relations and weaker performances on double reversals. Contrary to previous findings, both groups performed better on double reversals than reversals.

Janssen et al. (2014) compared performances on the protocol between 13 adults with Social Anxiety Disorder (SAD) and 14 control participants. Similar to previous evidence, both groups produced their highest levels of accuracy on simple relations,

with lowest on double reversals. Interestingly, while highest accuracies were observed on I-YOU relations, both groups emitted their next best performances on NOW-THEN relations and their lowest accuracies on HERE-THERE relations. The two groups only differed significantly on reversals, with the control group showing superiority in this regard. In a related study, Hendriks et al. (2016) compared the performances of 27 individuals with anxiety, eight with psychosis and 23 control participants. The results showed that all groups produced their highest levels of accuracy on simple relations and the lowest on double reversals, with some evidence that the group with psychosis produced lowest accuracies overall, while the control group produced the highest level of accuracy.

Possible limitations to using the deictic protocol in adults and clinical samples. Although the deictic protocol has been used less often with adults than children, and has not been used extensively with clinical samples, several authors have raised concerns about using the tool with these samples (Hussey et al., 2014). These concerns may be summarised as follows: (1) The protocol was explicitly designed for developmental purposes (i.e. use with young children) to establish deictic relations when they were found to be absent or deficient (see Barnes-Holmes, 2001); (2) Even typically-developing adults show deficits on specific deictic relations when these relations are not presented as they typically are in natural language (McHugh et al., 2004; Vitale et al., 2008); (3) RFT does not necessarily predict that psychological suffering involves *deficits* in relational responding; (4). It is *possible* that deficits or unexpected patterns of deictic relational responding *might* be observed in psychological suffering, but more meaningful effects would likely be obtained if the deictic relations were specific to the domain of interest (e.g. an individual's levels of anxiety relative to others).

Summary. Behavioural researchers working under the rubric of RFT have proposed a distinction between verbal self-discrimination as observed in humans and nonverbal self-discrimination as observed with nonhumans. According to RFT, verbal self-discrimination and perspective-taking comprise repertoires of AARR. For RFT, perspective-taking is AARR that becomes abstracted across multiple exemplars of talking about your perspective in relation to others. The core relations involved in

perspective-taking are referred to as deictic relations, and include responding from one's own perspective in terms of interpersonal relations, spatial relations and temporal relations. Most of the empirical research on deictic relational responding has employed various iterations of a developmental protocol. This research supports the distinctions among the three types of deictic relations and the finding that these relations vary on a continuum of complexity from simple relations to reversed relations, and double reversed relations. In general, accuracies in performances on the deictic relational protocol increase as a function of age. However, even adults produce error rates and, in some cases, adult performances are better on double reversals than reversals. Once established via MET, perspective-taking repertoires can generalise to both new stimuli and contexts. Studies assessing patterns of deictic relational responding in atypical populations have found that children with ASD produce somewhat weaker patterns than their peers. Results with atypical adults have found that performances on the deictic protocol correlate with verbal ability and IQ. Participants with a diagnosis of schizophrenia have been shown to produce significantly weaker results on reversed and double reversed relations than typically-developing counterparts. Several researchers have raised concerns about using the protocol to assess deictic relational repertoires in adult samples.

Exploring the putative relationship between performances on the deictic relations protocol and ToM tasks. A considerable number of the studies described above attempted to assess the relationship between children's performances on the deictic protocol and traditional ToM tasks, in order to determine the potential functional overlap between the skills targeted by each. Specifically, with typically-developing children, Weil, Hayes and Capurro (2011) reported that establishing competent performances on the protocol generalised to ToM tasks, although Jackson et al. (2014) reported that training on the protocol was *not* sufficient to improve weak performances on ToM tasks with atypically-developing children.

Four of the studies described in the earlier sections with adults from clinical samples have also attempted to assess the relationship between performances on the protocol and ToM tasks. With adults with mild-moderate intellectual disability and

schizophrenia, O'Neill and Weil (2014) reported that training on the protocol significantly improved weak baseline performances on the Deceptive Container Task (Perner, Frith, Leslie, & Leekam, 1989) and the Hinting Task (Corcoran, Mercer, & Frith, 1995). Similarly, when Lovett and Rehfeldt (2014) used MET to establish deictic relational responding in adults with ASD, they saw improvements on the Theory of Mind Inventory (TOMI). Villatte et al. (2010a) found that performances on reversal trials by adults with and without schizophrenia significantly predicted accuracy on the ToM task. Finally, Hendriks et al. (2016) reported that with individuals with anxiety and others with psychosis, performance on the protocol was positively correlated with both the Faux-pas (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999) and the Strange Stories (Happé, 1995) tests (although these correlations were not retained when intelligence was controlled for). Taken together, there is evidence of a functional overlap between deictic relational responding as measured by the protocol and ToM with numerous samples.

Several studies have systematically adapted trials from the deictic protocol to create tasks that closely resemble the attribution of true and false beliefs as targeted by ToM measures, such as the Deceptive Container Task. Specifically, McHugh, Barnes-Holmes, Barnes-Holmes and Stewart (2006) constructed trials that required participants from five different age groups to respond in accordance with the three deictic relations and on some trials to respond in accordance with logical NOT. Consider a *true belief trial* with the following scenario and subsequent questions: "If you put the doll in the cookie jar and I am here. What would I think is in the cookie jar? What would you think is in the cookie jar?" Now consider *a false belief trial containing logical NOT* as follows: "If you put the doll in the cookie jar and I am not here. What would you think is in the cookie jar? What would I think is in the cookie jar?". For RFT, responding correctly to the latter task involves responding in accordance with HERE-THERE, NOW-THEN and logical not (i.e. I did NOT see inside THERE and THEN, so this is what I think is inside HERE and NOW). The results from McHugh et al. indicated that accuracy on both types of trial appeared to increase as a function of age, but there was

no significant difference in accuracy rates when responding to true and false belief was compared, nor were differences recorded between these trial types for any age group.

In a related study, McHugh, Barnes-Holmes, Barnes-Holmes, Whelan and Stewart (2007) compared responding to self versus other in adapted false belief trials, in terms of both accuracy and response latency. Consider the following *self-attribution false belief* trial: "If you put the pencils in the Smarties box and I am not there, would you think the Smarties box contains pencils?". Now consider an *other-attribution false belief trial*: "If I put the pencils in the Smarties box and you are not there, would I think the Smarties box contains pencils?" The results of the study indicated high accuracy overall in false belief responding, but significantly longer latencies on trials involving the perspective of other versus self.

Across two experiments, Villatte, Monestès, McHugh, Freixa, i Baqué and Loas (2010b) used a similar adaptation of true and false belief trials as above and compared 15 control participants, 15 high in social anhedonia, 15 low in social anhedonia and 15 with schizophrenia. The four trial types were as follows: *self-attribution true belief* (e.g. "If I put the pencils in the Smarties box and you are here, you would think the Smarties box contains …..?"); *self-attribution false belief* (e.g. "If I put the pencils in the Smarties box and you are not here, you would think the Smarties box contains….?"); *attribution-to-other true belief* (e.g. "If you put the pencils in the Smarties box and I am here, I would think the Smarties box contains….?"); and *attribution-to-other false belief* (e.g. "If you put the pencils in the Smarties box and I am not here, I would think the Smarties box contains ….?"). The results of Experiment 1 indicated that participants with high social anhedonia were significantly less accurate than the low social anhedonia group on other-attributions. While latencies did not differ across groups, all were significantly faster on self- rather than other-attributions. Experiment 2 indicated that both groups showed longer latencies on other-attributions than self-attributions and on false beliefs versus true. Participants with schizophrenia were significantly less accurate than controls on self-attribution of false belief but there was no difference on true belief. While the individuals diagnosed with schizophrenia also produced more errors on both

types of other-attribution, this was not statistically significant different from the performances of controls on other-attribution tasks.

In a related study, McHugh, Barnes-Holmes, Barnes-Holmes, Stewart, and Dymond (2007) constructed deictic trials that resembled ToM tasks for deception and presented them to five different age groups. There were four experimental trial types which, for example, presented a picture of a teddy on the screen above a picture of a toy box and a refrigerator, along with a trial-specific question. Participants provided their response by placing the picture at the top of the screen on top of one of the pictures below. (1) *First-order Positive trials* included, for example, the question "If I have a teddy and I want you to find it, where should I put the teddy?" A correct response in this example would involve the participant placing the picture of the teddy on top of the picture of the toy box. (2) *First-order Negative trials* included, for example, the question "If you have a teddy and you don't want me to find it, where should you hide it?" A correct response would involve the participant placing the picture of the teddy on top of the picture of the refrigerator. (3) *Second-order Positive trials* included, for example, the question "If I have a teddy and if you know that I know you're trying to hide it from me, where should you hide the teddy?" In this case, a correct response would involve the participant placing the picture of the teddy on top of the picture of the toy box. (4) *Second-order Negative trials* included, for example, the question "If I have a teddy and if I know that you don't know I'm trying to hide it from you, where should I hide the teddy?" A correct response would involve the participant placing the picture of the teddy on top of the picture of the refrigerator. The results showed that in general overall accuracy appeared to improve with age. These differences in improvement were significant for all age groups, except between late childhood and adolescence, and between adolescence and adulthood.

Summary. A considerable number of studies have assessed the relationship between performances on the deictic protocol and traditional ToM tasks, and found evidence of a functional overlap between deictic relational responding on the protocol and ToM with numerous samples. Several studies have systematically adapted trials from the deictic protocol to create tasks that closely resemble the attribution of true and

false beliefs as targeted by traditional ToM tasks. The evidence indicates that accuracy on both types of trials increases as a function of age, but there are no differences in accuracy between true and false belief attribution. Related research has compared responding to self versus other in adapted false belief tasks and found no difference in overall accuracies but significantly longer latencies on trials involving the perspective of other versus self. In studies with atypical samples, those with high social anhedonia have been found to perform significantly more poorly than controls on false belief attributions to others. Similarly, participants with a diagnosis of schizophrenia have performed significantly worse than controls on self-attributions of false belief.

The need for an alternative to the deictic protocol. Much of the work conducted in the area of deictic relational responding has focused on assessing the presence of these patterns of relational responding in a dichotomous manner. That is, participants were typically assessed for the presence of deictic relational responding, and if found to be deficient, these relations were trained and tested to determine if they produced predicted outcomes. However, little research has focused on the relative strength of a pattern once it was observed and there have been recent calls for analyses that focus on the relative strength or persistence of derived relational responding, rather than simply its presence versus absence (e.g. Barnes-Holmes, Barnes-Holmes, Hussey, & Luciano, 2016). In an attempt to develop methodologies for assessing the relative strength of derived relational responding, researchers have explored alternative methodologies, such as the Implicit Relational Assessment Procedure (IRAP; e.g. Barnes-Holmes, Hayden, Barnes-Holmes, Stewart, & Boles, 2008; Hussey, Barnes-Holmes, & Barnes-Holmes, 2015).

The IRAP requires participants to emit two opposing patterns of relational responding and the ease with which one pattern may be emitted relative to the other provides a measure of response strength. Specifically, the procedure typically presents label and target stimuli (e.g. the label word “puppy” with the target word “pleasant”) and requires participants to respond “True” (e.g. puppy-pleasant) or “False” (e.g. puppy-unpleasant) to the stimulus pairs. An IRAP typically comprises four trial types (e.g. *Puppy-Positive, Puppy-Negative, Spider-Positive,* and *Spider-Negative*) that are

generally analysed independently in terms of the difference in response latencies between responding that is deemed consistent versus inconsistent with a participant's verbal history. In general, response latencies are expected to be shorter during blocks of trials that require history-consistent versus history-inconsistent responding.

The body of empirical research employing the IRAP has grown considerably, with an increasing focus on clinically relevant phenomena (Vahey, Nicholson, & Barnes-Holmes, 2015). Using the IRAP to assess deictic relational responding, particularly in the clinical domain, would provide an alternative to the Barnes-Holmes (2001) protocol. A recent study in which the IRAP was used to target responding to self versus others seems particularly relevant (Barbero-Rubio, Lopez-Lopez, Luciano, & Eisenbeck, 2016).

The study presented participants with their own names and the name of the researcher as label stimuli, and statements pertaining to specific characteristics of the self versus other as targets (e.g. "is in front of the laptop"). There were two response options ("yes" and "no") on each trial. The four trial types in this study were referred to as: *I-I* (participant name-participant characteristics), *Other-Other* (researcher name-researcher characteristics), *I-Other* (participant name-researcher characteristics) and *Other-I* (researcher name-participant characteristics). In general, the pattern of IRAP effects reported by Barbero-Rubio et al. (2016) indicated that participants' response latencies showed significantly more rapid responding on the *I-I* trial type relative to the other three trial types during history-consistent blocks (i.e. responding "True" on *I-I* and *Other-Other* trial types, and responding "False" on *I-Other* and *Other-I* trial types). In addition, the difference in response latencies between consistent and inconsistent blocks for each trial type was in the predicted direction (i.e. shorter on history-consistent relative to history-inconsistent trials), and these differences were significant in terms of the normalized D_{IRAP}-scores.

The Current Thesis

Following a review of the literature on perspective-taking (as above), the primary aim of the current thesis was to determine the potential utility of the IRAP in the domain of perspective-taking. To this end**, Chapter 2,** systematically replicated previous

research by Barbero-Rubio et al. (2016), in which participants were presented with an IRAP that contained each participant's own name (self) versus the name of the researcher (other) as label stimuli, and statements describing specific characteristics of the self (e.g. "is in front of the laptop") versus other (e.g. "is standing up") as target stimuli, along with "Yes" and "No" as response options. In an extension of the research by Barbero-Rubio et al., the current research included a control condition in which no responding to self was involved, only to others. The results in Experiment 1 yielded significant IRAP effects for two of the trial types in both the deictic and control IRAPs, whereas four significant effects were observed in the Barbero-Rubio et al. study, in which there was no control condition. Given the differences between the results of Experiment 1 and those found by Barbero-Rubio et al., Experiment 2 investigated whether introduction of a *read-aloud procedure* would produce a pattern of effects similar to the original study? On balance, Experiment 2 replicated the effects on the *I-Other* and *Other-I* trial types reported in the original Barbero-Rubio study (i.e. they were both significantly different from zero). The additional analyses conducted here, however, indicated that the effects were both significantly weaker than the *I-I* and *Other-Other* trial types. The control IRAP also yielded significant effects for all four trial types, although two of the trial types (*Researcher-Researcher* and *Other-Other*) continued to be significantly stronger than the two remaining trial types. One possible concern that could be raised regarding the two studies involving the self- versus other-IRAPs conducted in Chapter 2 is that differences that emerged between responding to self and other within the IRAP could be attributed to factors other than perspective-taking per se. For example, a pattern known as the single-trial-type-dominance-effect (STTDE) emerged in Experiment 1 (see below). A second potential concern that could be raised regarding both experiments pertains to the simple target phrases that specified characteristics of self and other (e.g. "is sitting down", "is the participant", "is in front of the computer"). As such, it could be argued that responding on the IRAP simply required deictic relational responding, but not perspective-taking. Indeed, perspective-taking would appear to require more complex target statements or relational networks that involve taking the perspective of self versus other.

Chapter 3 sought to develop IRAPs that taregeted perspective-taking with regard to the mental states of self and other**.** This was pursed in a series of experiments that employed a novel version of the IRAP, known as the Natural Language-IRAP (NL-IRAP). Across a sequence of six experiments, a 'self-focused IRAP' required participants to respond to both positive and negative statements about themselves, whilst an 'other-focused IRAP' required participants to respond to similar statements about others. Experiment 1 and 2 investigated perspective-taking with regard to an *unspecified* other. Experiments 3-6 investigated perspective-taking with regard to a *specified* other, with the specified relationship between self and other manipulated across experiments. The results from the first two experiments indicated that there were significant differences between the self- versus other-focused IRAPs, when the other remained unspecified. The remaining four experiments, however, indicated that when the other was specified there was limited evidence that performances on the two IRAPs differed significantly. Overall, the IRAP effects for the most part, were in the predicted direction. However, on balance, the results could be seen as somewhat disappointing because there was little evidence of perspective-taking when other was specified, at least in terms of different performances across the two IRAPs, or in correlations among the IRAPs and the self-report measures. Perhaps, therefore, there was some sensitivity to self versus other, but the use of complex statements in the IRAPs potentially undermined or reduced the impact of deictic relational responding per se.

In reflecting upon the results obtained in Chapters 2 and 3, a key issue emerged that seemed important to address. Specifically, the challenge was to develop an IRAP that facilitated responding to complex relational networks while maintaining sensitivity to self versus other. **Chapter 4** addressed this issue in two ways:1. Self- and other-pictures were employed in an IRAP to ensure that the functions of self and other relational networks were controlling participants responding; 2. Perspective-taking scenarios were presented before IRAP blocks, rather than presenting complex perspective-taking statements within each IRAP trial. Experiment 1 was pilot research that assessed the feasibility of using matched pictures of self and others. Participants were asked to bring an electronic picture of their face and a picture of another unknown

person, which could be inserted into the IRAP at the beginning of the session. Each participant was then exposed to two IRAPs, with one IRAP containing the picture of themselves and the other containing the matched picture of the other person. Both IRAPs presented the same pictures of pens as the contrast category. The results of the IRAPs were broadly consistent with common-sense expectations. That is, participants confirmed more readily than they denied that a picture of a face was a face and that a picture of a pen was a pen. They also denied more readily than confirmed that a picture of a pen was a face and that a picture of a face was a pen. No significant differences in the sizes of the individual trial type effects, or differences among those effects, emerged between the two (self and other) IRAPs.

Given the relatively positive findings of Experiment 1 in Chapter 4 (i.e. in general, responding in a history-consistent manner), Experiment 2 sought to return to the main objective of the thesis by examining the impact of a brief vignette on performances on a belief-IRAP containing self/other pictures and belief-related statements. Specifically, participants were presented with a famliarisation IRAP, one of two vignettes which described a false belief scenario or a control scenario, followed by a belief-IRAP and eight self-reports. The results indicated vignette-consistent scores for three of the four trial types, with vignette-inconsistent effects recorded on the *Self-Scarf* trial type in both conditions. There was little evidence that the two vignettes impacted differentially upon the IRAP performances. The correlational analyses failed to indicate any clear relationships between the self-report measures and the IRAP. Despite there being no significant difference between IRAP performances across the two conditions, the pattern of results suggest that both vignettes, to some degree, impacted the IRAP effects.

In reflecting upon the results obtained in Experiment 2 in Chapter 4, a number of issues emerged that seemed important to address in a follow-up experiment. First, it became apparent that participants in the Control Condition may have found the relationship between the vignette and the IRAP trial types somewhat ambiguous. Specifically, the control vignette specified that there were two boxes present in the room (one in front of the participant and another in front of the other person), whereas the

belief-IRAP presented statements that specified only one box. As such, it is difficult to interpret the IRAP effects that were observed for the Control group. A related issue pertains to the fact that the order in which the IRAP blocks were presented was not counterbalanced (i.e. the first block of the belief-IRAP was always vignette-consistent). With these issues in mind, we designed a second experiment that once again sought to develop an IRAP that would show some sensitivity to a false belief vignette. The main aim of Experiment 3 was to determine the extent to which false belief vignettes presented before each block of trials in a belief-IRAP would impact the performances observed on that IRAP. In doing so, two specific variables were manipulated across four conditions; 1. The sequence in which the critical stimuli were specified in false belief vignette; and 2. the order in which the IRAP blocks (i.e. vignette-consistent followed by vignette inconsistent versus the opposite block sequence) were presented. The main rationale for employing these versions of the same vignette was to determine if clear differential patterns of responding, consistent with the specified sequence in the vignette, would be observed in the belief-IRAP performances. The results of the third experiment were more conclusive than those of the second, in indicating that the primary controlling variable was the sequence in which the IRAP blocks were presented rather than the content of the vignettes. On balance, post-hoc analyses of performance on the first practice block pairs indicated that the block sequence and vignettes did interact with each other, but only when the vignette and the initial contingencies of the IRAP cohered with each other.

Chapter 5 provides a General Discussion of the findings gathered across the eleven experiments reported across the three empirical chapters (2, 3 and 4). The empirical research reported in the current thesis was intertwined throughout with the on-going development of conceptual analyses within RFT itself, and thus it is important to place the current empirical research in that context. In order to achieve this objective, Chapter 5 introduces two separate useful conceptual analyses. The first part focuses on what has been labelled the differential arbitrarily applicable relational responding effects (DAARRE) model and the second part focuses on a recent attempt to provide a systematic framework for research within RFT.

Chapter 2
Exploring Differential Trial Type Effects and the Impact of a Read-aloud Procedure on Deictic Relational Responding on the IRAP[3]

[3] This chapter is published in Kavanagh, D., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart C., & Finn, M. (2018). Exploring differential trial type effects and the impact of a read-aloud procedure on deictic relational responding on the IRAP. *The Psychological Record, 68(2),* 163-176.

In a systematic replication and extension of previous research by Barbero-Rubio et al. (2016), participants were presented with an IRAP that contained each participant's own name (self) versus the name of the researcher (other) as label stimuli, and statements describing specific characteristics of the self (e.g. "is in front of the laptop") versus other (e.g. "is standing up") as target stimulu, along with "Yes" and "No" as response options. The extension of the original work involved the inclusion of a control condition in which no responding to self was involved, only responding to others. That is, the control condition involved an IRAP in which none of the trial types required responding to self, but only responding to two other separate individuals (i.e. the researcher and a picture of another unknown participant). If the comparison between self and other in a deictic-IRAP is an important variable, one might expect a different pattern of results in a control condition in which there is no contrast between self and other. Another way in which the current research extended Barbero-Rubio et al. was the inclusion of self-report measures of self-esteem and of the presence of psychotic-like experiences. One measure of perspective-taking employed in the original Barbero-Rubio study was also retained. All measures were included on an exploratory basis; hence no specific predictions were made.

Experiment 1

Method

Participants. Forty participants were recruited for Experiment 1, 28 females and 12 males. They ranged from 18-36 years old (M = 23.34). All participants were recruited through random convenience sampling from the Department of Experimental. Clinical and Health Psychology, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. Experiment 1 comprised three computer-based tasks: a familiarisation IRAP, a deictic IRAP and a control IRAP. In all three, participants were required to respond to two others rather than to the self versus others. The study also included three questionnaires: the Community Assessment of Psychic

Experiences (CAPE); the Interpersonal Reactivity Index (IRI; perspective-taking sub-scale only) and the Rosenberg Self-esteem Scale (RSES). All materials were presented in Dutch (but they are translated into English when referred to in the text of the current chapter).

Familiarisation IRAP. The IRAP was presented on standard personal computers. The IRAP software was used to present the instructions and stimuli and to record responses. The familiarisation IRAP did not contain stimuli relevant to deictic relations and was employed simply to familiarise participants with the procedure, because no specific pre-block rules for responding were presented in any of the IRAPs. The familiarisation IRAP presented two label words at the top of the screen: *Fruits* and *Vegetables* (see Table 1). Eight target words were individually presented in the centre of the screen; four were fruits (e.g. "Pear") and four were vegetables (e.g. "Broccoli"). The response options "Yes" and "No" were presented at the bottom left- and right-hand corners. The familiarisation IRAP comprised four trial types: *Fruit-Fruit, Vegetable-Vegetable, Fruit-Vegetable and Vegetable-Fruit* (see Figure 1).

Table 1

Labels, targets and response options in the familiarisation IRAP

Labels	
Fruit	Vegetables
Targets	
Apple	Carrot
Banana	Potato
Orange	Broccoli
Pear	Sprout
Reponses	
Yes	No

Note: Stimuli were presented to participants in Dutch.

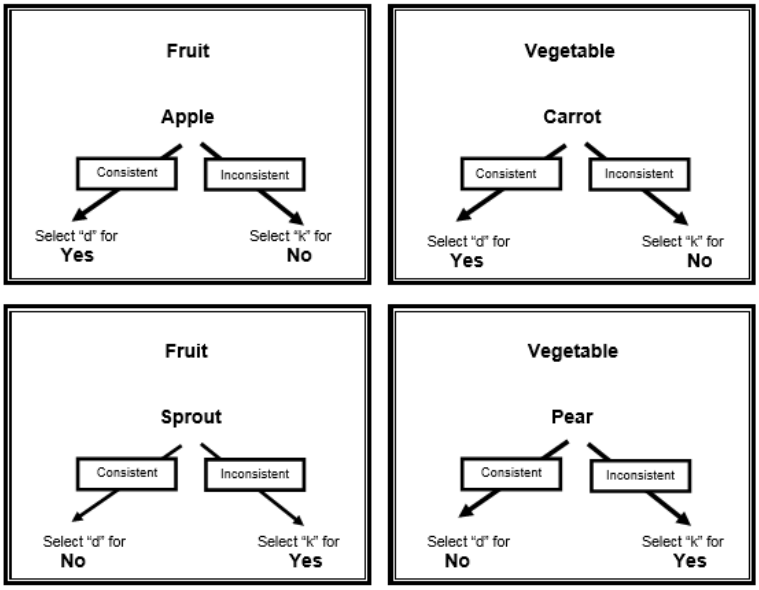

Figure 1. Examples of the four trial types in the familiarisation IRAP: *Fruit-Fruit, Vegetable-Vegetable, Fruit-Vegetable* and *Vegetable-Fruit.* The words *Consistent* and *Inconsistent* were not shown on-screen.

Deictic IRAP. The deictic IRAP presented two label stimuli (participant name/researcher name) on the top of the screen (see Table 2). The target stimuli comprised 12 statements; six described characteristics of the participant at the present time (e.g. "has a yellow Post-it*"*) and six described characteristics of the researcher (e.g. "has an orange Post-it*"*). Again, the response options were "Yes" and "No". The four trial types were denoted as: *I-I* (participant name-participant characteristics), *Other-Other* (researcher name-researcher characteristics), *I-Other* (participant name-researcher characteristics) and *Other-I* (researcher name-participant characteristics).

Table 2

Labels, targets and response options in the deictic IRAP

Labels	
David (Participant's name)	Dee (Researcher's name)
Targets	
is seated	is standing up
is the participant	is the researcher
is on the keyboard	has a pen
looking at screen	has a notebook
is here	is there
has a yellow Post-it	has an orange Post-it
Reponses	
Yes	No

Note: Stimuli were presented to participants in Dutch.

Control IRAP. The control IRAP presented the label stimuli "Dee" (researcher's name) and "Ciara" (name of an individual, whose picture was placed on the wall in front of participants, see Table 3). To match the deictic IRAP, the target stimuli comprised 12 statements; six described features of the researcher (e.g. "has brown hair") and six described features of the person in the picture (e.g. "has blond hair"). Again, "Yes" and "No" were the response options. In denoting the four trial types, the term *Researcher* will be used to refer to the actual researcher and the term *Other* will be used to refer to the person shown in the picture. Please note that all four trial types involved responding to another and *not* the self. The four trial types were thus denoted as: *Other-Other* (picture of other and characteristics of other), *Researcher-Researcher* (researcher and researcher characteristics), *Other-Researcher* (picture of other and researcher characteristics) and *Researcher-Other* (researcher and characteristics of other).

Table 3

Labels, targets and response options in the control IRAP

Labels	
Ciara (Picture)	Dee (Researcher's name)
Targets	
is seated	is standing up
is the participant	is the researcher
is on the keyboard	has a pen
looking at screen	has a notebook
is here	is there
has a yellow Post-it	has an orange Post-it
Reponses	
Yes	No

Note: Stimuli were presented to participants in Dutch.

The Community Assessment of Psychic Experiences (CAPE; Stefanis et al., 2002). The CAPE measures psychotic-like experiences in the general population and was employed because perspective-taking has been implicated in psychotic-like experiences (e.g. Savla, Vella, Armstrong, Penn, & Twamley, 2013). The scale consists of 42 symptom items rated along three sub-scales: positive symptoms (20 items, e.g., "Do you ever feel as if there is a conspiracy against you?"), negative symptoms (14 items, e.g., "Do you ever feel that you experience few or no emotions at important events?") or depressive symptoms (eight items, e.g., "Do you ever feel sad?"). Each item is rated on two 4-point Likert scales from 0 (*never*) to 3 (*nearly always*) to indicate (1) the frequency of symptoms and (2) the level of distress associated with each symptom. The CAPE provides overall frequency and distress scores of psychic experiences, and total frequency and distress scores for each of the three sub-scales. In order to account for partial non-responses, all scores are weighted for the number of valid answers per subscale (i.e. sum score divided by number of items completed). Overall frequency and distress scores are also weighted. In all cases, higher scores indicate greater frequency or distress regarding symptoms, although there are no clinical cut-offs for this measure. The Dutch version was completed by participants. The scale

has demonstrated adequate reliability: positive dimension alpha = 0.63, negative dimension alpha = 0.64, and depressive dimension alpha = 0.62 (Konings, Hanssen, Van Os, & Krabbendam, 2006).

The Rosenberg Self-esteem Scale (RSES; Robins, Hendin, & Trzesniewski, 2001). The RSES is a 10-item measure of self-esteem. All items (e.g. "I take a positive attitude toward myself") are rated on a four-point scale from 0 (*strongly agree*) to 3 (*strongly disagree*). The RSES yields an overall score, with a maximum of 30 and a minimum of 0. The Dutch version has shown high internal consistency (Cronbach's alpha = .86) and high congruent validity (Franck, De Raedt, Barbez, & Rosseel, 2008).

Perspective-taking (PT) sub-scale of the Interpersonal Reactivity Index (IRI; Davis, 1980). The PT sub-scale of the IRI measures perspective-taking. The sub-scale consists of seven items (e.g. "I sometimes find it difficult to see things from the other guy's point of view") rated on a five-point scale from 0 (*does not describe me*) to 4 (*describes me very well*). The sub-scale yields an overall score, with a maximum of 28 and a minimum of 0. High scores indicate strong perspective-taking and low scores indicate weak perspective-taking. There are no clinical cut-offs for this measure. The Dutch version has demonstrated good internal consistency (Cronbach's alpha = .73), and construct validity (De Corte et al., 2007).

Procedure. Experiment 1 took place on an individual basis in sound-proof cubicles at the Department of Experimental, Clinical and Health Psychology, Ghent University. Informed consent was obtained from all individual participants. All participants were exposed to the same experimental sequence, as follows: familiarisation IRAP, deictic IRAP, control IRAP, RSES, PT scale and the CAPE, always presented in this order.

Familiarisation IRAP. The familiarisation IRAP was employed to establish competent performances on a simple word-based IRAP (*Fruits* vs. *Vegetables*) prior to completion of the deictic IRAP. Participants were simply instructed to determine, based on individual trial feedback, what the task involved. Consider a trial with the label

"Fruits" and the target "Pear*"*. 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, "Yes" and "No") 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.

Consistent trial blocks required responding that was in accordance with the pre-experimental verbal history of the participants: *Fruit-Fruit/Yes, Vegetable-Vegetable/Yes, Fruit-Vegetable/No* and *Vegetable-Fruit/No. Inconsistent* trial blocks required responding that was *not* in accordance with pre-experimental verbal relations: *Fruit-Fruit/No, Vegetable-Vegetable/No, Fruit-Vegetable/Yes* and *Vegetable-Fruit/Yes*. The familiarisation IRAP always commenced with a consistent block of trials. When participants selected the response option that was deemed correct within that block, the label, target and response option stimuli were immediately removed from the screen, and the next trial was presented after an inter-trial interval of 400 ms (the label, target and response option stimuli then appeared simultaneously at the beginning of the next trial). When participants selected the response option that was deemed incorrect for that block, the stimuli remained on the 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). Participants were required to achieve both accuracy (≥ 80% correct responding) and latency criteria (≤ 2,000 milliseconds) in every block. As is typical in IRAPs, performance feedback was presented at the end of each block. The program automatically recorded response accuracy (based on the first response emitted on each trial) and response latency (time in ms between trial onset and the emission of a correct response) on each trial.

The familiarisation IRAP differed from a typical IRAP in that it contained only practice blocks (i.e. these were not followed by test blocks). Participants were exposed to a maximum of three pairs of blocks, with 24 trials per block (12 for each type of target stimulus, fruit or vegetable). If participants achieved both accuracy and latency

criteria on the first, second or third pair of blocks, they proceeded to the deictic IRAP. All participants completed the familiarisation IRAP within three sets of blocks.

Deictic IRAP. The format of the deictic IRAP was identical to Barbero-Rubio et al. (2016), except that explicit rules were not provided (it was assumed that the necessary competence on the task had been established by the familiarisation IRAP). The deictic IRAP comprised a maximum of four pairs of practice blocks, followed by three pairs of test blocks. On each trial, there was a label at the top of the screen (*participant's name* or *researcher's name*), a target in the centre of the screen (e.g. "is the participant" or "is the researcher") and two response options ("Yes" and "No") at the bottom left and right of the screen. 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, "Yes" and "No") 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.

When participants selected the response option that was deemed correct within that block, an inter-trial interval of 400 ms was presented, after which the next trial occurred. When participants selected the response option that was deemed incorrect for that block, the stimuli remained on the screen and a red "X" appeared beneath the target stimulus. Only when the correct response option was selected did the program proceed to the 400 ms inter-trial interval (followed by the next trial). This pattern of trial presentations, with corrective feedback, continued until the entire block of 24 trials was presented. Trials were presented 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 blocks required responding that was in accordance with pre-experimental verbal relations: *I-I/Yes, Other-Other/Yes, I-Other/No,* and *Other-I/No.* Inconsistent blocks required the opposite: *I-I/No, Other-Other/No, I-Other/Yes* and *Other-I/Yes.* Again, all participants experienced a consistent block first.

When participants completed a block of trials, the IRAP program delivered feedback on their performance during that block. The feedback consisted of a message

informing them how accurately and how quickly they had responded. 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% correct and a maximum median latency of no more than 2000 ms on each block. If participants achieved both accuracy and latency criteria on the first, second, third, or fourth pair of practice blocks, they proceeded to the first pair of test blocks; if they failed on the fourth pair of practice blocks participation in the experiment was terminated.

A fixed set of six test blocks was presented with no accuracy or latency criteria required 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.

Control IRAP. The format of the control IRAP was similar to the deictic IRAP. Consistent blocks required responding that was in accordance with pre-experimental verbal relations: *Other-Other/Yes, Researcher-Researcher/Yes, Other-Researcher/No* and *Researcher-Other/No*. Inconsistent blocks required the reverse: *Other-Other/No, Researcher-Researcher/No, Other-Researcher/Yes* and *Researcher-Other/Yes*. Again, all participants experienced a consistent block first.

Questionnaires. Participants completed the three questionnaires in the following sequence: the PT sub-scale, the RSES and the CAPE.

Ethical Considerations. All aspects of Experiments 1 and 2 in the current chapter were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All procedures received ethical approval from the Ethical Commission, Faculty of Psychology and Educational Sciences, Ghent Univeristy (Approval date: April 27th, 2016, Reference: 2016/38/Deirdre Kavanagh). Informed consent was obtained from all individual participants.

Results

Questionnaire data. A summary of the means and standard deviations for the CAPE, the RSES and the Perspective-taking (PT) sub-scale is provided in Appendix A. The overall CAPE and sub-scale scores were relatively low, indicating low psychotic-like symptoms. Scores on the RSES were relatively high, indicating overall high self-esteem. The PT sub-scale scores were relatively high, indicating good perspective-taking abilities.

IRAP Data. Mean response latencies for consistent and inconsistent blocks, divided according to trial type, were calculated for each participant. Specifically, in order to pass the practice blocks and advance to test blocks, participants were required to maintain an accuracy level of ≥ 80% correct and a median latency of ≤ 2,000 ms. Based on these criteria, three participants failed to complete practice blocks in the deictic IRAP and did not proceed to the test blocks. Exclusion criteria also applied to the test blocks, such that participants were required to maintain an accuracy level of ≥ 70% correct and a median latency ≤ 2,200 ms on two of the three successive pairs of the six test blocks. No participants failed to maintain these criteria; hence all data were included in the analysis of the deictic IRAP (N = 37). Any participant who failed the practice blocks of the deictic IRAP did not complete the control IRAP (i.e. three participants). The same criteria were applied to the analysis of the control IRAP and one participant failed to pass the practice blocks on this basis. Another participant failed to maintain criteria across two of the three successive pairs of six text blocks in the control IRAP. The final number of participants included in the analysis for the control IRAP was 35.

***Deictic D_{IRAP}*-scores.** D_{IRAP}-scores were calculated for each of the four trial types, such that positive D_{IRAP}-scores during consistent blocks indicated responding "Yes" more quickly than "No" on *I-I* and *Other-Other* trial types and responding "No" more quickly than "Yes" on *I-Other* and *Other-I* trial types. Negative D_{IRAP}-scores indicated the opposite pattern: responding "No" more quickly than "Yes" on *I-I* and *Other-Other* trial types and responding "Yes" more quickly than "No" on *I-Other* and *Other-I* trial types.

The mean D_{IRAP}-scores and standard errors for each trial type are illustrated in Figure 2. The *I-I* and *Other-Other* trial types produced relatively strong IRAP effects, but the *I-Other* and *Other-I* trial types did not. A one-way repeated measures analysis of variance (ANOVA) yielded a significant main effect for trial type, $F(3,36) = 20.84$, $p < .001$, $\eta p2 = .37$. Post-hoc comparisons (Fisher's PLSD tests, see Table 4), indicated that *I-I* ($M = .57$, $SE = 0.06$) differed significantly from the three other trial types: *Other-Other* ($M = 0.34$, $SE = 0.05$, $p < .01$), *I-Other* ($M = .02$, $SE = 0.07$, $p < .001$) and *Other-I* ($M = 0.08$, $SE = 0.05$, $p < .001$). *Other-Other* also differed significantly from *I-Other* ($p < .001$) and *Other-I* ($p < .01$). Four one-sample *t*-tests indicated that only *I-I*, $t(36) = 9.34$, $p < .001$, $d = 1.6$, and *Other-Other*, $t(36) = 6.4$, $p < .00$, $d = 1.1$, differed significantly from zero. In contrast to Barbero-Rubio et al. (2016), only two of the trial types were significantly different from zero, whereas all four of their IRAP effects were[4].

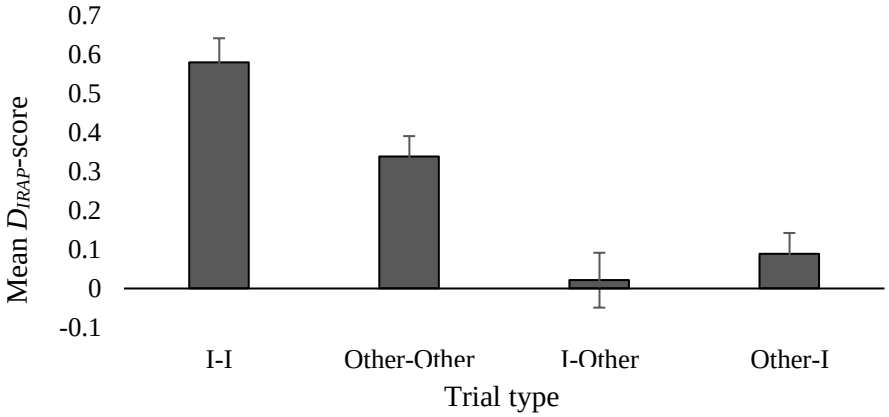

Figure 2. Mean D_{IRAP}-scores on the deictic IRAP trial types in Experiment 1. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores indicate history inconsistent responding.

[4] The raw reaction time data from this IRAP were subjected to the same analyses as those conducted by Barbero-Rubio et al. (2016) and these produced broadly similar results for the trial types that produced significant D_{IRAP}-scores in our study, but not for the trial types that produced non-significant effects.

Table 4
Experiment 1, Fisher's PLSD Comparisons for the deictic IRAP

Trial type	Mean Difference	*P*
I-I, Other-Other	0.55	<.001*
I-I, I-Other	0.48	<.001*
I-I, Other-I	0.24	<.01*
Other-Other, I-Other	-0.07	.40
Other-Other, Other-I	-0.32	<.001*
I-Other, Other-I	-0.25	<.01*

* Indicates significant *p* values.

***Control D_{IRAP}*-scores.** The mean D_{IRAP}-scores per trial type are illustrated in Figure 3. Relatively strong IRAP effects were recorded on the *Other-Other* and *Researcher-Researcher* trial types, with weak effects on the two remaining trial types. A repeated measures ANOVA yielded a significant main effect for trial type, $F(3,34) = 12.49$, $p < .001$, $\eta p2 = .27$. Post-hoc comparisons (see Table 5) indicated that *Other-Other* ($M = .44$, $SE = 0.06$) differed significantly from *Researcher-Other* ($M = .06$, $SE = 0.06$, $p < .001$) and *Other-Researcher* ($M = .04$, $SE = 0.06$, $p < .001$). *Researcher-Researcher* ($M = .32$, $SE = .05$) also differed significantly from *Other-Researcher* ($p < .001$) and *Researcher-Other* ($p < .01$). Four one-sample *t*-tests indicated that only *Other-Other*, $t(34) = 7.55$, $p < .001$, $d = 1.3$, and *Researcher-Researcher*, $t(34) = 6.05$, $p < .001$, $d = 1$, differed significantly from zero.

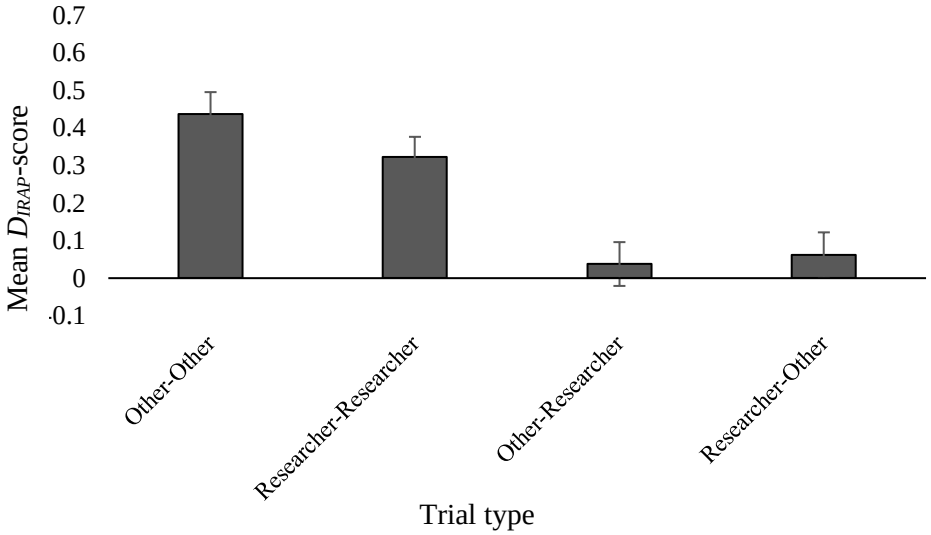

Figure 3. Mean D_{IRAP}-scores on the control IRAP trial types in Experiment 1. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores scores indicate history-inconsistent responding.

Table 5

Experiment 1, Fisher's PLSD Comparisons for the control IRAP

Trial type	**Mean Difference**	***P***
Other-Other, Researcher-Researcher	0.40	<.001*
Other-Other, Other-Researcher	0.37	<.001*
Other-Other, Researcher-Other	0.12	.15
Researcher-Researcher, Other-Researcher	-0.02	.76
Resaercher-Researcher, Researcher-Other	-0.29	<.001*
Other-Researcher, Researcher-Other	-0.26	<.01*

* Indicates significant *p* values.

Correlations. A correlation matrix was calculated to determine if any of the D_{IRAP}-scores from the deictic IRAP predicted self-reported psychotic experiences (on the CAPE), self-esteem (on the RSES) or perspective-taking (on the PT sub-scale). The only significant correlations involved the *Other-Other* trial type with the overall CAPE

frequency, r (28) = -.384, p = .035, and the CAPE positive sub-scale, r (28) = -.475, p < .01, but neither remained significant after Bonferroni corrections.

Summary. The data from Experiment 1 replicated the findings of Barbero-Rubio et al. (2016) to some extent. Specifically, the D_{IRAP}-scores for the *I-I* and *Other-Other* trial types were relatively strong and significant. In contrast to their study, however, the remaining two trial types were both weak and non-significant. Interestingly, a similar pattern was observed with the control IRAP in that two of the IRAP effects were strong and significant, and two were not. A detailed discussion of why the current pattern of trial type effects obtained for the D_{IRAP}-scores did not closely match the original results will be provided in the Chapter 5. Finally, all of the significant correlational analyses between the deictic IRAP and the questionnaires were recorded for the *Other-Other* trial type. At this point, a subsequent experiment to try to replicate Experiment 1 seemed worthwhile.

A second experiment was undertaken that involved a novel method for collecting IRAP data, which had been shown to yield significant effects for all four trial types in a separate line of research being conducted by our group (see Finn, Barnes-Holmes, & McEnteggart, 2018). Specifically, we had found that relatively extreme differential trial type effects were reduced when participants were asked to read aloud the stimuli and response options that appeared on each IRAP trial. Given that relatively extreme differential trial type effects were observed in Experiment 1 of the current chapter, a read aloud procedure was introduced in Experiment 2.

Experiment 2

While the research reported in the current chapter was being conducted, an unrelated study in our research group had found that extreme differential trial type effects may be moderated by the introduction of what could be called a *read-aloud procedure.* Specifically, participants are required to read aloud the label, target and

chosen response option at the time of selection on each IRAP trial throughout the entire procedure. Hence, in Experiment 2, all participants completed all IRAPs using a read-aloud procedure. Given the differences between the results of Experiment 1 and those found by Barbero-Rubio et al. (2016), it seemed useful to investigate whether introducing the read-aloud procedure would produce a pattern of effects similar to Experiment 2.

Method

Participants. A total of 66 individuals participated in Experiment 2, 58 females, seven males and one individual who did not wish to be categorised as either male or female. Ages ranged from 18 to 48 years old (M = 22.98). All participants were recruited through random convenience sampling from the Department of Experimental, Clinical and Health Psychology, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. All materials and apparatus for Experiment 2 were identical to Experiment 1.

Procedure. The procedure for Experiment 2 was largely identical to Experiment 1, except that a read-aloud procedure was added to both IRAPs. This simply required participants to read aloud the label, target and chosen response option at the time of selection on each trial throughout the entire procedure. Similar to Experiment 1, the researcher remained in the room throughout the procedure. However, rather than remaining silent, if the participant failed to read aloud, the researcher reminded them to keep reading aloud during all trials. This was only necessary for a small number of participants and each of these required only one reminder across both IRAPs. Experiment 2 also differed from Experiment 1 in that the order in which the IRAP blocks (i.e. consistent followed by inconsistent versus inconsistent followed by consistent) was counterbalanced across participants. Preliminary analyses yielded no significant effects for this procedural variable and thus it is not included in subsequent analyses.

Results

Questionnaire data. A summary of the means and standard deviations for the CAPE, the RSES and the PT sub-scale is provided in Appendix B. The overall pattern of results from the questionnaires was broadly similar to Experiment 1.

IRAP data. All aspects of data processing for the IRAPs were similar to those employed in Experiment 1. All participants reached the required performance criteria on the practice blocks of the deictic IRAP. All participants also maintained the performance criteria during the test blocks, hence all data were included in the analyses (N = 66). Four participants failed to reach the required performance criteria on the practice blocks of the control IRAP, although all remaining participants maintained the performance criteria during the test blocks. The final number of participants included in the analysis for the control IRAP was 62.

Deictic D_{IRAP}-scores. The mean D_{IRAP}-scores for each trial type are illustrated in Figure 4. Unlike Experiment 1, all trial types produced relatively strong effects, although *I-I* and *Other-Other* were again stronger than the remaining two. A repeated measures ANOVA indicated a significant main effect for trial type, $F(3,65) = 8.98$, $p < .001$, $\eta p2 = .12$. Post-hoc comparisons (see Table 6) indicated that I-I ($M = 0.4$, $SE = 0.04$) differed significantly from *I-Other* ($M = .19$, $SE = .05$, $p < .001$) and *Other-I* ($M = 0.21$, $SE = 0.05$, $p < .001$). *Other-Other* also differed significantly from *I-Other* ($p < .001$) and *Other-I* ($p < .002$). Unlike Experiment 1, *I-I* and *Other-Other* did not differ significantly from each other ($p = .5$). Four one-sample *t*-tests indicated that all trial types differed significantly from zero: *I-I*, $t(65)= 9.43$, $p < .001$, $d=1.17$, *I-Other*, $t(65) = 4.15$, $p < .001$, $d= .51$, *Other-I*, $t(65)= 4.5$, $p < .001$, $d= .56$, and *Other-Other*, $t(65) = 9.06$, $p < .001$, $d= 1.12$.

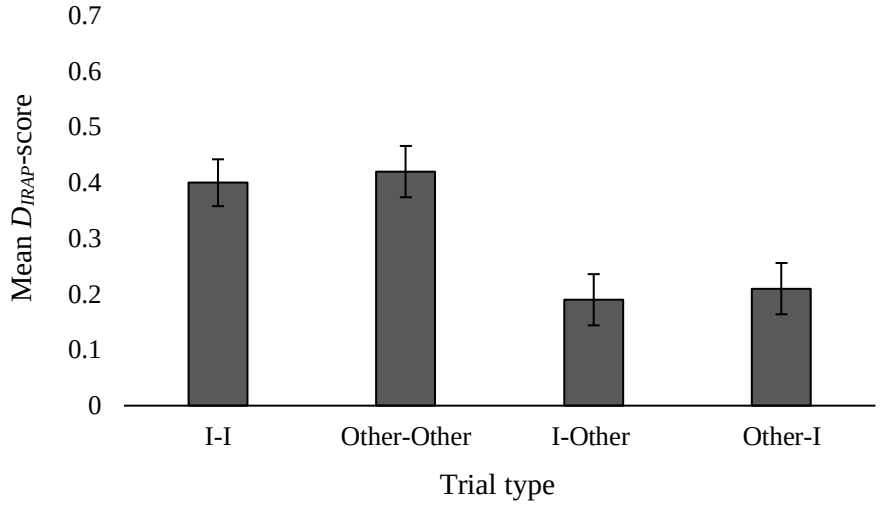

Figure 4. Mean D_{IRAP}-scores on the deictic IRAP trial types in Experiment 2. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores indicate history-inconsistent responding.

Table 6

Experiment 2, Fisher's PLSD Comparisons for the deictic IRAP

Trial type	**Mean Difference**	***P***
I-I, Other-Other	0.21	<.001*
I-I, I-Other	0.19	<.001*
I-I, Other-I	-0.02	.73
Other-Other, I-Other	-0.14	.81
Other-Other, Other-I	-0.23	<.001*
I-Other, Other-I	-0.22	. <.001*

* Indicates significant *p* values.

Control D_{IRAP}-scores. The mean trial type D_{IRAP}-scores are illustrated in Figure 5. All trial types produced relatively strong effects, with the strongest observed on the *Other-Other* and *Researcher-Researcher* trial types. A repeated measures ANOVA indicated that trial type was significant, $F(3,60) = 6.41$, $p < .001$, $\eta p2 = .01$. Post-hoc comparisons, see Table 7, indicated that *Other-Other* ($M = 0.34$, $SE = 0.04$) differed

significantly from *Other-Researcher* (M = .13, SE = .05, p < .001) and *Researcher-Other* (M = 0.13, SE = 0.05, p < .001). *Researcher-Researcher* (M = 0.28, SE = 0.04) differed significantly from *Other-Researcher* (p < .05) and *Researcher-Other* (p < .05). Four one-sample *t*-tests indicated that all trial types differed significantly from zero: *Other-Other*, $t(60)$= 8.44, p < .01, d=1.09, *Other-Researcher*, $t(60)$ = 2.45, p < .05, d= .32, *Researcher-Other*, $t(60)$= 2.83, p < .01, d=.37, and *Researcher-Researcher*, $t(60)$ = 6.49, p < .01, d=.84.

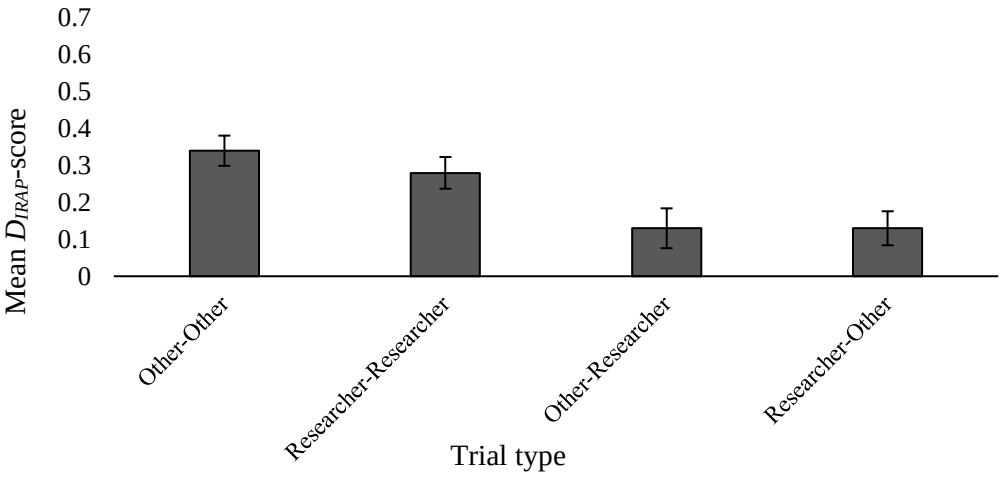

Figure 5. Mean D_{IRAP}-scores on the control IRAP trial types in Experiment 2. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores indicate history-inconsistent responding.

Table 7

Experiment 2, Fisher's PLSD Comparisons for the control IRAP

Trial type	**Mean Difference**	***P***
Other-Other, Researcher-Researcher	0.21	<.001*
Other-Other, Other-Researcher	0.21	<.001*
Other-Other, Researcher-Other	0.06	.311
Researcher-Researcher, Other-Researcher	0.00	.968
Resaercher-Researcher, Researcher-Other	-0.15	<.02*
Other-Researcher, Researcher-Other	-0.15	<.02*

* Indicates significant *p* values.

Correlations. A correlation matrix only yielded significant results for the *I-I* trial type. Specifically, increased response biases in responding to I as I predicted: higher overall frequency of psychotic experiences, $r(63)= .316$, $p = .01$, higher levels of overall distress, $r(63) = .267$, $p = .03$, greater frequency in positive symptoms, $r(63)=.27$, $p = .03$ and greater frequency of depressive symptoms, $r(63)= .25$, $p = .04$, as measured by the CAPE. However, none remained significant after Bonferroni corrections.

Summary. The data from Experiment 2 once again showed relatively strong IRAP effects on the *I-I* and *Other-Other* trial types, replicating the findings from Experiment 1 and those from Barbero-Rubio et al. (2016), although these two trial types no longer differed significantly from each other. On balance, Experiment 2 replicated the effects on *I-Other* and *Other-I* reported in the Barbero-Rubio et al. study (i.e. they were both significantly different from zero). The additional analyses (repeated measures ANOVA and post hoc comparisons of D_{IRAP}-scores) conducted in Experiment 2, however, indicated that the *I-Other* and *Other-I* trial types were both significantly weaker than the *I-I* and *Other-Other* trial types. The control IRAP also yielded significant effects for all four trial types, although two of the trial types (*Researcher-Researcher* and *Other-Other*) continued to be significantly stronger than the two other trial types. Similar to previous research by Finn et al., (2018), the read-aloud procedure appeared to attenuate the differential trial type effect, such that all four trial types (for both IRAPs) were now significantly different from zero.

Discussion

The initial purpose of Experiment 1 was to conduct a systematic replication and extension of Barbero-Rubio et al. (2016). The results yielded significant effects for two of the trial types in both the deictic and control IRAPs, whereas four significant effects were observed by Barbero-Rubio et al. In comparing the current findings with those reported by Barbero-Rubio et al. (2016), it is interesting to note that a different pattern of results was obtained in Experient 1. Specifically, they found significant effects for all four trial types, whereas we did not; furthermore, the effect for the *I-I* trial type in our study was significantly different from the effect for the *Other-Other* trial type (these

trial types did not differ significantly in the original study). In attempting to explain this difference, it is important to note first that some of the procedures involved in running the IRAPs differed substantively between the studies. For example, in Barbero-Rubio et al., participants received explicit perspective-taking instructions at the beginning of each IRAP block (i.e. "For the next block of trials, you have to respond as if you were you and Adrian were Adrian" and "For the next block of trials, you have to respond as if you were Adrian and Adrian were you"). In addition, participants in the original study were required to complete a deictic relational task (DRT), consisting of 20 scenarios, 12 of which involved reversed deictic relations (e.g. “Mario is swimming in the pool, and Ramon is sailing in a boat. If Ramon were Mario, what would he be doing?”) and eight double reversed deictic relations (e.g. “Luis is in Teide analysing sediments, and Maria is in Kilimanjaro searching for the source of a river. If Luis were Maria and if the Kilimanjaro were the Teide, where would Luis be?”). In stark contrast, participants in Experiment 1 of the current chapter were exposed to a basic familiarisation IRAP that focused on fruit and vegetables, with no reference to perspective-taking. Furthermore, when participants were exposed to the deictic and control IRAPs in the current study, no specific instructions concerning perspective-taking were provided either at the beginning of the IRAPs or before each block.

Given the foregoing procedural differences between Barbero-Rubio et al. (2016) and the current study, it is difficult to draw firm conclusions concerning the variables that were responsible for the different patterns of results across the two studies. It is worth noting, however, that the type of instructions that are presented to participants before and during IRAP tasks may have quite dramatic effects on performance (see Finn et al., 2016). The exact manner in which instructions have these effects remains to be elucidated (see Finn et al., 2018); thus, further speculation at this point would be premature.

In Experiment 2, a read-aloud procedure was implemented and the data showed relatively strong IRAP effects on two trial types, replicating the findings from Experiment 1 and those from Barbero-Rubio et al. (2016). However, the data from Experiment 2 did not indicate significant differences between the *I-I* and *Other-Other*

trial types. The results of this second experiment appear relevant to future research that will attempt to use the IRAP to study deictic relational responding and perspective-taking more generally; we will return to this issue in Chapter 5.

It is important to acknowledge a critical limitation in the two experiments reported in the current chapter. Specifically, the control IRAPs were always presented after the deictic IRAPs. Thus, any difference between the deictic and control IRAPs may be due simply to an order effect. On balance, a simple sequence effect does not account for the differences observed between Experiments 1 and 2, because both involved the same deictic-control IRAP sequence. In any case, the current findings call for greater attention to the conditions under which IRAPs are run, including pre-exposure procedures and the types of instructions that are used, and the impact that these and other variables (such as the read-aloud procedure) may have on the functional properties of the IRAP in exploring specific domains, such as deictic relational responding and perspective-taking more generally.

A potential concern that could be raised pertains to the simple target phrases that specified characteristics of self and other (e.g. “is sitting down”, “is the participant”, “is in front of the computer”). As such, it could be argued that responding on the IRAP simply required deictic relational responding, but not perspective-taking. Indeed, perspective-taking would appear to require more complex target statements or relational networks that involve taking the perspective of self versus other. In principle, this sort of complex relational network requires that the participant responds to statements that coordinate with how the self responds in particular situations, versus how they perceive others will respond in the same situations.

Chapter 3

The Search for Perspective-taking IRAPs: Exploring the Potential of the Natural Language-IRAP[5]

[5] This chapter is based on Kavanagh, D., Roelandt, A., Van Raemdonck, L., Barnes-Holmes, Y., Barnes, Holmes, D., & McEnteggart C. (2019). The on-going search for perspective-taking IRAPs: Exploring the potential of the natural language IRAP. *The Psychological Record*, 69 (2), 291-314.

Methodologies for studying perspective-taking as deictic relational responding may require the presentation of complex target statements or relational networks in order to facilitate taking the perspective of self versus other. For example, consider an other-based statement such as "It makes *other people* happy if they win the lottery" versus a self-based statement "Winning the lottery makes *me* happy". In principle, this sort of complex relational network requires a participant to respond to statements that co-ordinate their own response and particular events, versus co-ordinating the responding of others with the same events. Such an IRAP, at least in terms of face validity, appears to target perspective-taking.

Developing such an IRAP would necessarily involve inserting relatively complex statements or networks into the procedure. In doing so, the separation of the stimuli into labels and targets may be problematic because participants may simplify the task by responding to single words or sub-clauses within the labels and targets in such a way that fails to capture perspective-taking. However, one way of potentially avoiding this problem would be to employ a natural language format, as previously reported by Kavanagh, Hussey, McEnteggart, Barnes-Holmes, and Barnes-Holmes (2016). In that study, the IRAP combined the label and target stimuli to form statements that are more similar to natural language.

All of the experiments reported in the current chapter employed two Natural Language-IRAPs (NL-IRAPs) to study perspective-taking, with one IRAP targeting self-perspective and the other IRAP targeting other-perspective. The general strategy pursued here was to determine the extent to which the two NL-IRAPs employed in each experiment would yield any differences in the direction and/or strength of the individual trial type effects. In addition, correlational analyses were conducted to determine the extent to which the two NL-IRAPs correlated, or failed to correlate, with each other and with a range of self-report measures that were deemed to be relevant to perspective-taking. Given the exploratory nature of the research, the self-report measures selected from experiment to experiment were not based on a well-established theoretical rationale. Rather, we selected measures that had frequently been used in previous

studies of perspective-taking within the field of contextual behavioural science. In the later experiments, measures of attachment were included based on comments provided by the Doctoral Advisory Committee, who advised that these measures could be relevant when attempting to analyse perspective-taking with regard to others with whom a participant might have a significant relationship.

Experiment 1 focused on self versus other when the other was unspecified. In developing two IRAPs (i.e. self- versus other-focused), we sought to manipulate only the 'self versus other' variable. Thus, any difference that emerged between the two IRAPs could be readily interpreted as based on the self versus other focus. Because Experiment 1 was exploratory, no formal predictions were made.

Experiment 1

Method

Participants. Fifty-four participants were recruited for Experiment 1, 44 females and 11 males. Participants ranged from 17-38 years old (M = 20.89) and were recruited through random convenience sampling from the Department of Experimental, Clinical and Health Psychology, Ghent University participant pool. Each participant was paid an hourly rate of 10 euro. The general strategy for recruiting numbers of participants was guided by the results of a recent meta-analysis of IRAP effects in the clinical domain, indicating that a minimum of 29 is required to achieve a power of 0.8 for first-order correlations (Vahey, Nicholson, & Barnes-Holmes, 2015). Because participants sometimes failed to reach various performance criteria for the IRAP (details provided subsequently), it was necessary to recruit more than 29 participants to yield an adequate dataset for analyses.

Materials and apparatus. Experiment 1 comprised two computer-based tasks: an other-focused IRAP and a self-focused IRAP. The experiment again included the CAPE and also included the Acceptance and Action Questionnaire-II (AAQ-II). As before, all materials were presented in Dutch, but are presented in English in the current

chapter. The use of the AAQ-II was exploratory, to simply determine if perspective-taking measured with self and other IRAPs correlated with psychological flexibility.

Other-focused IRAP. The other-focused IRAP required participants to respond to various statements about other people. These statements were presented in a natural language format, rather than as label and target stimuli, to form a regular sentence or statement. The IRAP presented 16 statements, each of which comprised two parts. The statement referred to an event (deemed positive or negative) and a positive or negative reaction to that event. The sequence in which the event and reaction were presented in each statement could vary, in that the event could be presented before the reaction, or the reaction could be presented before the event. Nevertheless, the 16 statements were divided into four trial types (see Figure 1), based on whether the event and the reaction were both positive, both negative, or a combination of positive and negative (see Table 1). For illustrative purposes, consider the four following statements "People will be proud if they succeed in their exams" (*Positive Event-Positive Reaction*); "Others are frustrated when they pass an exam" (*Positive Event-Negative Reaction*); "Others feel lucky if they fail an exam" (*Negative Event-Positive Reaction*); and "If others fail an exam, then they are disappointed" (*Negative Event-Negative Reaction*). The response options "Yes" and "No" were presented at the bottom left- and right-hand corners on each trial. The reader should note that the NL-IRAP presented some statements that could be seen as potentially confusing, in terms of how participants should respond, particularly the *Positive-Negative* and *Negative-Positive* statements. However, the general structure of the NL-IRAP remained the same as the standard IRAP in that two of the four trial types could be seen as less coherent than the other two.

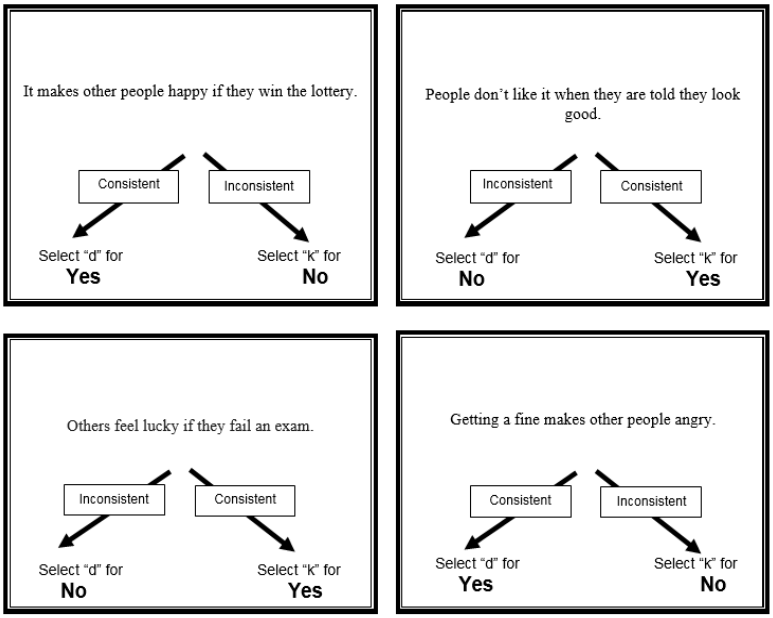

Figure 1. Examples of the four trial types in the other-focused IRAP: *Positive Event-Positive Reaction, Positive Event-Negative Reaction, Negative Event-Positive Reaction, Negative Event-Negative Reaction*. The arrows and words Consistent and Inconsistent were not shown on-screen. Trial type labels denote each of the two parts of the statement, but not necessarily the sequence in which they appeared in the statement.

Table 1

Natural language statements from the other-focused IRAP

Trial types	Stimuli
Positive Event - Positive Reaction	If others hear that they look good, their self-esteem enhances. If the enemy of other people dies, they feel liberated. It makes other people happy if they win the lottery. People will be proud if they succeed in their exams.
Positive Event - Negative Reaction	People don't like it when they are told they look good. If other people's enemy dies they become angry. Other people are disappointed when they win the lottery Others are frustrated when they pass an exam.
Negative Event - Positive Reaction	Other people are happy when a loved one dies. Getting fines makes other people happy. Others feel lucky if they fail an exam. If another person's enemy wins the lottery, they will be happy.
Negative Event - Negative Reaction	Other people feel despair when a loved one dies. Getting a fine makes other people angry. If others fail an exam then they are disappointed. Other people become bitter if their enemy wins the lottery.

Note: Statements were presented to participants in Dutch. Trial type labels denote each of the two parts of the statement, but not necessarily the sequence in which they appeared in the statement.

Self-focused IRAP. The self-focused IRAP was similar to the other-focused IRAP, but required participants to respond to various statements about themselves, rather than about others (e.g. "I'm proud when I succeed in my exams"). These 16 statements were also presented in a natural language format and comprised two parts, an event and a reaction (see Table 2). Again, "Yes" and "No" were the response options. The self-focused IRAP comprised the same four trial types as above.

Table 2

Natural language statements from the self-focused IRAP

Trial types	Stimuli
Positive Event - Positive Reaction	My self-esteem increases if someone says I look good. I feel liberated if my enemy dies. Winning the lottery makes me happy. I'm proud when I success in my exams.
Positive Event - Negative Reaction	I feel ugly if someone says I look good. I'm angry if my enemy dies. Winning the lottery disappoints me. It frustrates me if I succeed in my exams.
Negative Event - Positive Reaction	I'm happy if a loved one dies. Getting fines make me happy. Failing an exam is fantastic. I rejoice if someone I hate wins the lottery.
Negative Event - Negative Reaction	If a loved one dies, I'm miserable. Getting a fine makes me angry. Failing an exam is disappointing. It irritates me if someone I hate wins the lottery.

Note: Statements were presented to participants in Dutch. Trial type labels denote each of the two parts of the statement, but not necessarily the sequence in which they appeared in the statement.

The CAPE was presented as described in Chapter 2.

Acceptance and Action Questionnaire-II (AAQ-II 7-item version; Bond et al., 2011). The AAQ-II measures acceptance, experiential avoidance and psychological inflexibility. The scale consists of seven items rated on a 7-point Likert scale from 0 (*never true*) to 7 (*always true*). The AAQ-II yields an overall score with a maximum of 49 indicating *low psychological flexibility* and a minimum of 7 indicating *high psychological flexibility.* While the measure was not designed as a diagnostic tool, Bond et al. (2011) reported that scores ≥24 correlate with psychological distress. The English version of this scale has been shown to have good internal consistency (*alpha* = .84, Bond et al.). The Dutch translation, used here, has yielded similar reliability values (*alpha* = .85, Bernaerts et al.).

Procedure. Experiment 1 took place on an individual basis in sound-proof cubicles at the Department of Experimental, Clinical and Health Psychology, Ghent University. Informed consent was obtained from all participants. Each participant was

exposed to the other-focused and self-focused IRAPs, with the order of each counterbalanced across participants. Participants thereafter completed the CAPE, followed by the AAQ-II, always presented in this order.

Other-focused IRAP. The other-focused IRAP comprised a maximum of eight pairs of practice blocks, followed by three pairs of test blocks. On each trial, an other-related statement was presented in the middle of the screen (e.g. "People will be proud if they succeed in their exams"), with two response options ("Yes" and "No") at the bottom left and right of the screen. Participants were simply instructed to determine, based on individual trial feedback, what the task involved. 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 "Yes" and "No" 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. The instruction "The previously correct and incorrect answers have been reversed" was presented between blocks of trials.

When participants selected the response option that was deemed correct within that block, an inter-trial interval of 400 ms was presented, after which the next trial occurred. When participants selected the response option that was deemed incorrect for that block, the stimuli remained on the screen and a red "X" appeared beneath the statement. Only when the correct response option was selected did the program proceed to the 400 ms inter-trial interval (followed by the next trial). This pattern of trial presentations, with corrective feedback, continued until the entire block of 32 trials was presented. Trials were presented in a quasi-random order within each block, with the constraint that each statement was presented twice across the 32 trials. *Consistent* trial blocks required responding that was deemed to be in accordance with positive events producing positive reactions and negative events producing negative reactions (i.e. *Positive Event-Positive Reaction/Yes, Positive Event-Negative Reaction/No, Negative Event-Positive Reaction/No,* and *Negative Event-Negative Event/Yes)*. *Inconsistent* trial blocks required responding that was in accordance with positive events producing

negative reactions and negative events producing positive reactions (i.e. *Positive Event-Positive Reaction/No, Positive Event-Negative Reaction/Yes, Negative Event-Positive Reaction/Yes,* and *Negative Event-Negative Reaction/No)*. The other-focused IRAP always commenced with a consistent block of trials.

When participants completed each 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 and how quickly they had responded overall during that block. The average speed of responding was calculated from stimulus onset to the first correct response across all 32 trials within the block. Participants were required to achieve a maximum median latency of no more than 5000 ms *on each trial type*. Although 5,000 ms is unusually long for the latency criterion in an IRAP (most range between 2000 ms and 3000 ms), initial pilot work indicated that most participants failed to reach the latency criterion when set at <5000 ms., especially when the criterion was set at the trial type level (the IRAP latency criterion is typically set at the block level). It is also important to remember that the IRAP effect is calculated from the difference in mean latencies between blocks of trials, rather than from the absolute length of the latencies. As such, the IRAP effect, even when the latency criterion is set relatively high at 5000ms., consists of the difference in average latencies between the blocks, which remains relatively short (i.e. no more than a few hundred milliseconds).

Participants were also required to achieve a minimum accuracy of no less than 75%, also set at the trial type level (i.e. no more than 2 errors were permitted per trial type). If participants achieved both accuracy and latency criteria on any pair of practice blocks, they proceeded to the first pair of test blocks; if they failed on the eighth pair of practice blocks, participation in the experiment was terminated. Although setting the accuracy criterion at 75% is lower than many other studies that have used 80% or more, it is important to note that the accuracy criterion was set at the trial type, rather than the block, level. Setting the accuracy criterion at the trial type level requires a high level of accuracy across all trial types and thus, in a sense, is more stringent than 80% at the block level.

A fixed set of six test blocks was presented with no accuracy or latency criteria required for participants to progress from one block to the next. However, percentage correct and median latency were again presented at the end of each block to encourage participants to maintain criterion-level responding from the practice blocks.

Self-focused IRAP. The format of the self-focused IRAP was identical to the other-focused IRAP, but with statements regarding the self, rather than statements regarding others (e.g. "It irritates me if someone I hate wins the lottery"). It was particularly important in this IRAP to ensure that participants were responding to each of the statements *from their own perspective.* Hence, participants were instructed at the beginning of the IRAP, as follows: "The program will present statements on the screen which refer to you. Please remember that when you see "I" or "me" on-screen, this refers to you (the participant)". The designation of consistent and inconsistent blocks was identical to the previous IRAP. Again, all participants were first presented with a *consistent* block of trials.

Questionnaires. Participants completed the two questionnaires in the following sequence: the CAPE followed by the AAQ-II.

Ethical Considerations. All aspects of Experiments 1-6 in the current chapter were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All procedures received ethical approval from the Ethical Commission, Faculty of Psychology and Educational Sciences, Ghent Univeristy (Approval date: May 2017, Reference: 2017/24/Deirdre Kavanagh). Informed consent was obtained from all individual participants.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations for the AAQ-II and the CAPE weighted overall and with sub-scale scores is provided in Appendix C. The mean AAQ-II score fell below 24, indicating little or no psychological distress. All weighted overall and sub-scale scores of the CAPE were relatively low.

IRAP data. Consistent with standard IRAP practice, mean response latencies for consistent and inconsistent blocks were initially divided according to trial type and calculated for each participant (see Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). Based on the latency and accuracy criteria, three participants failed to complete the practice blocks (and did not proceed to the test blocks) on the self-focused IRAP, and four failed to complete the practice blocks on the other-focused IRAP. Hence, all seven datasets were excluded from further analyses. For the remaining participants, the same accuracy and latency criteria were applied in the test blocks, except that the criteria now applied across *all six* test blocks. This meant that no more than eight errors were permitted per trial type across the six test blocks. Using these criteria, five participants failed to complete the self-focused IRAP and five failed the other-focused IRAP. All 10 datasets were excluded from further analyses, leaving the final number of datasets in the analyses at $N = 37$.

D_{IRAP}-scores. Consistent with the majority of published IRAP studies, D_{IRAP}-scores for both IRAPs were calculated for each of the four trial types (see Barnes-Holmes et al., 2010), such that positive D_{IRAP}-scores during consistent blocks indicated responding "Yes" more quickly than "No" on *Positive Event-Positive Reaction* and *Negative Event-Negative Reaction* trial types, and responding "No" more quickly than "Yes" on *Positive Event-Negative Reaction* and *Negative Event-Positive Reaction* trial types. Negative D_{IRAP}-scores indicated the opposite pattern: "No" more quickly than "Yes" on *Positive Event-Positive Reaction* and *Negative Event-Negative Reaction* trial types, and responding "Yes" more quickly than "No" on *Positive Event-Negative Reaction* and *Negative Event-Positive Reaction* trial types.

The mean D_{IRAP}-scores and standard errors for each trial type for both IRAPs are presented in Figure 2. Positive scores were recorded for all four trial types, with the weakest observed for *Negative Event-Positive Reaction*. For each of the trial types, the mean D_{IRAP}-scores were greater for the other-focused IRAP, with the largest difference for the *Negative Event-Negative Reaction* trial type. A 2x4 repeated measures Analysis of Variance (ANOVA) yielded a significant main effect for IRAP type [$F(1, 36) = 4.12$,

$p < .05$, $\eta_p^2 = .1$] and for trial type, $F(1, 36) = 10.14$, $p < .01$, $\eta_p^2 = .22$, but no significant interaction ($p > .3$). Post-hoc (Fisher's PLSD) comparisons among the four trial types, see Table 3, indicated that *Negative Event-Positive Reaction* differed significantly from the other three trial types and *Positive Event-Positive Reaction* differed from *Negative Event-Negative Reaction*. In examining Figure 2, it appears that the significant difference between the *Positive Event-Positive Reaction* and *Negative Event-Negative Reaction* trial types was driven largely by the other-focused IRAP. On balance, the interaction between IRAP type and trial type was non-significant, and thus it would be unwise to draw any strong conclusions at this point. Indeed, the next study was a direct replication conducted primarily to determine if the difference between the IRAPs was robust and if the relatively strong D_{IRAP} effect for the *Negative Event-Negative Reaction* trial type in the other-focused IRAP would be observed again.

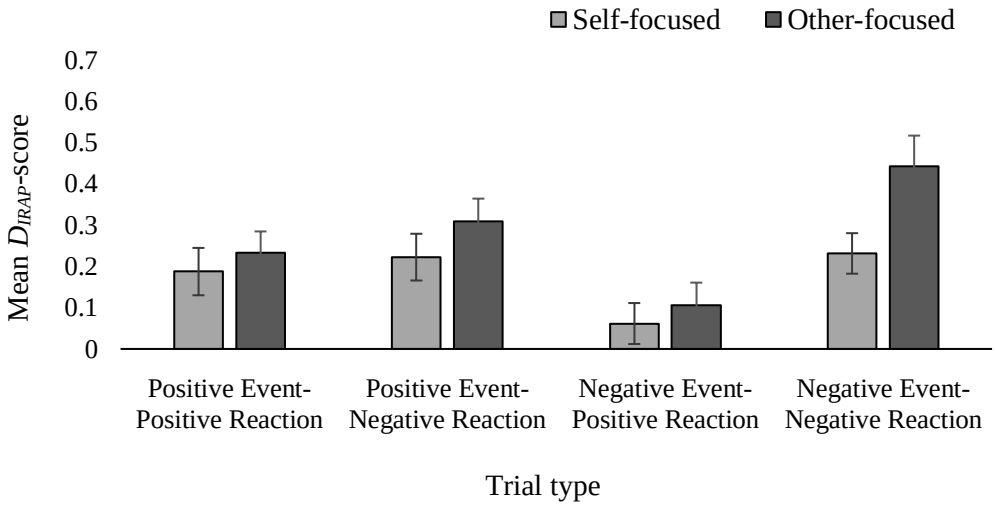

Figure 2. Mean D_{IRAP}-scores on the self-focused IRAP trial types and other-focused trial types in Experiment 1. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores indicate history-inconsistent responding.

Table 3

Experiment 1 Fisher's PLSD Comparisons

Trial type	Mean Difference	*P*
Positive Event-Positive Reaction/ Positive Event–Negative Reaction	-0.06	.25
Positive Event- Positive Reaction/ Negative Event-Positive Reaction	0.13	<.01*
Positive Event- Positive Reaction/ Negative Event-Negative Reaction	-0.13	<.01*
Positive Event-Negative Reaction/ Negative Event- Positive Reaction	0.18	<.001*
Positive Event-Negative Reaction/ Negative Event- Negative Reaction	-0.07	.14
Negative Event-Positive Reaction/ Negative Event-Negative Reaction	-0.25	<.001*

* Indicates significant *p* values.

Correlations. Given that there was a significant main effect for IRAP type, but no interaction with trial type, a single overall *D*-score (the mean of the four trial types) was calculated for each IRAP, and then subjected to correlational analyses with the questionnaires (a total of nine correlations for each IRAP). Neither set of correlations proved to be significant (all *r*s < .32 and all *p*s > .05). In order to determine if *differences in responding to self and others* across the two IRAPs correlated with psychological symptoms, we calculated an IRAP difference score by subtracting the other-focused overall *D*-score from the self-focused overall *D*-score for each participant, and then subjected this to correlational analyses with the questionnaires (a total of nine correlations, not shown). The frequency of positive psychotic-like symptoms was the only significant correlation, $r(37) = 0.34$, $p = 0.039$, but did not remain so after a Bonferroni correction. Finally, the overall *D*-scores from the two IRAPs were correlated and this single correlation proved to be non-significant ($r = 0.208$ and $p = 0.219$).

Summary. The results from Experiment 1 indicated that the other-focused IRAP produced bias scores that were significantly stronger than the self-focused IRAP. There was no significant interaction between IRAP type and trial type, although visual

inspection of Figure 2 indicated a large difference between IRAPs for the *Negative Event-Negative Reaction* trial type. Only one of the 27 correlations proved to be significant but did not remain so after a Bonferroni correction. Overall therefore, the data indicated that there was a significant difference between the self- and other-focused IRAPs, that the IRAP difference score correlated with psychotic-like symptoms, and that the two IRAPs failed to correlate with each other. This pattern of results suggests that the self- and other-focused IRAPs were tapping into different behavioural repertoires, and in this sense this first experiment could be seen as a success. On balance, given the novelty of the procedure employed in Experiment 1 (i.e. the NL-IRAP), it seemed important to conduct a direct replication of the study, with additional questionnaires. Because Experiment 2 remained exploratory, no formal predictions were made.

Experiment 2

Experiment 2 was a direct replication of Experiment 1, except that the AAQ-II was omitted and replaced with several measures of emotional attachment, relationships and self-warmth. A measure that was in development at that time, the Psychological Flexibility Index (PFI), was also included in the battery of tests. All questionnaires were completed after the participants had finished the IRAPs.

Method

Participants. Fifty-one individuals were recruited for Experiment 2, 37 females, 14 males. Ages ranged from 18-49 years (M = 24.5). All participants were recruited through random convenience sampling from the Department of Experimental, Clinical and Health Psychology, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. Both IRAPs were identical to Experiment 1. The CAPE was included again. Five additional questionnaires assessed: psychological flexibility (using the Psychological Flexibility Index, PFI); self-warmth (using a Self-warmth Thermometer); emotional attachments (using the Experiences in Close

Relationships-Relationship Structures questionnaire, ECR-RS); and relationships with others (The Inclusion of Other in the Self, IOS; and the Experiencing of Self Questionnaire, ESQ). The PFI replaced the AAQ-II as a measure of psychological flexibility because it was being developed by Bond and colleagues as an alternative to the AAQ. The Self-warmth Thermometer was included to determine whether performance in the self-IRAP correlated with self-warmth (Vahey, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009). The various attachment questionnaires were included because pre-existing difficulties in attachment relationships may manifest in difficulties in perspective-taking with regard to others (Bernstein, Laurent, Nelson, & Laurent, 2015). All materials were presented in Dutch (translated into English when referred to in the text). The CAPE was the only questionnaire with a validated Dutch version. The instructions and items of the remaining measures were created using a backward forward translation procedure (World Health Organization, WHO, 2017). There are no clinical cut-offs for any of the measures.

Psychological Flexibility Index (PFI). The PFI is a measure of psychological flexibility currently being developed by Bond and colleagues. In its current format, the measure includes 82 items. Each item is rated on a 6-point Likert scale from 1 (*disagree strongly*) to 6 (*agree strongly*), with a minimum of 82 and a maximum of 492, generated by reversing relevant items and then summing the scores. Higher scores indicate higher levels of psychological flexibility, with lower scores indicating lower flexibility. At present, there are no reliability data on this measure.

Experiences in Close Relationships-Relationship Structures questionnaire (ECR-RS; Fraley, Heffernan, Vicary, & Brumbaugh, 2011). The ECR-RS assesses attachment patterns in four close relationships (mother, father, romantic partner, and best friend). Each of the four relationships is rated as a separate domain along two subscales: a) anxious attachment and b) avoidant attachment. The *anxious attachment* subscale comprises 3 items (e.g. “I'm afraid that this person may abandon me”) with a maximum possible score of 21 and a minimum of 3. The *avoidant attachment* subscale comprises 6 items (e.g. “I don't feel comfortable opening up to this person”), with a maximum possible score of 42 and a minimum of 6. Each item is rated on a 7-point

Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Higher scores indicate higher levels of avoidant attachment and anxious attachment. According to Fraley et al., the *alpha* reliabilities for the four relationship domains in the avoidant subscale are between .81 and .92, with the anxiety subscale between .83 and .87. Test-retest reliability is available for only two domains on each subscale, but is adequate (*alpha* =.65 for romantic relationships and .80 for parental relationships).

Inclusion of Other in the Self (IOS; Aron, Aron, & Smollan, 1992). The IOS is a measure of closeness in relationships, comprising two sets of seven Venn diagrams. All Venn diagrams contain one circle that represents the self, while the other circle represents a "best friend" or "other people generally". As such, each set of Venn diagrams represents the relationship between self and a significant other (best friend) or between self and a non-significant other (other people generally). Seven Venn diagrams were presented in each set, with each Venn diagram differing systematically in terms of the extent of the overlap between the two circles. Specifically, in the first Venn diagram, the two circles are completely separate, whereas in the seventh Venn diagram, the two circles are almost fully overlapping, with each Venn diagram in-between showing some variation from one extreme to the other. In order to yield one overall score for the relationship between self and best friend, and one overall score for the relationship between self and other people generally, each Venn diagram is allocated a number between 1 and 7, where 1 represented the least overlap, and 7 represented the most. Hence, the maximum score for best friend/other people generally was 7, with the minimum score 1. The IOS has demonstrated adequate reliability (*alpha* = .93, see Aron et al.).

Experiencing of Self Scale (EOSS; Kanter, Parker, & Kohlenberg, 2001). The EOSS measures the control of others over the experience of the self. It consists of 20 items rated along four subscales (each with 5 items): casual acquaintances-absent (e.g. "My feelings are influenced by casual acquaintances when I am alone"); casual acquaintances-present (e.g. "My wants are influenced by casual acquaintances when I am with them"); close relationships-absent (e.g. "My attitudes are influenced by close relationships when I am alone"); and close relationships-present (e.g. "My actions are

influenced by close relationships when I am with them"). Each item is rated on a 7-point Likert scale from 1 (*never true*) to 7 (*always true*). The maximum overall score is 140 and the minimum is 20, with high scores indicating greater control of others over the experience of self. According to Kanter et al., the scale overall has high internal consistency (*alpha* = 0.91), with internal consistency in the subscales ranging from *alpha* .83-.93.

Self-warmth Thermometer. A feeling-thermometer adapted from Vahey et al. (2009) was used as a measure of subjective self-warmth. The current measure composed an illustrated thermometer with a continuous horizontal scale from 0 (*cold*), rising in intervals of 10, to 100 (*warm*). Participants were required to indicate their self-warmth from 0-100. Given that this is not a standardised measure, there are no reliability data.

Procedure. The two IRAPs were identical to Experiment 1. Participants thereafter completed the CAPE, PFI, IOS, ECR-RS, EOSS and the Self-warmth Thermometer, always presented in this order. All questionnaires were presented to participants via computer using the program Psychopy (Pierce, 2007).

Results and Discussion

Questionnaire data. A summary of the means and standard deviations of each questionnaire and relevant sub-scales is provided in Appendix D. The mean score on the Self-warmth Thermometer was around the mid-way point at 56.65 (/100) and as such was consistent with previous samples (see Vahey et al., 2009). The overall PFI scores were relatively high, indicating high psychological flexibility. The overall CAPE and sub-scale scores were relatively low, indicating low psychotic-like symptoms. The ECR-RS scores were also relatively low in terms of both attachment-anxiety and attachment-avoidance. The overall EOSS and subscale scores were low, indicating low control by others over the experience of self. The IOS scores for best friend were higher than for other people, suggesting a closer relationship in this regard. Nothing unusual or unexpected, therefore, emerged from the questionnaires, given the use of a non-clinical sample.

IRAP data. One participant failed the practice blocks on the self-focused IRAP and another participant failed to complete self-reports, and thus their IRAP data was excluded from further analysis. The final number of participants included in the analyses was $N = 49$.

D_{IRAP}-scores. The mean D_{IRAP}-scores and standard errors for each trial type for both IRAPs are presented in Figure 3. Positive scores were recorded for all four trial types, with the weakest observed for *Negative Event-Positive Reaction*. For three of the four trial types, the mean D_{IRAP}-scores were greater on the other-focused IRAP, with the largest difference between IRAPs for *Positive Event-Negative Reaction*. A 2x4 mixed repeated measures ANOVA produced a main effect for IRAP type , $F(1, 48) = 4.61$, $p = .04$, $\eta_p^2 = .09$, and for trial type, $F(1, 48) = 4.42$, $p < .01$, $\eta_p^2 = .09$, but no significant interaction ($p > .1$). Post-hoc (Fisher's PLSD) comparisons among the four trial types, see Table 4, indicated that *Negative Event-Positive Reaction* differed significantly, or marginally so, from the other three trial types.

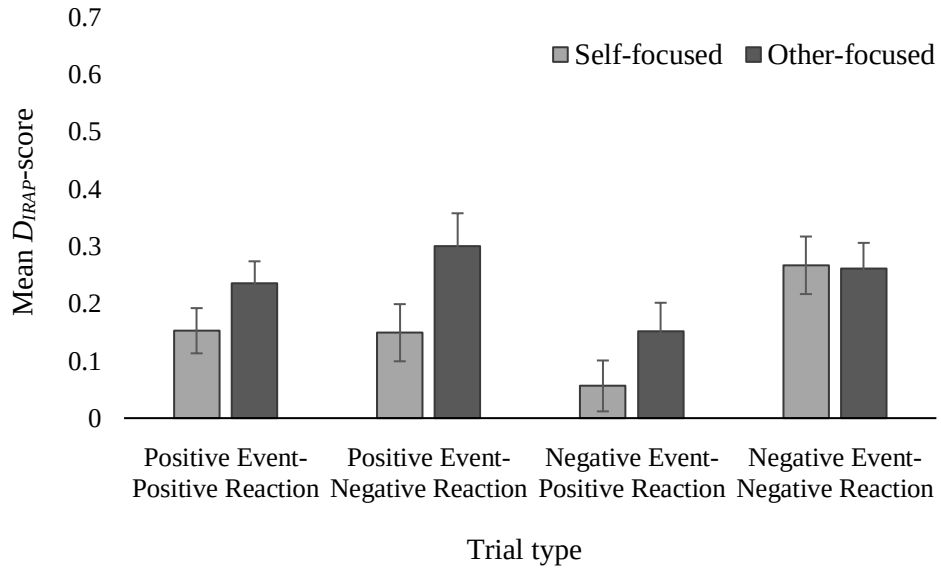

Figure 3. Mean D_{IRAP}-scores on the self-focused and others-focused IRAP trial types in Experiment 2. Positive *D_{IRAP}-scores* indicate history-consistent responding and negative *D_{IRAP}-scores* indicate history-inconsistent responding.

Table 4

Experiment 2 Fisher's PLSD Comparisons

Trial type	Mean Difference	*P*
Positive Event-Positive Reaction/ Positive Event–Negative Reaction	-0.03	.51
Positive Event- Positive Reaction/ Negative Event-Positive Reaction	0.09	.05
Positive Event- Positive Reaction/ Negative Event-Negative Reaction	-0.07	.13
Positive Event-Negative Reaction/ Negative Event- Positive Reaction	0.12	<.01*
Positive Event-Negative Reaction/ Negative Event- Negative Reaction	-0.04	.39
Negative Event-Positive Reaction/ Negative Event-Negative Reaction	-0.16	<.001*

* Indicates significant *p* values

Correlations. The overall *D*-score from each IRAP, as well as the difference scores between the two IRAPs, were subjected to correlational analyses with the questionnaires (i.e. 25 correlations for each IRAP and 25 for the IRAP difference score, not shown). The correlation matrix showed three significant correlations. The other-focused IRAP significantly correlated with the frequency of positive psychotic-like symptoms, $r(49) = -0.3$, $p < 0.04$, but did not do so after a Bonferroni correction. The other-focused IRAP also significantly correlated with the ECR-RS attachment-avoidance subscale for romantic partner, $r(49) = 0.34$, $p < 0.02$, and for mother, $r(49) = 0.31$, $p < 0.03$, but neither remained significant after Bonferroni corrections. There were no significant correlations with the IRAP difference score (all *r*s <.28 and all *p*s > .05). Finally, the overall *D*-scores from the two IRAPs were correlated and this single correlation proved to be significant ($r = 0.315$ and $p = 0.027$).

Summary. The results from Experiment 2 again indicated that the other-focused IRAP produced bias scores that were significantly stronger than the self-focused IRAP. Again, there was no significant interaction between IRAP type and trial type. Interestingly, the relatively large difference between the IRAPs for the *Negative*

Event-Negative Reaction trial type that was observed in Experiment 1, was completely absent in Experiment 2, despite the fact that Experiment 2 was a direct replication. The reason for the failure to replicate the effect for this trial type remains unclear at this time. It is noteworthy that we also failed to replicate this effect across the remaining four studies, and thus it is reasonable to conclude that the effect observed in Experiment 1 was likely due to genuine error variance in the sample. Only three of the 75 correlations proved to be significant, but did not remain so after Bonferroni corrections, and thus should be interpreted with extreme caution. At this point, therefore, the main effect for IRAP type had been replicated, in that the other-focused IRAP produced larger positive D_{IRAP} effects relative to the self-focused IRAP. Differences between the two IRAPs for individual trial types did not appear to be particularly important, in that both experiments failed to yield even a trend towards a significant interaction, and the relatively large difference in the IRAPs observed in Experiment 1 for the *Negative Event-Negative Reaction* trial type was completely absent in Experiment 2. The lack of significant correlations between the IRAPs and the questionnaires in both studies suggests that the measures were not targeting behaviours that overlapped to any great degree. On balance however, the overall *D*-scores from the two IRAPs did correlate in Experiment 2 (but not in Experiment 1), suggesting some overlap in the two measures.

At this point in the research programme, it seemed wise to begin to address the fact that the other-focused IRAP produced stronger D_{IRAP}-scores than the self-focused IRAP. One simple explanation for this difference might be that questions about how other people *in general* react may tend to produce more stereotypical responding. That is, when the other is unspecified, participants tend to produce responses that reflect some general view of other people, rather than of a particular individual. In contrast, when responding to self, a range of potentially important contextual variables may be involved. Thus, for example, when asked how you as an individual react to positive events, there may be a tendency to confirm that in general you react positively but that you are not the type of person to get overexcited, relative to unspecified others. Similarly, an individual may confirm that they react negatively to negative events, but not to the same degree, in general, as other people do. Insofar as this may be the case, it

is possible that an other-focused IRAP that specified a particular other would produce a different result. The next experiment, therefore, was a systematic replication of Experiments 1 and 2, but other was now specified and identified as someone who the participant believed was "the person they are closest to". Identifying this person involved the participant completing a bespoke specified other-focusing task, which is described subsequently. Conducting a third experiment also allowed us to determine if we would replicate the correlation between the two IRAPs found in Experiment 2, but not in Experiment 1. Finding such a correlation would undermine, to some extent, the argument that the two IRAPs were targeting fundamentally different behavioural repertoires. The research remained relatively exploratory and thus no formal predictions were made.

Experiment 3

Method

Participants. Forty participants were recruited for Experiment 3, 29 females, 10 males and one participant who identified as neither male nor female. Ages ranged from 18-45 years (M = 22.67). All participants were recruited through random convenience sampling from the Department of Experimental, Clinical and Health Psychology, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. The self-focused IRAP was identical to that used in Experiments 1 and 2. The other-focused IRAP was similar to the previous experiments, except that the phrase "other people" was replaced with the name of a significant other identified by each participant (e.g. "Sarah"). All questionnaires were identical to those used in Experiment 2.

Specified other-focusing task. This task was designed to help ensure that all participants, in a broadly similar way, identified a significant other to whom they felt closest. The name of the person identified was then presented in the other-focused IRAP. Specifically, participants were asked to write down the name of the individual to whom they felt closest, followed by answers to six further questions about their relationship: 1. length of relationship, rated on Likert scale from 1 (*0-1 year*) to 4 (*10+*

years); 2. closeness of relationship (using one set of the IOS Venn diagrams and referring specifically to the name of the significant other); 3. frequency of contact, rated from 1 (*several times a day*) to 6 (*less than once a month*); 4. level of sharing personal information, rated from 1 ("*I share everything I possibly could*") to 5 ("*I share nothing about myself with this person*"); 5. warmth toward the significant other, on a thermometer from 0 ("*very cold toward the person*") to 100 ("*very warm toward the person*"); and 6. An open-ended description of the most pleasant event shared with this person. The scores on Questions 1-5 were combined and divided by 5 to provide an overall score for the specified other-focusing task, with higher scores indicating greater closeness in the relationship.

Self-focusing task. This task was also developed for the purpose of the current experiment, primarily to control for the extra 5 minutes it took to complete the specified other-focusing task. In this case, participants were simply asked to write down their own first name, followed by answers to five further questions: 1. List your top three hobbies; 2. How frequently do you engage in these hobbies, rated from 1 (*several times a day*) to 6 (*less than once a month*); 3. What is your favourite food; 4. Where is your favourite place; and 5. How warm do you feel toward yourself, from 0 (*very cold toward myself*) to 100 (*very warm toward myself*). Outcomes on this control task were not quantified.

Specified other-focused IRAP. The specified other-focused IRAP was identical to the other-focused IRAP used previously, except that each statement included the name of the individual from the specified other-focusing task (e.g. "Getting a fine makes *David* angry").

Self-focused IRAP. The self-focused IRAP was identical to the previous experiments.

Questionnaires. All questionnaires included in Experiment 3 were identical to Experiment 2.

Procedure. The procedure for Experiment 3 was similar to the procedure from the previous studies, except for the inclusion of the two focusing tasks before each IRAP. Experiment 3 comprised three phases: 1. specified other-focusing task and specified other-focused IRAP; 2. self-focusing task and self-focused IRAP; and 3.

Questionnaires (CAPE, PFI, ECR-RS, EOS, IOS and the Self-warmth Thermometer). The order of the presentation of Phases 1 and 2 was counterbalanced across participants.

Phase 1: Specified other-focusing task and specified other-focused IRAP. The researcher guided participants through the seven items of the specified other-focusing task in a semi-structured manner. Upon completion of the specified other-focusing task, participants were instructed as follows: "Before we begin the next task, I just want you to take a moment to think about [name of significant other]. The next phase of the experiment will be focused, in part, on this person." The specified other-focused IRAP was very similar to the other-focused IRAP from Experiment 2, except that each statement used the name of the significant other and participants were instructed as follows: "The program will present statements on the screen which refer to [name of significant other]. Please remember that when you see [name of significant other], this refers to [nature of the relationship, e.g., your best friend]".

Phase 2: Self-focusing task and self-focused IRAP. The researcher guided participants through the items of the self-focusing task also in a semi-structured manner. Participants were then instructed as follows: "The next task will contain statements about you. Consider the statement, "If I win the lottery I feel happy". It is important to remember that the "I" we are referring to in this statement is you the participant. The procedure for the self-focused IRAP was identical to Experiment 2.

Questionnaires. Participants completed the questionnaires in the following sequence: the CAPE, PFI, ECR-RS, EOSS, IOS and the Self-warmth Thermometer.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations of each questionnaire and relevant subscales is provided in Appendix E. The overall pattern of results from the questionnaires was broadly similar to Experiment 2.

IRAP data. One participant failed the practice blocks on the specified other-focused IRAP; one failed to complete the self-focused IRAP because of limited time; three failed to maintain criteria on the self-focused IRAP; and two failed to do so on the

specified other-focused IRAP. Thus, the final number of participants included in the analyses was $N = 33$.

D_{IRAP}-scores. The mean D_{IRAP}-scores for each trial type are presented in Figure 4. Positive scores were recorded for all four trial types, with the weakest on *Negative Event-Positive Reaction*. For each of the four trial types, the difference between the two IRAPs appeared relatively modest. A 2x4 mixed repeated measures ANOVA produced a main effect for trial type, $F(1, 32) = 9.98$, $p < .001$, $\eta_p^2 = .24$, but not for IRAP type ($p > .8$), nor for the interaction ($p > 0.7$). Post-hoc (Fisher's PLSD) comparisons indicated that *Positive Event-Positive Reaction* differed significantly from the other three trial types, and *Negative Event-Positive Reaction* differed from *Positive Event-Negative Reaction* and from *Negative Event-Negative Reaction* (see Table 5).

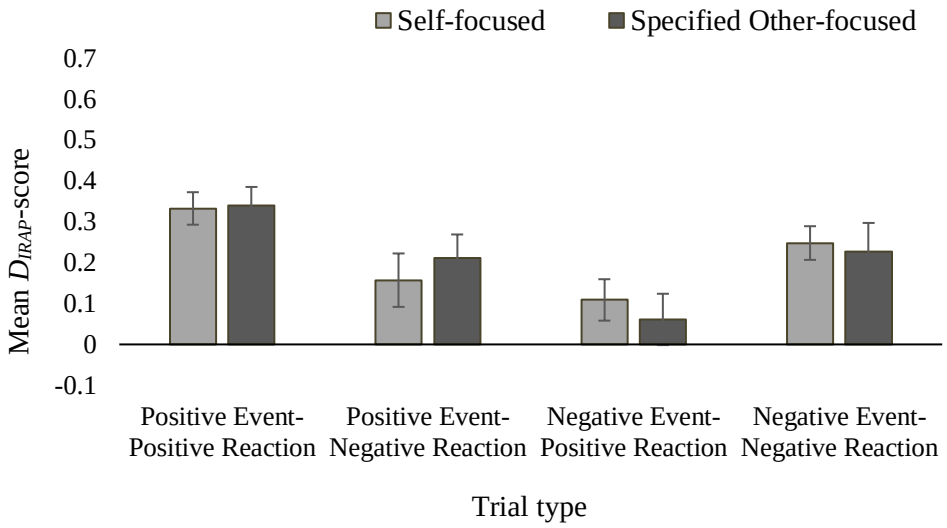

Figure 4. Mean D_{IRAP}-scores on the self-focused IRAP trial types and specified other-focused trial types in Experiment 3. Positive *D_{IRAP}-scores* indicate history-consistent responding and negative *D_{IRAP}-scores* indicate history-inconsistent responding.

Table 5

Experiment 3 Fisher's PLSD Comparisons

Trial type	Mean Difference	*p*
Positive Event-Positive Reaction/ Positive Event–Negative Reaction	*0.15*	*<.01**
Positive Event- Positive Reaction/ Negative Event-Positive Reaction	*0.25*	*<.0001**
Positive Event- Positive Reaction/ Negative Event-Negative Reaction	*0.1*	*<.04**
Positive Event-Negative Reaction/ Negative Event- Positive Reaction	*0.1*	*<.03**
Positive Event-Negative Reaction/ Negative Event- Negative Reaction	*-0.05*	*.31*
Negative Event-Positive Reaction/ Negative Event-Negative Reaction	*-0.15*	*<.001**

* Indicates significant *p* values.

Correlations. The overall *D*-score from each IRAP, as well as the difference scores between the two IRAPs, were subjected to correlational analyses (not shown) with the questionnaires (i.e. 25 correlations for each IRAP and 25 for the IRAP difference score). A correlation matrix showed four significant correlations. The self-focused IRAP correlated positively with frequency of negative psychotic-like symptoms, $r(33) = .38$, $p < 0.03$, and with warmth towards oneself, $r(33) = .49$, $p < 0.01$. The specified other-focused IRAP correlated negatively with avoidant-attachment with one's mother, $r(33) = -.43$, $p < 0.01$. The IRAP difference score correlated positively with avoidant-attachment with one's mother, $r(33) = 0.41$, $p < 0.02$. None of these correlations remained significant after Bonferroni corrections. Finally, the overall *D*-scores from the two IRAPs were correlated and this single correlation proved to be significant ($r = 0.414$ and $p = 0.016$).

Summary. In contrast to the two previous experiments, the difference for IRAP type was non-significant in this third experiment. This suggests that when the other is specified, participants respond in a broadly similar fashion to both self- and other-focused IRAPs. The overall *D*-scores from the two IRAPs also correlated, suggesting some overlap in the two measures. Only four of the 75 correlations proved to be

significant but did not remain so after Bonferroni corrections, and thus again should be interpreted with extreme caution. On balance, Experiment 3 differed from the previous two experiments in two ways. That is, Experiment 3 introduced a specified other, but also involved a focusing task designed to help ensure that all participants identified, in a similar manner, with the significant other to whom they felt closest. It is possible that completing this task before the IRAPs, in some undefined way, undermined the differences in IRAP performances previously recorded in Experiments 1 and 2. Experiment 4, therefore, replicated Experiment 3, which again involved a specified other, but without exposure to the focusing tasks.

Experiment 4

Method

Participants. Thirty-four participants were recruited for Experiment 4, 21 females and 13 males. Ages ranged from 18-39 years old (M = 22.64). All participants were recruited through random convenience sampling from the Experimental, Clinical and Health Psychology Department, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. Both IRAPs were identical to Experiment 3.

Procedure. The procedure was very similar to Experiment 3, except that the self-focusing and specified other-focusing tasks were no longer included. That is, participants were again asked to write down the name of the individual to whom they felt closest, but this was *not* followed by any questions about the nature of their relationship. Note also that no additional questions preceded the self-focused IRAP.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations of each questionnaire and relevant sub-scales is provided in Appendix F. The overall pattern of results from the questionnaires was broadly similar to Experiment 3.

IRAP data. Three participants failed to maintain criteria on the self-focused IRAP and one failed to complete the self-focused IRAP because of time constraints. Thus, the final number of participants included in the analyses was $N = 30$.

DIRAP-scores. The mean D_{IRAP}-scores and standard errors for each trial type for both IRAPs are presented in Figure 5. Positive scores were recorded for all four trial types, with the weakest observed for *Negative Event-Positive Reaction.* For each of the four trial types, the difference between the two IRAPs appeared relatively modest, except for *Positive Event-Negative Reaction*. A 2x4 repeated measures ANOVA yielded a significant main effect for trial type, $F(1, 29) = 10.99$, $p < .001$, $\eta_p^2 = .27$, but no main effect for IRAP type ($p > .8$) or the interaction ($p > .4$). Post-hoc (Fisher's PLSD) comparisons among the four trial types (see Table 6) indicated that *Negative Event-Positive Reaction* differed significantly from the other three trial types.

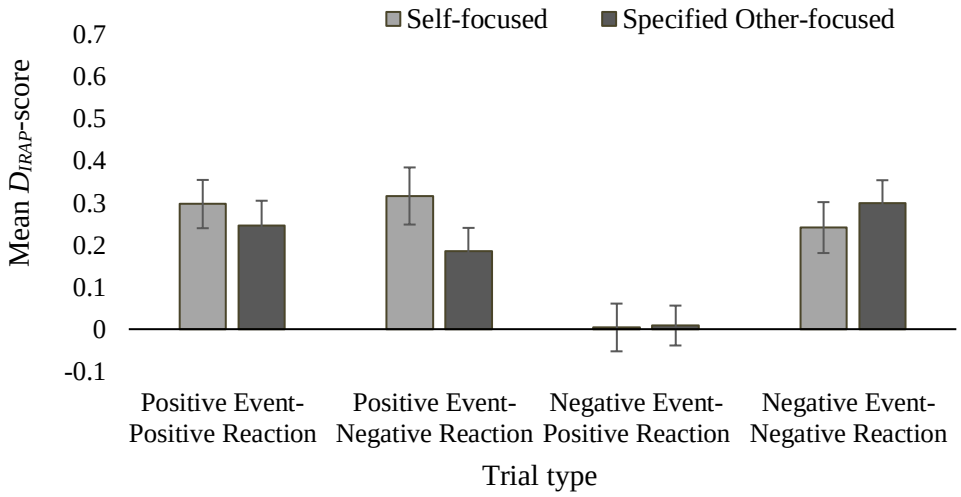

Figure 5. Mean D_{IRAP}-scores on the self-focused IRAP trial types and specified other-focused trial types in Experiment 4. Positive D_{IRAP}-scores indicate history-consistent responding negative D_{IRAP}-scores indicate history-inconsistent responding.

Table 6

Experiment 4 Fisher's PLSD Comparisons

Trial type	Mean Difference	*P*
Positive Event-Positive Reaction/ Positive Event–Negative Reaction	0.04	.49
Positive Event- Positive Reaction/ Negative Event-Positive Reaction	0.27	<.0001*
Positive Event- Positive Reaction/ Negative Event-Negative Reaction	- 0.01	.96
Positive Event-Negative Reaction/ Negative Event- Positive Reaction	0.23	<.0001*
Positive Event-Negative Reaction/ Negative Event- Negative Reaction	-0.04	.45
Negative Event-Positive Reaction/ Negative Event-Negative Reaction	-0.28	<.0001*

* Indicates significant *p* values.

Correlations. The overall *D*-score from each IRAP, as well as the difference scores between the two IRAPs, were subjected to correlational analyses with the questionnaires (i.e. 25 correlations for each IRAP and 25 for the IRAP difference score, not shown). A correlation matrix showed six significant correlations. The self-focused IRAP correlated positively with the EOSS close relationships-absent subscale, $r(30) = 0.47$, $p < 0.01$. The specified other-focused IRAP correlated negatively with the ECR-RS attachment-avoidance subscale for one's mother, $r(30) = -0.42$, $p < 0.05$, one's father, $r(30) = -.391$, $p < 0.04$, and with attachment-anxiety of best friend, $r(30) = -0.38$, $p < 0.04$, and again with the overall PFI score, $r(30) = -0.42$, $p < 0.03$. The difference score between the IRAPs correlated negatively with attachment-avoidance with romantic partner, $r(30) = -0.38$, $p < 0.04$. None of these correlations remained significant after Bonferroni corrections. Finally, the overall *D*-scores from the two IRAPs were correlated and this single correlation proved to be non-significant ($r = 0.284$ and $p = 0.129$).

Summary. Similar to Experiment 3, the difference for IRAP type was again non-significant, suggesting that when the other is specified, participants respond in a broadly similar fashion to both self- and other-focused IRAPs. As a result, the lack of

difference between the IRAPs observed in Experiment 3 was unlikely due to the focusing tasks (because in Experiment 4, these tasks were not included). However, the overall *D*-scores from the two IRAPs did not correlate significantly (in Experiment 3 the correlation was significant), suggesting no overlap in the two measures. Only six of the 75 correlations proved to be significant but again these did not remain so after Bonferroni corrections.

At this point in the research, the first two experiments had yielded significant differences between the IRAPs when other was unspecified, but non-significant differences were recorded across the third and fourth experiments when other *was* specified. In these latter experiments, the specified other was identified as the person to whom each participant was closest, and thus the variable of closeness, but not similarity, was manipulated. It is possible, therefore, that the lack of significant differences between the self- and other-IRAPs was not due to closeness but to similarity, if one assumes that in general people are close to others who are similar to them. Experiment 5, therefore, replicated Experiment 4, but asked participants to identify an individual to whom they were close but who they perceived to be very different: would targeting difference in this way produce a significant difference between the two IRAPs?

Experiment 5

Method

Participants. Thirty-two participants were recruited for Experiment 5, 24 females, eight males. Ages ranged from 18-32 years (M = 20.82). All participants were recruited through random convenience sampling from the Department of Experimental, Clinical and Health Psychology, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. The self-focused IRAP was identical to that used in Experiments 1-4. The other-focused IRAP was similar to Experiments 3 and 4, except that the name presented in each of the statements was now of a specified other (e.g.

"Sarah") who the participant considered to be very different from them. All questionnaires were identical to those used in Experiments 2-4.

Specified other-focused IRAP. The specified other-focused IRAP was similar to Experiments 3 and 4, except that each statement now included the name of an individual to whom the participant was close, but who was nonetheless perceived to be very different. Participants were instructed as follows "I want you to think of a person in your life that you are close to, but who you also consider to be very different from you. For example, this could be a friend who has very different interests from you."

Self-focused IRAP. The self-focused IRAP was identical to all previous experiments.

Questionnaires. All questionnaires were identical to Experiments 2-4.

Procedure. The procedure was similar to Experiment 4, except that the specified-other IRAP involved the name of an individual to whom the participant was close, but who was perceived to be very different.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations of each questionnaire and relevant sub-scales is provided in Appendix G. The overall pattern of results from the questionnaires was broadly similar to previous experiments.

IRAP data. One participant failed to complete the specified other-focused IRAP because of limited time and one failed to maintain criteria on the self-focused IRAP. Thus, the final number of participants included in the analyses was $N = 30$.

D_{IRAP}-scores. The mean D_{IRAP}-scores for each trial type are presented in Figure 6. Positive scores were recorded for all four trial types, with the weakest on *Negative Event-Positive Reaction*. For each of the four trial types, the difference between the two IRAPs appeared relatively modest. A 2x4 mixed repeated measures ANOVA produced a main effect for trial type, $F(1, 29) = 5.33$, $p < .001$, $\eta_p^2 = .16$, but not for IRAP type ($p > .5$), or for the interaction $p > 0.2$). Post-hoc (Fisher's PLSD) comparisons indicated

that *Negative Event-Positive Reaction* differed significantly from *Positive Event-Positive Reaction* and *Negative Event-Negative Reaction* (see Table 7).

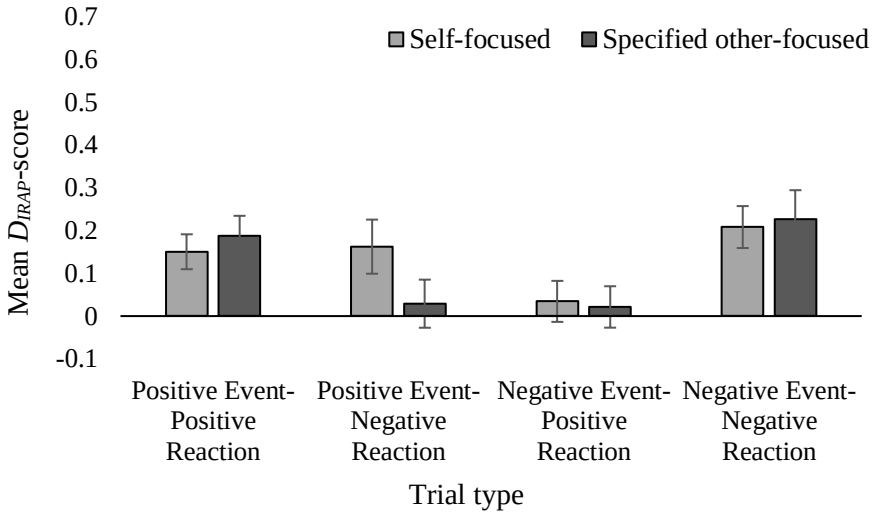

Figure 6. Mean D_{IRAP}-scores on the self-focused IRAP trial types and specified other-focused trial types in Experiment 5. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores indicate history-inconsistent responding.

Table 7

Experiment 5 Fisher's PLSD Comparisons

Trial type	**Mean Difference**	***P***
Positive Event-Positive Reaction/ Positive Event–Negative Reaction	.03	.51
Positive Event- Positive Reaction/ Negative Event-Positive Reaction	.12	.01*
Positive Event- Positive Reaction/ Negative Event-Negative Reaction	- .06	.2
Positive Event-Negative Reaction/ Negative Event- Positive Reaction	.09	.05
Positive Event-Negative Reaction/ Negative Event- Negative Reaction	-.09	.05
Negative Event-Positive Reaction/ Negative Event-Negative Reaction	-.18	<.0001*

* Indicates significant *p* values.

Correlations. The overall *D*-score from each IRAP, as well as the difference scores between the two IRAPs, were subjected to correlational analyses with the questionnaires (i.e. 25 correlations for each IRAP and 25 for the IRAP difference score, not shown). A correlation matrix showed only one significant correlation. The self-focused IRAP correlated negatively with avoidant-attachment with one's mother, $r(30) = -.38$, $p < 0.04$, but again this was non-significant after Bonferroni correction. In addition, the overall *D*-scores from the two IRAPs were correlated and this single correlation proved to be significant ($r = 0.410$ and $p = 0.023$).

Summary. Similar to Experiments 3 and 4, the difference for IRAP type was again non-significant, suggesting that when the other is known by the participant, but perceived as different from them, participants respond in a broadly similar fashion to both self- and other-focused IRAPs. The overall *D*-scores from the two IRAPs also correlated, suggesting some overlap in the two measures. Only one of the 75 correlations proved to be significant but did not remain so after a Bonferroni correction, and thus once again this result should be interpreted very cautiously. In Experiments 3 and 4, the specified other was identified as the person to whom each participant felt closest, but in Experiment 5 specified other was identified as the person to whom each participant felt closest but perceived them as different. It was assumed that people would feel positive toward others to whom they are close, and thus the variable of valence (i.e. positivity) was kept constant across these three experiments. It is possible, therefore, that the lack of difference in overall IRAP effects was not due to closeness and similarity, but to positive valence, if we assume that in general people feel positive toward those to whom they are close. Experiment 6, therefore, replicated Experiment 5, but asked participants to identify an individual who they knew but disliked: would targeting valence in this way produce a significant difference between the two IRAPs?

Experiment 6

Method

Participants. Thirty-four participants were recruited for Experiment 6, 30 females, four males. Ages ranged from 18-50 years ($M = 21.29$). All participants were

recruited through random convenience sampling from the Department of Experimental, Clinical and Health Psychology, Ghent University participant pool and were paid an hourly rate of 10 euro.

Materials and apparatus. The self-focused IRAP was identical to that used in Experiments 1-5. The other-focused IRAP was similar to Experiments 3-5, except that the name presented in each of the statements was now of a specified other (e.g. "Sam") who the participant disliked. All questionnaires were identical to those used in Experiments 2-5.

Specified other-focused IRAP. The specified other-focused IRAP was similar to Experiments 3-5, except that each statement now included the name of an individual who was known to the participant, but who they disliked. Participants were instructed as follows "I want you think of a person in your life who you often interact with, but who you don't like. For example, this could be a classmate, work colleague or family member who you do not get along with."

Self-focused IRAP. The self-focused IRAP was identical to all previous experiments.

Questionnaires. All questionnaires were identical to Experiments 2-5.

Procedure. The procedure was similar to Experiment 5, except that the specified-other IRAP involved the name of an individual to whom the participant was close, but who they disliked.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations of each questionnaire and relevant subscales is provided in Appendix H. The overall pattern of results from the questionnaires was broadly similar to previous experiments.

IRAP data. One participant failed to complete the specified other-focused IRAP because of limited time and two failed to maintain criteria on the specified other-focused IRAP. One participant failed to maintain criteria on the self-focused IRAP. Thus, the final number of participants included in the analyses was $N = 30$.

D_{IRAP}-scores. The mean D_{IRAP}-scores for each trial type are presented in Figure 7. Positive scores were recorded for seven of the eight trial types, with the only negative score recorded on *Negative Event-Positive Reaction* of the specified other-focused IRAP. For each of the four trial types, the difference between the two IRAPs appeared relatively modest. A 2x4 mixed repeated measures ANOVA produced a main effect for trial type, $F(1, 29) = 7.28$, $p < .001$, $\eta_p^2 = .2$, but not for IRAP type ($p > .09$), or for the interaction ($p > 0.46$). Post-hoc comparisons (Fisher's PLSD) indicated that *Negative Event-Positive Reaction* differed significantly from *Positive Event-Positive Reaction, Positive Event-Negative Reaction,* and *Negative Event-Negative Reaction* (see Table 8).

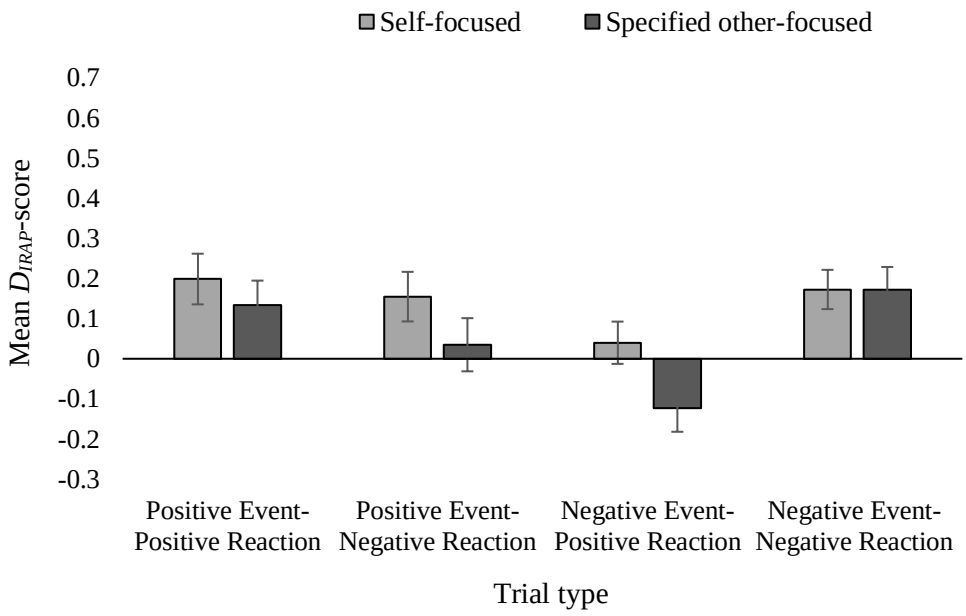

Figure 7. Mean D_{IRAP}-scores on the self-focused IRAP trial types and specified other-focused trial types in Experiment 6. Positive D_{IRAP}-scores indicate history-consistent responding and negative D_{IRAP}-scores indicate history-inconsistent responding.

Table 8

Experiment 6 Fisher's PLSD Comparisons

Trial type	Mean Difference	*P*
Positive Event-Positive Reaction/ Positive Event–Negative Reaction	.07	.18
Positive Event- Positive Reaction/ Negative Event-Positive Reaction	.21	<.01*
Positive Event- Positive Reaction/ Negative Event-Negative Reaction	- .0006	.90
Positive Event-Negative Reaction/ Negative Event- Positive Reaction	.14	<.02*
Positive Event-Negative Reaction/ Negative Event- Negative Reaction	-.08	.14
Negative Event-Positive Reaction/ Negative Event-Negative Reaction	-.21	<.01*

* Indicates significant *p* values.

Correlations. The overall *D*-score from each IRAP, as well as the difference scores between the two IRAPs, were subjected to correlational analyses with the questionnaires (i.e. 25 correlations for each IRAP and 25 for the IRAP difference score, not shown). The correlation matrix yielded five significant correlations. The self-focused IRAP correlated negatively with avoidant-attachment to one's romantic partner, $r(30) = -.41$, $p < 0.03$, and the control of a person close to them over the experience of self, $r(30) = - .43$, $p < 0.02$. The specified-other IRAP correlated negatively with the control of a person close to them over the experience of self, $r(30) = - .39$, $p = 0.03$, closeness to other people, $r(30) = - .41$, $p = 0.03$, and with warmth toward oneself , $r(30) = -.46$, $p < 0.01$. Once again, none of these correlations remained significant after Bonferroni corrections. Finally, the overall *D*-scores from the two IRAPs were correlated and this single correlation proved to be non-significant ($r = 0.103$, $p = 0.59$).

Summary. Similar to Experiments 3-5, the difference for IRAP type was again non-significant, suggesting that when the other is negatively valenced, participants again respond in a broadly similar fashion to both self- and other-focused IRAPs. On balance, the overall *D*-scores from the two IRAPs did not correlate significantly, suggesting limited overlap in the two measures. Once again, only a small number of the

75 correlations proved to be significant, but again these did not remain so after Bonferroni corrections.

Discussion

The overarching purpose of the current series of experiments was to develop IRAPs that would clearly differentiate responding from the perspective of self versus other. Previous research appeared to demonstrate deictic relational responding using the IRAP (i.e. responding to the characteristics of self versus other within an IRAP), but to date clear evidence of responding from the perspective of self versus from the perspective of another has not been shown. To demonstrate such an effect would seem to require inserting statements into an IRAP that ask participants to confirm and deny how the self versus others would react to similar events. Achieving this objective required that relatively complex statements be employed, and for this reason a natural language version of the IRAP was used. The results from the first two experiments reported here indicated that there were significant differences between the self- versus other-focused IRAPs, when the other remained unspecified. The remaining four experiments, however, indicated that when the other was specified there was limited evidence that performances on the two IRAPs differed significantly. The correlational analyses between the IRAPs and the self-report measures yielded very few significant effects (and none at all after Bonferroni corrections). Finally, the correlational analyses between the two IRAPs in each experiment were significant in some cases (Experiments 2, 3 and 5), but not others (Experiments 1, 4 and 6). Overall, the relatively small number of uncorrected correlations between the IRAPs and the explicit measures, and the complete absence of any corrected significant correlations, suggest that such effects should be taken with extreme caution.

Overall, the current findings could be seen as encouraging because each expeirment produced perfromances that would be deemed consistent with the pre-experiemental histories of the participants. That is, the IRAP effects for most part, were in the predicted direction (e.g. participants confirmed more rapidly than they denied that when positive events occur, they react positively, but when negative events occur, they

react negatively. On balance however, the results could be seen as somewhat disappointing because we failed to find strong evidence of perspective-taking when other was specified, ay least in terms of different perfromances across the two IRAPs, or in correlations among the IRAPs and self-report measures. The IRAPs produced predictable outcomes and thus it seems important, going forward, to ask why the IRAPs did not distinguish in a clear and consistent way between the perspective of self versus other. One possibility is that the IRAPs simply tapped into 'sense-making'. For example, confirming, rather than denying that positive events evoke positive reactions in the self and others makes sense in natural language. The only caveat to this interpretation is that significant differences between the IRAPs were observed in Experiments 1 and 2 when the other was unspecified, but not in Experiments 3-6 when other was specified. Perhaps, therefore, there was some sensitivity to self versus other, but the use of complex statements in the IRAPs potentially undermined or reduced the impact of deictic relational responding per se. That is, in presenting such complex stimuli or networks in the IRAP, participants more or less interpreted the task as a sense-making or problem-solving task, in which the self versus other had little or no impact, particularly when the other was specified. The challenge going forward, therefore, is to develop IRAPs that maintain the deictic functions of self and other, in the context of perspective-taking rather than simple sense-making.

Chapter 4

Exploring Perspective-taking in a False Belief IRAP using Pictures of Self and Other[6]

[6] Experiment 1 of this chapter is based on Kavanagh, D., Matthyssen, Barnes-Holmes, Y., Barnes, Holmes, D., McEnteggart C. & Vastano, R. (2019). Exploring the use of pictures of self and other in the IRAP: Reflecting upon the emergence of differential trial type effects. *International Journal of Psychology and Psychological Therapy, 19(3),* 323-336

In pursuing IRAP research that has focused on assessing responses to self versus other, an issue arose concerning the nature of the stimuli we were using, which we hoped to begin to address in the empirical work outlined in the current chapter. Experiment 1 was deemed to be exploratory in seeking to determine if it was feasible to use matched pictures of self versus another across two IRAPs. Specifically, we sought to determine if the IRAPs would produce intuitively predictable effects when pictures of self and matched pictures of another were inserted into the procedure (see below). The broader motivation for the first experiment naturally stemmed from the programme of research presented in the two previous chapters that sought to develop the IRAP as a measure of perspective-taking (see also Barbero-Rubio et al., 2016). In each of the studies we conducted thus far, and in the original work by Barbero-Rubio et al., the stimuli that were presented in the IRAP itself all consisted of words or statements that pertained to the characteristics of the participant or another individual. Typically, these self-related terms involved using the participant's name or words such as "I", "my" or "me", whilst other-related terms involved using another's name or words such as "they", "others" or "them". Although this research yielded interesting findings, we had concerns that the use of such stimuli might allow for some ambiguity in how they were interpreted by participants. For example, when the on-screen stimulus was "I" the assumption was that the participant would interpret this as referring to self, rather than to the computer or another person. In general, it appeared that this assumption was upheld, but of course room for ambiguity remained. One way in which we thought we could remove this potential ambiguity was to insert a picture of the participant into the IRAP itself, so that on those trials when the picture appeared, it would be clearly interpreted as a self-related stimulus.

In the first experiment presented in the current chapter, participants were asked to bring an electronic picture of their face which could be inserted into the IRAP at the beginning of the session. They were also asked to bring a picture of another person, who they considered to be relatively similar to themselves in terms of age, gender and general facial features. Each participant was then exposed to two IRAPs, with one IRAP containing the picture of themselves and the other containing the matched picture of the

matched other. Both IRAPs presented the same pictures of pens as the contrast category. The two IRAPs were thus very similar in that they both involved confirming and disconfirming if the on-screen pictures contained faces versus pens. We expected to find IRAP effects that were consistent with previously-established verbal relations. That is, participants would confirm more readily than deny that a picture of a face was a face, and that a picture of a pen was a pen, as well as denying more readily than confirming that a picture of a pen was a face and that a picture of a face was a pen. Participants were also asked to complete the six questionnaires that were employed in Experiments 2-6 of the previous chapter. Given the exploratory nature of the study, we made no formal predictions concerning the extent to which the two IRAPs (self-picture versus other-picture) would produce different outcomes or how performance on those IRAP might correlate with responses to the questionnaires.

Experiment 1

Method

Participants. Thirty-four participants were recruited for Experiment 1, 29 females and five males. Participants ranged from 17-45 years (M = 21.4) and were recruited through random convenience sampling from the Ghent University participant pool. Each participant was paid an hourly rate of 10 euro. Given that this was the first study in which pictures of the self versus others had been inserted into an IRAP, it was not possible to conduct a meaningful power analysis. Nevertheless, the general strategy for recruiting numbers of participants was similar to that previously described in Chapter 2.

Materials and apparatus. Experiment 1 comprised of two IRAPs, a self-picture IRAP and an other-picture IRAP, and included the same six questionnaires previously employed in Experiments 2-6 of Chapter 2.

Picture stimuli used in the IRAPs. The face picture stimuli were collected for both IRAPs prior to the experiment. Participants were asked to bring to the experiment two pictures; a picture of themselves that they liked and a picture of an unknown other who they considered to be similar in looks to themselves (i.e. same gender, age, hair

colour, skin colour and eye colour). These pictures were included in the self-picture and other-picture IRAPs, respectively.

Self-picture IRAP. The self-picture IRAP presented a total of six label stimuli on the top of the screen; three of the labels were words pertaining to faces (i.e. “face”, “head” and “person”), with the remaining three words pertaining to pens (i.e. “bic”, “pen” and “stylo”, see Table 1). The target stimuli consisted of four images presented in the centre of the screen, one picture depicted the individual participant’s face and the other three pictures depicted three different pens. The response options “Yes” and “No” were presented at the bottom left- and right-hand corners. The four trial types were denoted as: *Face words-Self picture, Face words-Pen pictures, Pen words-Self picture* and *Pen words-Pen pictures* (see Figure 1).

Table 1
Experiment 1, label stimuli for the self-picture and other-picture IRAPs

Face words	**Pen words**
Person	Pen
Face	Stylo
Head	Bic

Note: Stimuli were presented to participants in Dutch.

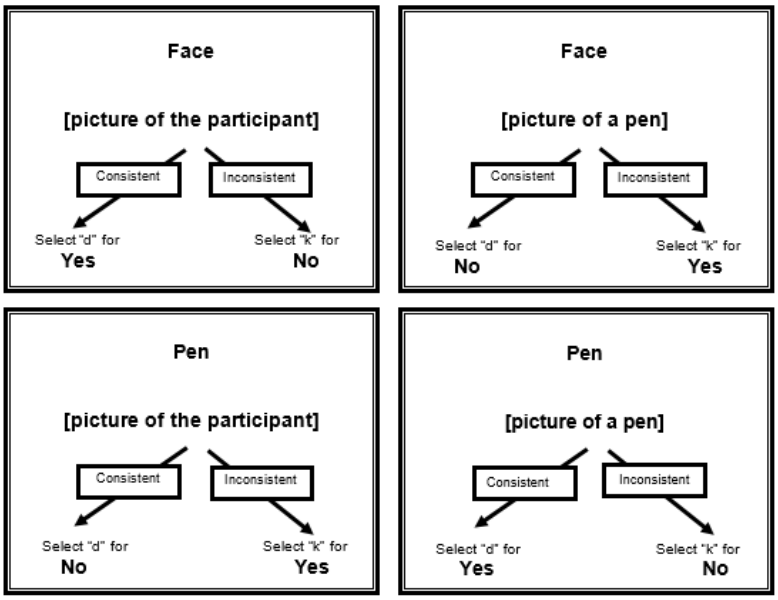

*Figure 1***.** Examples of the four trial types in the self-picture IRAP: *Face words-Face picture*, *Face words-Pen pictures*, *Pen words-Face picture* and *Pen words-Pen pictures.* The words *Consistent* and *Inconsistent* were not shown on-screen.

Other-picture IRAP. The other-picture IRAP was similar to the self-picture IRAP, except that the four target pictures now depicted the face of another person (rather than the face of each participant), as well as the three different pens. The four trial types were denoted as: *Face words-Other picture*, *Face words-Pen pictures*, *Pen words-Other picture* and *Pen words-Pen pictures* (see Figure 2).

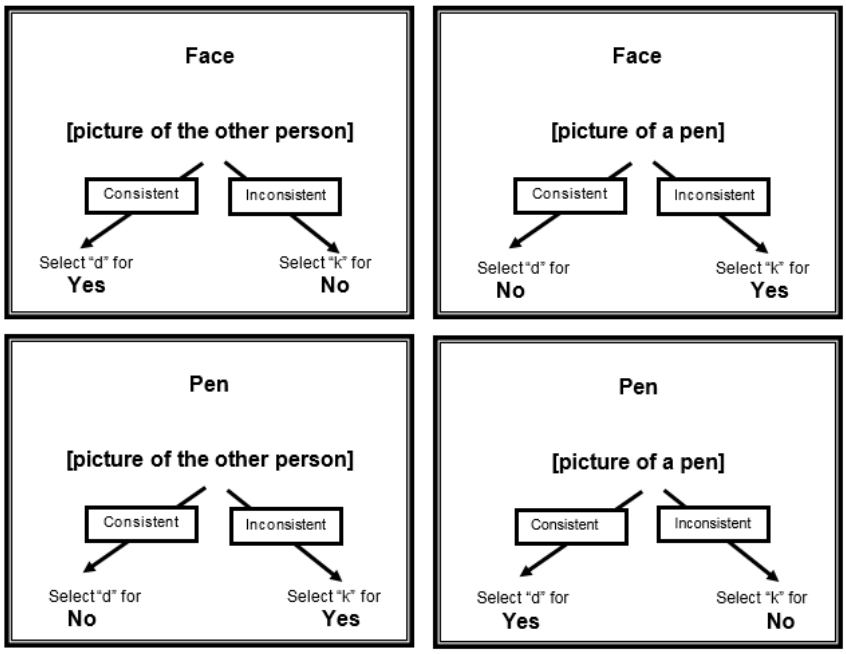

*Figure 2***.** Examples of the four trial types in the other-picture IRAP: *Face words-Face picture, Face words-Pen pictures, Pen words-Face picture* and *Pen words-Pen pictures.* The words *Consistent* and *Inconsistent* were not shown on-screen**.**

Questionnaires. All questionnaires included in Experiment 1 were identical to those presented in Experiments 2-6 of Chapter 2.

Procedure. Experiment 1 took place on an individual basis in sound-proof cubicles at the Department of Experimental, Clinical and Health Psychology, Ghent University. Each participant was exposed to the self-picture and other-picture IRAPs, with the order of each counterbalanced across participants. Participants thereafter completed the six questionnaires, always presented in the same order as Experiments 2-6 of Chapter 2.

Self-picture IRAP. The self-picture IRAP comprised of a maximum of four pairs of practice blocks, followed by three pairs of test blocks. On each trial, the label (face-related word or pen-related word) appeared at the top of the screen, with a target (picture of participant's face or picture of a pen) in the centre. The response options and

between block instructions were presented in a similar manner to those described in Chapters 2 and 3.

Trials were presented in a quasi-random order within each block, such that each of the four trial types appeared eight times within each block (32 trials were presented in a block). Consistent blocks required responding that was in accordance with pre-experimental verbal relations: *Face words-Face picture/Yes, Face words-Pen pictures/No, Pen words-Face picture/No* and *Pen words-Pen pictures/Yes.* Inconsistent blocks required the opposite: *Face words-Face picture/No, Face words-Pen pictures/Yes, Pen words-Face picture/Yes* and *Pen words-Pen pictures/No.* The presentation of consistent and inconsistent blocks was counterbalanced across participants.

Other-picture IRAP. The other-picture IRAP was similar to the self-picture IRAP, except that it contained the picture of the other face rather than the face of the participant.

Questionnaires. Participants completed the six questionnaires immediately after completing the two IRAPs.

Ethical Considerations. All aspects of Experiments 1-3 in the current chapter were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All procedures received ethical approval from the Ethical Commission, Faculty of Psychology and Educational Sciences, Ghent University (Approval date: January 4th 2018, Reference: 2017/109/Deirdre Kavanagh). Informed consent was obtained from all individual participants.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations for the six questionnaires is provided in Appendix I. The mean score on the Self-warmth Thermometer was around the mid-way point at 56.72 (/100) indicating relatively average warmth toward self. The overall PFI scores were relatively high, indicating high

psychological flexibility. The overall CAPE and sub-scale scores were relatively low, indicating low psychotic-like symptoms. The ECR-RS scores were also relatively low in terms of both attachment-anxiety and attachment-avoidance. The overall EOSS and sub-scale scores were low, indicating low control by others over the experience of self. The IOS scores for best friend were higher than for other people, suggesting a closer relationship in this regard. Nothing unusual or unexpected, therefore, emerged from the questionnaires, given the use of a non-clinical sample.

IRAP data. In order to pass the practice blocks and advance to test blocks, participants were required to maintain an accuracy level of ≥ 80% correct and a median latency of ≤ 2,000 ms. Based on the latency and accuracy criteria, one participant failed the practice blocks on the self-picture IRAP and was excluded from further analysis. Exclusion criteria were also applied to the test blocks, such that participants were required to maintain an accuracy level of ≥ 80% correct and a median latency ≤ 2,000 ms on two of the three successive pairs of the twelve test blocks. Of the remaining participants, one failed to maintain criteria across the test block pairs in the self-picture IRAP, again these data were excluded from further analysis (final N = 32). Consistent with Nicholson and Barnes-Holmes (2012), if participants failed to maintain criteria for one or both test blocks from a given pair, the data from those two blocks were excluded and the data from the remaining two test-block pairs were analysed. These exclusions were applied to four of the final data sets.

Consistent with the previous chapters, D_{IRAP}-scores for both IRAPs were calculated for each of the four trial types, such that positive D_{IRAP}-scores during consistent blocks indicated responding "Yes" more quickly than "No" on *Face words-Face picture* and *Pen words-Pen pictures* trial types, and responding "No*"* more quickly than "Yes*"* on *Face words-Pen pictures* and *Pen words-Face picture* trial types. Negative D_{IRAP}-scores indicated the opposite pattern: responding "No" more quickly than "Yes" on *Face words-Face picture* and *Pen words-Pen pictures* trial types, and responding "Yes" more quickly than "No" on *Face words-Pen pictures* and *Pen words-Face picture* trial types.

The mean D_{IRAP}-scores and standard errors for each trial type are presented in Figure 3. Positive scores were recorded for all four trial types, with the weakest on *Pen words-Face picture* in both of the IRAPs. For each of the four trial types, the difference between the two IRAPs appeared relatively modest. A 2x4 mixed repeated measures ANOVA produced a main effect for trial type, $F(1, 31) = 17.14$, $p < .0001$, $\eta_p^2 = .36$, but not for IRAP type ($p > .5$), or for the interaction ($p > 0.1$). Post-hoc comparisons, with the trial type effects collapsed across the self- and other-picture IRAPs (see Table 2), indicated that each trial type differed from every other trial type, except for the comparison between *Faces words-Pen pictures* and *Pen words-Pen pictures.* Four one-sample *t*-tests indicated that the effects were significantly different from zero (*ps* < .05) for all four trial types.

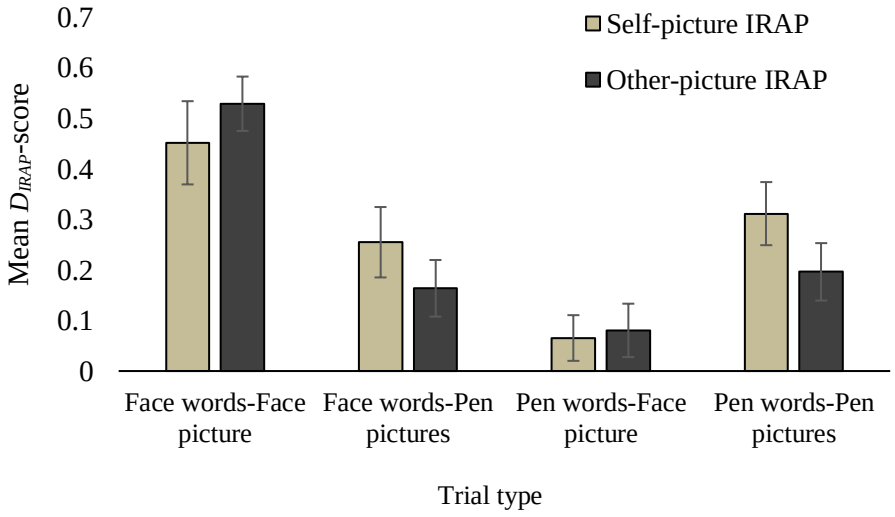

Figure 3. Mean D_{IRAP}-scores on the self-picture and the other-picture IRAP trial types. Positive D_{IRAP}-scores indicate history-consistent responding.

Table 2
Experiment 1, Fisher's PLSD Comparisons

Trial type	Mean Difference	*P*
Face words-Face picture/ Face words-Pen pictures	.28	<.001*
Face words-Face picture/ Pen words-Face picture	.418	<.001*
Face words-Face picture/ Pen words-Pen pictures	.236	<.001*
Face words-Pen pictures/ Pen words-Face picture	.136	.024*
Face words-Pen pictures/ Pen words-Pen pictures	-.045	.454
Pen words-Face picture/ Pen words-Pen pictures	-.181	.003*

* Indicates significant *p* values.

Correlations. Given that there was no significant main effect for IRAP type or interaction with trial type, a single overall *D*-score (the mean of the four trial types) was calculated for each IRAP, and then subjected to correlational analyses with the questionnaires (a total of 25 correlations for each IRAP). Neither set of correlations proved to be significant (all *p*s > .05).

Summary. The primary objective of Experiment 1 was to determine if it was feasible to use matched pictures of self versus another across two IRAPs. In both IRAPs, the contrast categories were pen-related stimuli. In general, the pattern of effects indicated that pictures of faces could be incorporated into the IRAP. Specifically, with the exception of one trial type, the effects recorded were consistent with what might be expected based on previously-established verbal relations. That is, participants tended to confirm more readily than to deny that a picture of a face was a face and that a picture of a pen was a pen. They also tended to deny more readily than confirm that a picture of a pen was a face and that a pen-related word was a picture of a face. There was no significant difference between the self- and other-IRAPs, at least in terms of the *D*-scores.

In previous chapters, and in research by others (see Barbero-Rubio et al., 2016), differences have been found in reactions to self versus others, using IRAPs. No clear differences emerged in the current experiment between self versus others. On balance, differences between self and other IRAPs were not observed in the experiments reported in the previous chapter when the 'other' was specified rather than unspecified. Insofar as the 'other' was specified in the current experiment, because a specific picture of another individual was employed in the other-picture IRAP, the current findings could be seen as generally consistent with the findings reported thus far. On balance, it could be argued that it was highly unlikely that in the current experiment any differences between the self- versus other-picture IRAPs would emerge because responding in the IRAPs was controlled more by the presence of a face versus a pen, rather than by the presence of one's own face versus the face of another. In any case, the results from Experiment 1 are promising in indicating that pictures of self and other can be successfully employed in an IRAP, and this finding was used as a basis for selecting self- versus other-related stimuli in the following two experiments reported in the current chapter.

As noted earlier, the general rationale for the six experiments reported in Chapter 3 was to develop an IRAP that was sensitive to perspective-taking rather than simply tracking deictic relations. In an effort to achieve this goal NL-IRAPs were employed, which allow relatively complex statements to be presented that appeared to require responding from a self-perspective in one IRAP and another's perspective on a second IRAP. Unfortunately, however, the complexity of the statements employed in those IRAPs may have undermined the extent to which 'genuine' perspective-taking was being assessed by the IRAPs. Specifically, it was possible that the IRAPs were treated largely as problem-solving tasks rather than contexts for taking the perspective of self versus other. On this basis, we aimed to identify relatively simple stimuli that could be inserted into an IRAP that would clearly distinguish between self versus other. This was the purpose of the next experiment.

Experiment 2

At this point, we faced the challenge of deciding how to assess perspective-taking using IRAPs that presented pictures of self and other without adding complex statements that may again undermine perspective-taking (as in the NL-IRAPs). The strategy we adopted thus involved avoiding the use of complex statements within the IRAP, but instead presenting vignettes that required perspective-taking just prior to the completion of an IRAP. In other words, the vignettes were designed to produce different patterns for self versus other trial types within a single IRAP. The vignette was based on one of the most widely used formats for exploring perspective-taking in the ToM literature, namely the Change in Location task. This task was designed to assess the attribution of false beliefs (see Perne et al., 1989). Specifically, the false belief vignette comprised of a written paragraph that described a scenario involving the participant and the other person depicted in the other-face picture. In this scenario, the participant observed that items in a box switched locations when the other person left the room. A belief IRAP was then presented that required the participant to respond to what they thought was in the box and what they thought that the other person thought was in the box. Given that the items had been switched when the other person had the left the room, the self and other should differ in terms of what they believed to be in the box. The key question was, would the observed IRAP effects be consistent with the false belief vignette? A control vignette was presented to half of the participants, in which there was no exchange of the items in the box and therefore no false belief attribution was required. At this stage, we were interested in determining if differential patterns of responding would be observed in the IRAP performances across the two conditions (false belief versus control). Participants were also asked to complete the six questionnaires that were employed in Experiments 2-6 of the previous chapter. Given the exploratory nature of the study, we made no formal predictions concerning the extent to which the IRAP in the two conditions would produce different outcomes or how performance on those IRAPs might correlate with responses to the questionnaires.

Method

Design. There were three stages in Experiment 2: 1. Familiarisation IRAP; 2. Condition vignette and belief IRAP; 3. Questionnaires. Experiment 2 also comprised of two vignette conditions: (i) False Belief and (ii) Control.

Participants. Seventy-five participants were recruited, 63 females and 12 males, all aged between 17 and 34 years (M = 21.27). The recruitment process was similar to that described in the previous experiment.

Materials and apparatus. Experiment 2 involved two computer-based tasks presented on standard computers, the familiarisation IRAP and the belief IRAP. The belief IRAP involved two pictures of faces selected by each participant, one picture presented the participant's face, while the other picture was the face of a stranger considered by the participant to be similar in looks to themselves (i.e. same gender, age, hair colour, skin colour and eye colour). Two short vignettes were also constructed for current purposes, with each pertaining to one of the two conditions (False Belief or Control). Two bespoke questionnaires assessed performance strategies and perceived physical similarities between the picture of self and the picture of other. The same six questionnaires previously employed in Experiments 2-6 of Chapter 2 were also included.

Picture stimuli used in the IRAP. The face picture stimuli were collected in a manner identical to the previous experiment.

Familiarisation IRAP. The familiarisation IRAP did not contain stimuli relevant to perspective-taking and was employed simply to familiarise participants with the procedure, because no practice blocks were presented in the subsequent belief IRAP. The stimuli presented in the familiarisation IRAP were similar to those described in Experiment 1 of Chapter 2.

Belief IRAP. Each trial in the belief IRAP presented the picture of the participant or the picture of the other person as a label stimulus at the top of the screen. The target stimuli comprised 12 statements, with one presented on each trial. Six of the statements referred to beliefs about a scarf (e.g. "thinks there is a scarf") and six referred to beliefs about gloves (e.g. "thinks there is a glove"), see Table 3. The response options "Yes" and "No" were presented at the bottom left- and right-hand corners. The four trial

types were denoted as: *Self-Glove* (participant's picture-statement about a glove); *Self-Scarf* (participant's picture-statement about a scarf); *Other-Glove* (picture of other-statement about a glove); and *Other-Scarf* (picture of other-statement about a scarf); see Figure 4.

Table 3
Experiment 2, target stimuli for the belief IRAP

Glove	Scarf
Thinks there is a glove	Thinks there is a scarf
Believes there is a glove	Believes there is a scarf
Says there is a glove	Says there is a scarf

Note: Stimuli were presented to participants in Dutch.

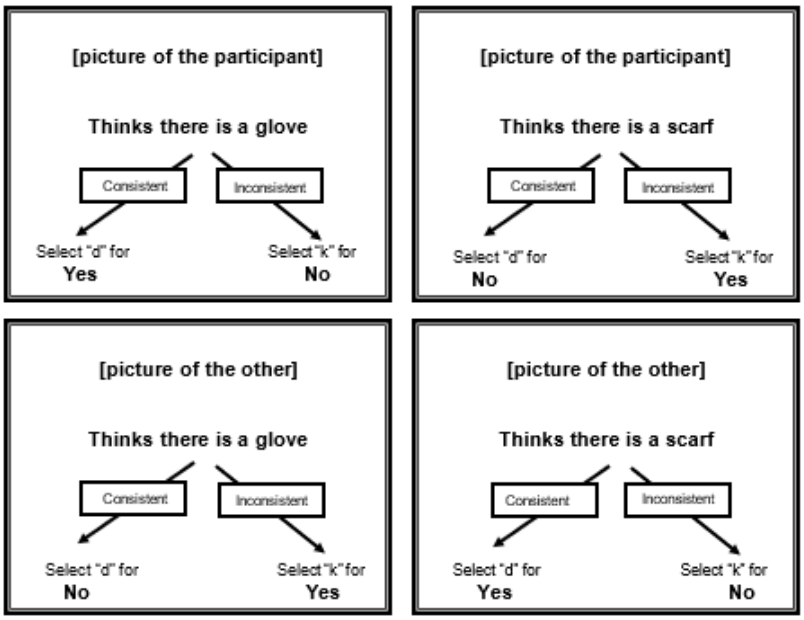

Figure 4. Examples of the four trial types in the belief IRAP: *Self-Glove, Self-Scarf, Other-Glove* and *Other-Scarf.* The words *Consistent* and *Inconsistent* were not shown on-screen.

Condition Vignettes. The two vignettes were presented on-screen in a word document. Each comprised of a written paragraph that described one of two scenarios involving the participant and the other person depicted in the other-face picture, followed by questions that served to check that participants had read and understood the vignette.

False Belief Vignette. The false belief vignette was based on a ToM Change in Location task designed to assess the attribution of false beliefs (see Perner, et al., 1989). The English translation of the vignette presented in Dutch is as follows:

> *The following sentences describe a scenario that involves you and the person in the second picture that you brought with you today. I want you to imagine that you and this person are in a room. In front of you both is a box. When you both open the lid of the box, you both see together that there is a scarf in the box and then you place the lid back on the box again. At this point, the other person leaves the room. When they are no longer in the room (but you still are) the scarf is removed from the box and is replaced with a glove. The lid is then put back on the box. At this point, the other person now returns to the room.*

After reading he vignette, participants were required to indicate which of the following statements best described what happened in the scenario presented: 1. “I stayed in the room and the other person left the room”; 2. “Both of us stayed in the room”; or 3. “The other person stayed in the room and I left the room”[7]. Participants were then instructed as follows: “Please remember the details of the scenario you read

[7] The researchers checked responses to these questions after the experiment to ensure that they were consistent with the vignette; the data were lost for five participants due to a software overwriting error, but consistency was obtained for all other participants.

above as you will require information from this scenario to successfully complete the next part of the experiment."

Control Vignette. The control vignette described a similar scenario to the false belief vignette, but critically there was no change in the location of items. The English translation is as follows:

> *The following sentences describe a scenario that involves you and the person in the second picture that you brought with you today. I want you to imagine that you and this person are in a room. You have a box in front of you and the other person has a different box in front of them. When you both open the lid of the boxes, you both see together that you have a glove in your box and the other person has a scarf in their box. Then the lids are placed back on the boxes.*

The same three statements, and the instruction presented above, were used to ensure that participants understood the vignette.

Strategy Questionnaire. This questionnaire was designed to identify any strategies that participants may have engaged in to successfully complete the belief IRAP (see Appendix J for questionnaire translated to English). Specifically, participants were asked three questions concerning the potential influence of the preceding vignette on their responding during the belief IRAP, rated from 1 ("*not much"*) to 5 ("*a lot"*). Specifically, one question asked "How much of your responding on the computer task was influenced by the scenario that you read before and throughout the task?" The questionnaire also contained a single open-ended question regarding any strategy they may have used to complete the task. Participants, in the False Belief Condition only, were also asked about the degree of success they believed they had in taking the perspective of the other person, rated from 1 ("*not successful*") to 5 (*"very successful").*

Similarity Questionnaire. This questionnaire was designed to identify perceived similarities between the self- and other-picture (see Appendix J). Specifically, participants were asked a single question concerning overall perceived similarity

between self and other, rated from 1 ("*not similar at all*") to 5 ("*very similar"*). Another five questions pertained to perceived similarity in terms of 1. hair colour, 2. age, 3. eye colour, 4. skin colour and 5. facial expression, rated from 1 ("*not similar at all*") to 5 ("*very similar"*). Two final questions focused on the attractiveness of the self-picture and the other-picture, both rated from 1 ("*not attractive*") to 5 (*"very attractive").*

Questionnaires. All other questionnaires included in Experiment 2 were identical to those presented in the previous experiment.

Procedure. Experiment 2 comprised of three stages, with all participants completing those stages as follows: 1. Familiarisation IRAP; 2. Condition vignette and belief IRAP; and 3. Questionnaires (Strategy Questionnaire, Similarity Questionnaire, CAPE, PFI, IOS, ECR-RS, EOSS and the Self-warmth Thermometer, always presented in this order).

Stage 1: Familiarisation IRAP. The familiarisation IRAP was employed to establish competent performances on a simple word-based IRAP (*Fruits* vs. *Vegetables*) prior to completion of the belief IRAP. Participants were simply instructed to determine, based on individual trial feedback, what the task involved. The procedure for the familiarisation IRAP was largely identical to Experiments 1 and 2 of Chapter 2, in that it only presented practice blocks. However, in the current experiment participants could complete up to five sets of practice blocks within the familiarisation IRAP (rather than just three sets of blocks as in Experiments 1 and 2 of Chapter 2).

Stage 2: Condition Vignette and Belief IRAP. Following the familiarisation IRAP, participants completed the condition vignettes and belief IRAP. These were presented in a sequence, such that a condition vignette was presented before each of the three test block pairs of the belief IRAP (i.e. condition vignette-first test block pair-condition vignette-second test block pair-condition vignette-third test block pair). After reading the condition vignette, participants were required to indicate which statement out of three possible statements best described what happened in the scenario that they just read. Participants were then instructed as follows: "Please remember the details of the scenario you read above as you will require information from this scenario to

successfully complete the next part of the experiment." Participants were then presented with a test block pair from the Belief IRAP. The first block always required a response pattern that was deemed consistent with the vignette (e.g. participant's picture/"Thinks there is a glove"/"Yes").

On each trial of the belief IRAP, the label (participant's picture or other's picture) appeared at the top of the screen, with a target statement (belief about a scarf or belief about a glove) in the centre, and the two response options ("Yes" and "No") at the bottom. 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 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. The instruction "The previously correct and incorrect answers have been reversed" was presented between the first block and second block of each test block pair. Consistent blocks required responding that was in accordance with pre-experimental verbal relations: *Self-Glove/Yes, Self-Scarf/No, Other-Glove/No,* and *Other-Scarf/Yes*. Inconsistent blocks required the opposite: *Self-Glove/No, Self-Scarf/Yes, Other-Glove/Yes,* and *Other-Scarf/No*.

When participants completed a block of trials, the IRAP program delivered feedback on their performance during that block. The feedback consisted of a message informing them how accurately and how quickly they had responded. 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% correct and a maximum median latency of ≤ 2000 milliseconds on each block. A fixed set of three test block pairs was presented with no accuracy or latency criteria required 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.

Stage 3: Questionnaires. Participants completed the eight questionnaires in the following sequence: Strategy Questionnaire, Similarity Questionnaire, CAPE, PFI,

IOS, ECR-RS, EOSS and the Self-warmth Thermometer. The Strategy and Similarity Questionnaires were presented in paper format and all other questionnaires were presented to participants via computer using the program Psychopy (Peirce, 2007).

Results and Discussion

Questionnaire data. A summary of the means and standard deviations for questionnaires, divided according to the two conditions (False Belief and Control), are presented in Appendix K. The scores divided across the conditions did not appear to differ substantively. Independent *t*-tests indicated that there were no significant differences between conditions, except for one comparison; casual acquaintances-present sub-scale of the EOSS, $t(56) = 2.33$, $p = .02$ (all other *ps* > .3). Given the large number of comparisons (25), this single significant effect was considered to be a false positive.

The open-ended strategy question was also assessed ($N = 49$). The open-ended answers were read by one researcher who developed an initial coding frame to organise the data. These codes were then grouped into categories, according to how they were related. Following this, a second researcher independently reviewed the coding and categories developed by the first researcher. Any inconsistencies or issues raised by the second researcher were discussed and the categories were adjusted accordingly. A list of the final 11 categories is presented in Appendix L. The most common strategy recorded by participants was linking the object words and picture words together ($N = 25$). Ten participants reported that they focused only on the object word and did not read the full sentences. Five participants reported that they rehearsed the link between the object and person before starting an IRAP block.

IRAP data. Consistent with standard practice in IRAP research, mean response latencies for consistent and inconsistent blocks were initially divided according to trial-type and calculated for each participant. Based on the latency and accuracy criteria, eight participants failed to complete the familiarisation IRAP (and did not proceed to the belief IRAP). Exclusion criteria were also applied to the belief IRAP, such that

participants were required to maintain an accuracy level of ≥ 79% correct and a median latency ≤ 2,000ms on the *third block pair*.[8] Eleven participants failed to maintain these criteria, five from the False Belief Condition and six from the Control Condition. Their data were excluded from further analysis (False Belief, final N = 30; Control, final N = 30).

D_{IRAP}-scores. D_{IRAP}-scores for the belief IRAP were calculated for each of the four trial types, such that positive D_{IRAP}-scores during vignette-consistent blocks indicated responding "Yes" more quickly than "No" on *Self-Glove* and *Other-Scarf* trial types and responding "No*"* more quickly than "Yes*"* on *Self-Scarf* and *Other-Glove* trial types. Negative D_{IRAP}-scores indicated the opposite pattern: responding "No" more quickly than "Yes" on *Self-Glove* and *Other-Scarf* trial types and responding "Yes" more quickly than "No" on *Self-Scarf* and *Other-Glove* trial types.

The mean D_{IRAP}-scores and standard errors for each trial type, from the third test block pair, are presented in Figure 5. Positive scores were recorded for three of the four trial types, with negative scores recorded on *Self-Scarf* in both of the conditions. For each of the four trial types, the difference between the two IRAPs appeared relatively modest. A 2x4 mixed repeated measures ANOVA produced a main effect for trial type [$F(1, 58) = 8.65$, $p < .0001$, $\eta_\rho^2 = .29$], but not for condition ($p > .8$), or for the interaction ($p > 0.3$). Post-hoc comparisons, with the trial type effects collapsed across the two conditions (see Table 4), indicated that *Self-Scarf* differed from every other trial type. Eight one-sample *t*-tests indicated that the effects were significantly different from zero

[8] Initially, we planned to include analysed data from all three block pairs, or at least two block pairs, but 47 of the 75 participants failed to meet the performance criteria on the first and/or second test block pair. Thus, only data for the third block pair were analysed and, even then, 17 participants failed to meet the performance criteria. Nevertheless, focusing on the third block pair yielded data that could be analysed for 60 participants. Note, the high attrition rate was likely due to the use of a familiarisation IRAP in the place of the usual practice blocks; this issue is addressed, however, in the next study.

(*ps* < .05) for both the *Self-Glove* and *Other-Scarf* trial type in both conditions and for the *Self-Scarf* trial type in the Control Condition.

Table 4
Experiment 2, Fisher's PLSD Comparisons

Trial type	Mean Difference	*P*
Self-Glove, Self-Scarf	0.513	<.001*
Self-Glove, Other-Glove	0.160	.158
Self-Glove, Other-Scarf	0.033	.769
Self-Scarf, Other-Glove	-0.353	.002*
Self-Scarf, Other-Scarf	-0.480	<.001*
Other-Glove, Other-Scarf	-0.127	.263

* Indicates significant *p* values.

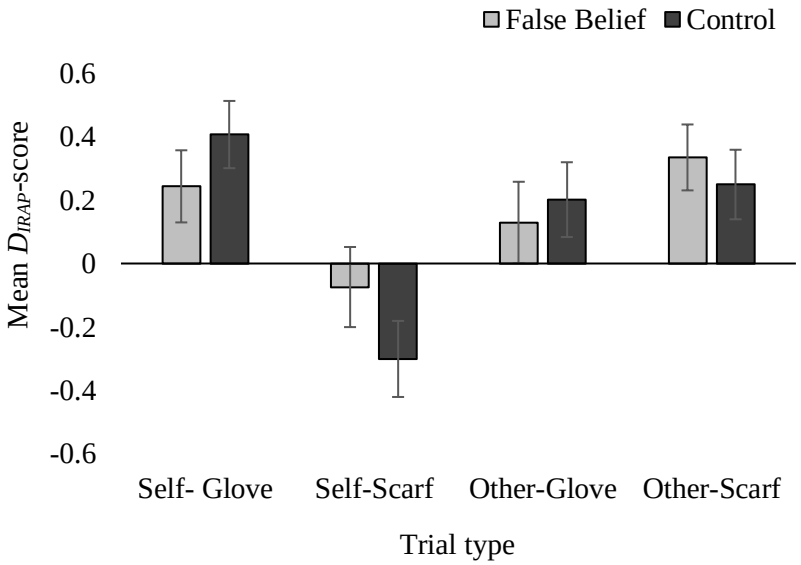

Figure 5. Mean D_{IRAP}-scores for the *third test block pair* of the belief IRAPs in the False Belief and Control Condition. Positive D_{IRAP}-scores indicate history-consistent responding.

Correlations. Given that no main effect emerged in comparing the False Belief and Control Conditions, the D_{IRAP}-scores for both conditions were collapsed before being subjected to correlational analyses with the questionnaires (i.e. a total of 100 correlations; 25 for each trial type). Only five significant correlations (at $p < 0.05$) emerged, with no obvious pattern or clustering around a particular trial type or self-report measure. Specifically, *Self-Glove* correlated positively with the frequency of negative psychotic-like symptoms [$r(58) = .32$, $p < 0.02$] and with distress associated with these symptoms [$r(58) = .27$, $p < 0.04$]. *Self-Scarf* correlated with the casual acquaintances-present sub-scale of the EOSS [$r(58) = .29$, $p < 0.03$]. Finally, *Other-Scarf* correlated with closeness to best friend [$r(58) = .27$, $p < 0.05$] and with greater control of others over the experience of self [$r(58) = .26$, $p < 0.05$]. None of these remained significant after Bonferroni corrections.

Summary and Conclusion. The primary objective of Experiment 2 was to explore the potential impact of a false belief and control vignette on performances in a related belief IRAP. The results indicated vignette-consistent scores for three of the four trial types, with vignette-inconsistent effects recorded on *Self-Scarf* in both conditions. There was little evidence that the two vignettes impacted differentially upon the IRAP performances. The correlational analyses failed to indicate any clear relationships between the self-report measures and the IRAP. Despite there being no significant difference between IRAP performances across the two conditions, the pattern of results suggest that both vignettes, to some degree, impacted the IRAP effects, given that 6 of the 8 trial types were in a vignette-consistent direction. Of course, the vignette-inconsistent effect for the *Self-Scarf* trial type seems somewhat anomalous. On balance, this result could be interpreted as a type of self-positive bias effect, in which participants tended to choose a positively valenced response option (i.e. Yes) on a trial type that presented a self-related label (i.e. a picture of the self). The bias towards responding "Yes" on the *Other-Scarf* trial type would thus be seen as driven largely by the vignette rather than a self-positivity bias effect (see Finn et al., 2018, for empirical evidence for, and a detailed discussion of, the complex manner in which individual trial types may

differentially influence responding on the IRAP). As an aside, it is interesting to note that the anomalous effect for the *Self-Scarf* trial type was substantively larger for the Control Condition, where "Yes" could be interpreted as a vignette-consistent response because the self knows that there are two items (one in each box). On balance, the inferential statistics did not yield a significant difference for condition and thus further speculation seems unwarranted.

In reflecting upon the results obtained in Experiment 2, a number of issues emerged that seemed important to address in a follow-up experiment. First, it became apparent that participants in the Control Condition may have found the relationship between the vignette and the IRAP trial types somewhat ambiguous. Specifically, the control vignette specified that there were two boxes present in the room (one in front of the participant and another in front of the other person), whereas the belief IRAP presented statements that specified only one box. As such, it is difficult to interpret the IRAP effects that were observed for the Control group. A related issue pertains to the fact that the order in which the IRAP blocks were presented was not counterbalanced (i.e. the first block of the belief IRAP was always vignette-consistent). It is possible, therefore, that the IRAP performances for the False Belief group were determined largely by the vignette, whereas the performance of the Control group was simply determined by the pattern that they were required to produce on the first block. Indeed, it could be the case that the IRAP effects for the False Belief group may also have been determined, at least to some degree, by the pattern required by the first block. If this was the case, it could explain why there was limited evidence for a significant difference between the two conditions. With these issues in mind, we designed a subsequent experiment that once again sought to develop an IRAP that would show some sensitivity to a false belief vignette.

Experiment 3

The main aim of Experiment 3 was to determine the extent to which false belief vignettes presented before each block of trials in a belief IRAP would impact the performances observed on that IRAP. In doing so, two specific variables were

manipulated across four conditions; 1. The sequence in which the critical stimuli were specified in the false belief vignette; and 2. the order in which the IRAP blocks were presented (i.e. vignette-consistent followed by vignette-inconsistent versus the opposite block sequence). Specifically, participants were presented with one of two false belief vignettes, both similar to that presented in Experiment 2. Half of the participants were presented a vignette in which a scarf was initially found in the box and this was later replaced with a glove; the other half were presented with a vignette in which a glove was initially found in the box and this was later replaced with a scarf. The only difference between both vignettes, therefore, was the sequence in which the stimuli were specified (i.e. scarf then glove versus glove then scarf). The main rationale for employing these versions of the same vignette was to determine if clear differential patterns of responding, consistent with the specified sequence in the vignette, would be observed in the belief IRAP performances. As noted above, the sequence in which the IRAP blocks were presented in Experiment 2 was not manipulated and thus the extent to which the vignette *per se* determined performance remained unclear. In Experiment 3, therefore, IRAP block sequence was also manipulated, thus creating a mixed 2x2 factorial design: (i) scarf-glove sequence/vignette-consistent first, (ii) glove-scarf sequence/vignette-consistent first, (iii) scarf-glove sequence/vignette-inconsistent first and (iv) glove-scarf sequence/vignette-inconsistent first.

Before continuing, it is important to note that pilot research for Experiment 3 suggested that specific parameters of the belief IRAP could be altered to both reduce attrition rates and increase the impact of the vignettes. First, the familiarisation IRAP employed in Experiment 2 was now replaced by exposure to practice blocks in the belief IRAP. Second, the vignettes were presented before exposure to each block of the IRAP (rather than before each pair of blocks). To avoid any perceived conflict between the vignette and the between-block instructions that were presented in Experiment 2, the latter were removed and replaced with general task instructions presented once at the beginning of the experiment. Given the large number of changes in Experiment 3, relative to 2, the former was deemed to be largely exploratory and thus we did not make any formal predictions.

Method

Design. Experiment 3 comprised of four conditions: (i) scarf-glove sequence/vignette-consistent first, (ii) glove-scarf sequence/vignette-consistent first, (iii) scarf-glove sequence/vignette-inconsistent first, and (iv) glove-scarf sequence/vignette-inconsistent first. Participants were given general task instructions before exposure to the vignette and the IRAP. The Similarity and Strategy Questionnaires that were employed in the previous experiment were presented after completing the IRAP; given the focus of the current experiment, the six questionnaires employed in the previous experiment were not presented here.

Participants. Seventy-four participants were recruited for the current experiment, 58 females and 16 males. Participants ranged from 18-25 years (M = 20.97). The general strategy for recruiting numbers of participants was similar to that previously described in Experiment 1.

Materials and Apparatus

General Task Instructions. Participants were presented with a sheet that provided general instructions for completing the IRAP. The English translation of the instructions presented in Dutch is as follows:

> *You will soon be performing different tasks on the computer. Before each part of the task, you will be presented with a story about you and the person whose picture you brought here today. This story will sometimes be consistent with that computer task you need to perform and other times it will be inconsistent with the computer task you need to perform.*

Condition vignettes. Each of the two vignettes was presented on-screen in a word document. Each comprised of a written paragraph that described one of two scenarios involving the participant and the other person depicted in the other-face picture. Both vignettes were similar to the false belief vignette employed in the previous

experiment. Based on pilot research the vignettes were modified to reduce potential ambiguity in the event described (see below).

Scarf-Glove Sequence Vignette. The scarf-glove sequence vignette specified that there was a scarf in the box first and this was replaced with a glove. The English translation of the vignette presented in Dutch is as follows:

> *The following sentences describe a scenario that involves you and the person in the second picture that you brought with you today. I want you to imagine that you and this person are in a room. In front of you both is a box. When you both open the lid of the box, you both see together that there is a scarf in the box and then you place the lid back on the box again. At this point, the other person leaves the room. Therefore, they cannot see what happens in the room but you still are in the room and you can still see what happens. When they are no longer in the room (but you still are) the scarf is removed from the box and is replaced with a glove. The lid is then put back on the box. At this point, the other person now returns to the room and they are not allowed to take the lid off the box.*

After reading through the vignette, participants were required to indicate which of the following statements best described what happened in the scenario: 1. "I stayed in the room and the other person left the room when the items in the box were changed"; 2. "We both stayed in the room when the items in the box were changed"; or 3. "The other person stayed in the room and I left the room when the items in the box were changed"[9]. Participants were then instructed as follows: "Please remember the details

[9] The researchers checked responses to these questions after the experiment to ensure that they were consistent with the vignette; only one participant in the scarf-glove sequence/vignette-inconsistent-first condition identified the incorrect scenario for one block out of the 16 presented.

of the scenario you read above. as you will require information from this scenario to successfully complete the next part of the experiment."

Glove-Scarf Sequence Vignette. The glove-scarf sequence vignette described a similar scenario to the scarf-glove sequence vignette. The only difference between both vignettes, therefore, was the sequence in which the stimuli were specified (i.e. scarf then glove versus glove then scarf). The same three statements, and the instruction presented above, were used to ensure that participants understood the vignette.

Belief IRAP. The format for the belief IRAP was similar to that presented in the previous experiment, except for the three following modifications; 1. A maximum of five practice blocks pairs were now presented before a fixed number of six test block pairs; 2. The order in which the IRAP blocks (i.e. consistent followed by inconsistent versus inconsistent followed by consistent) was manipulated across the four conditions; 3. Only the vignettes and the three statements, which ascertained participant understanding, were presented before each block (i.e. no additional rules or instructions were used).

The Strategy and Similarity Questionnaires described in the previous experiment were again used to assess performance strategies and perceived physical similarities between the self and other.

Procedure. Experiment 3 took place on an individual basis in sound-proof cubicles at the Department of Experimental, Clinical, and Health Psychology at Ghent University.

General task instructions. The researcher gave participants a copy of the general task instructions. If participants asked for any clarification, the researcher provided this verbally in a brief and concise manner.

Condition vignette and belief IRAP. Participants were exposed to the belief IRAP, with the same vignette presented before each practice and test block throughout the IRAP.

Consistent blocks required responding that was in accordance with the vignette, which was labelled as follows: *Self-Correct/Yes, Self-Incorrect/No, Other-Incorrect/No*

and *Other-Correct/Yes.* Inconsistent blocks required the opposite, labelled as follows: *Self-Correct/No, Self-Incorrect/Yes, Other-Incorrect/Yes* and *Other-Correct/No.*

Questionnaires. Participants completed the two questionnaires immediately after completing the belief IRAP.

Results and Discussion

Questionnaire data. A summary of the means and standard deviations for questionnaires, divided according to the four IRAP conditions, are presented in Appendix M. The similarity scores were relatively high, indicating that participants confirmed that they looked similar to the person depicted in the other picture. In general, the scores divided across the four IRAP conditions did not appear to differ substantively. For the purposes of statistical analysis only the scores from the general similarity question were entered into a one-way between-participants ANOVA and this proved to be non-significant ($p = .23$).

The means and standard deviations for the strategy scores indicate that participants perceived that they were relatively successful at taking the other person's perspective and that they felt that the vignette controlled their responding on the IRAP. The scores divided across the four IRAP conditions did not appear to differ substantively and two one-way between-participant ANOVAs, one for each question, both proved to be non-significant; given the separate analyses for each question a Bonferroni correction was applied ($p < .025$).

In addition, the open-ended strategy question was assessed ($N = 51$). The open-ended answers were read by one researcher who developed an initial coding frame to organise the data. These codes were then grouped into categories, according to how they were related. Following this, a second researcher independently reviewed the coding and categories developed by the first researcher. Any inconsistencies or issues raised by the second researcher were discussed and the categories were adjusted accordingly. A list of the final 11 categories is presented in Appendix N. The most common strategies recorded by participants was linking the object words and picture words together ($N =$

16) and focusing on the scenario to help complete the IRAP (N = 16). Forteen participants reported that they relied on the IRAP feedback and 11 reported that they focused only on the object word and did not read the full sentences. Nine participants reported that they rehearsed the link between the object and person before starting an IRAP block.

IRAP data. Due to a technical issue, the IRAP data for two participants were lost and thus were not included in the analyses. The data for a third participant were also excluded after the participant reported prior familiarity with similar IRAPs, which recent research indicates may influence IRAP performance (see Finn et al., 2018).

As noted previously, practice blocks required an accuracy level of ≥ 80% and a median latency of ≤ 2,000ms; three participants failed to achieve these criteria across five exposures and thus they did not proceed to the test blocks. Test blocks required an accuracy level of ≥ 79% and a median latency of ≤ 2,000ms (on two of the three successive pairs), which four participants failed to achieve, and thus the data for these participants were not included in subsequent analyses. Seven participants failed to maintain the accuracy and latency criteria for one of the pairs of test blocks, and thus their scores were calculated from the remaining two pairs (see Nicholson & Barnes-Holmes, 2012). The final analyses contained N = 64 (scarf-glove sequence/vignette-consistent-first, N = 17; scarf-glove sequence/vignette-inconsistent-first, N = 15; glove-scarf sequence/vignette-consistent-first, N = 16; and glove-scarf sequence/vignette-inconsistent-first, N = 16).

D_{IRAP}-scores. As noted previously, participants were divided into four conditions based on the vignette sequence (i.e. scarf-glove versus glove-scarf) and the IRAP block sequence (i.e. vignette-consistent-first versus vignette-inconsistent-first). D_{IRAP}-scores were calculated for each trial type, such that positive scores indicated a response bias that was consistent with the scarf-glove vignette and negative scores indicated a response bias that was consistent with the glove-scarf sequence. The data were entered into a preliminary 2x2x4 mixed repeated-measures ANOVA and this yielded a significant three-way interaction, $F(1, 60) = 18.85$, $p < .0001$, $\eta_p^2 = .24$. A

main effect for vignette, $F(1, 60) = 4.41$, $p = .04$, $\eta_\rho^2 = .06$, and a two-way interaction for vignette and IRAP block sequence, $F(1, 60) = 23.54$, $p < .0001$, $\eta_\rho^2 = .28$ were also recorded. Given the highly significant three-way interaction with IRAP trial type it was decided at this point to conduct four separate 2x2 independent ANOVAs, one for each trial type; a Bonferroni correction ($p < .0125$) was applied to the multiple follow-up ANOVAs.

A graphical representation of the four ANOVAs is presented in Figure 6. The following explanation may assist in the interpretation of the figure. If the vignette controlled the IRAP performances, then the two bars for the scarf-glove sequence should be in a positive direction, whereas the two bars for the glove-scarf sequence should be in a negative direction. The top left panel shows that the D_{IRAP}*-scores* for the *Self-Correct* trial type were vignette-consistent for both vignette sequences (scarf-glove and glove-scarf), but only when the IRAP commenced with a vignette-consistent block. The opposite appeared to be the case when the IRAP commenced with a vignette-inconsistent block. The descriptive analysis was supported by the 2x2 ANOVA, which yielded a significant interaction, $F(1, 60) = 83.83$, $p < .0001$, $\eta_\rho^2 = .58$, but no significant main effects (*ps* > .4). The top right panel shows that the D_{IRAP}-scores for the *Self-Incorrect* trial type were vignette-consistent, and only marginally so, for the scarf-glove sequence when the IRAP commenced with a vignette-consistent block. The 2x2 ANOVA yielded no significant main or interaction effects (all ps > .04). The bottom left panel shows that the D_{IRAP}*-scores* for the *Other-Incorrect* trial type were relatively small and all vignette-inconsistent. The 2x2 ANOVA yielded no significant main or interaction effects (all *ps* > .09). The bottom right panel shows that the pattern of D_{IRAP}-scores for the *Other-Correct* trial type were similar (albeit weaker) to the pattern observed for the *Self-Correct* trial type. The 2x2 ANOVA, yielded a significant interaction, $F(1, 60) = 15.23$, $p = .0002$, $\eta_\rho^2 = .20$, but no significant main effects (*ps* > .4). Overall, therefore, there was little evidence that the vignette controlled the IRAP performances for any of the four trial types. Indeed, for the *Self-Correct* and *Other-Correct* trial types the pattern of IRAP effects suggests that the primary controlling

variable was the order in which the IRAP blocks were presented (vignette-consistent-first versus vignette-inconsistent-first).

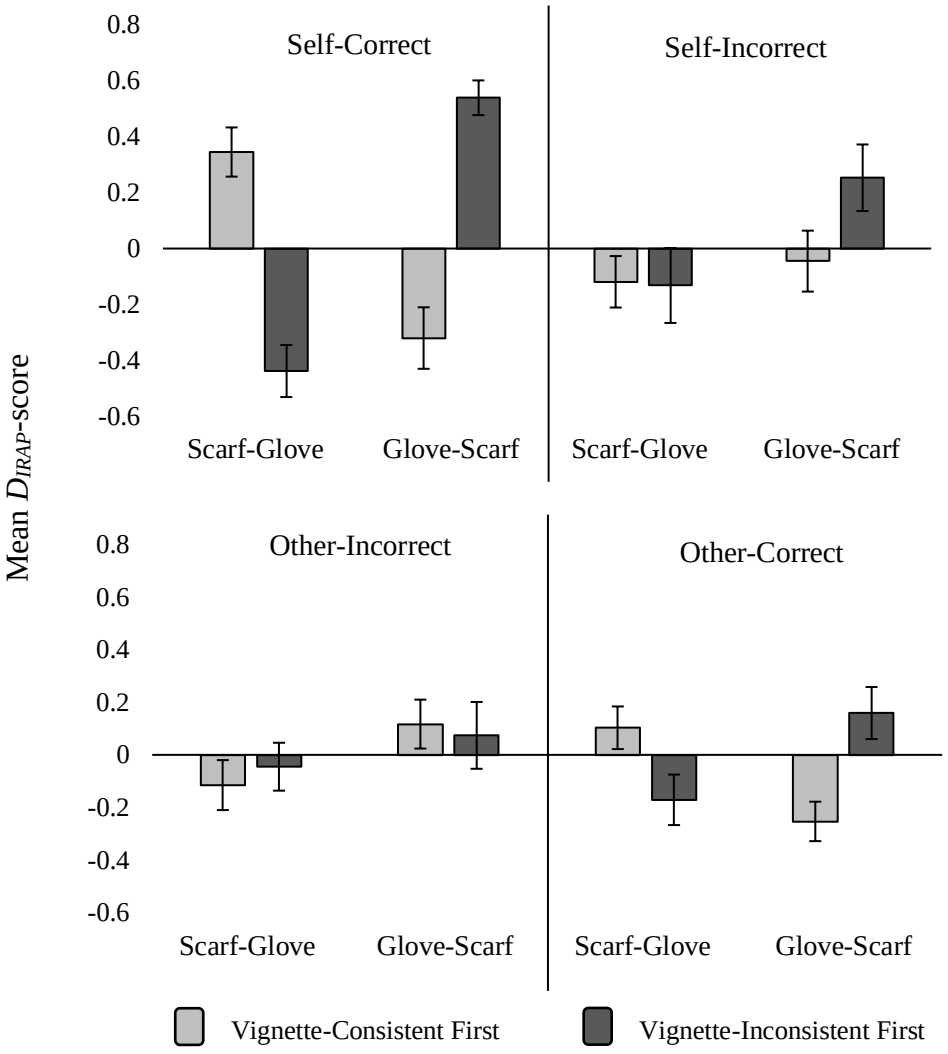

Figure 6. Mean D_{IRAP}-scores for the four trial types of the belief IRAP across the four conditions (scarf-glove sequence/vignette-consistentfirst*;* glove-scarf sequence/vignette-consistent-first*;* scarf-glove sequence/vignette-inconsistent-first*;* glove-scarf sequence/vignette-inconsistent-first). Positive D_{IRAP}-scores indicate a response bias that is consistent with the scarf-glove sequence vignette; negative scores indicate a response bias that is consistent with the glove-scarf sequence vignette.

At this point in the analyses of the data from the IRAP it appeared that the false belief vignette had virtually no impact on the response biases recorded during the test

blocks. Instead, the response pattern required during exposure to the first block of trials seemed to drive the IRAP response biases. In drawing this conclusion, however, it may be premature to assume that the vignette had *no impact whatsoever* on the IRAP performances. For example, perhaps the vignette was a controlling variable, but only when it cohered with the initial exposure to the IRAP. Or to put it another way, if participants perceived the vignette to be an accurate guide on how to respond on the IRAP they simply continued to be guided by both sources. If, however, participants perceived the vignette to be an inaccurate guide, then they simply ignored the vignette and treated the first block of IRAP trials as the 'correct' pattern. If this post-hoc interpretation of the findings is correct, then perhaps performances on the IRAP may differ during the practice blocks (i.e. when participants first encounter either coherence or incoherence between the vignette and the feedback contingencies of the IRAP). To test this suggestion, we simply compared the number of practice blocks that participants required in the vignette-consistent-first versus the vignette-inconsistent-first conditions. The difference proved to be significant with the consistent-first group requiring a mean of 2.03 (SD = 1.12) practice blocks to reach the accuracy and latency criteria versus a mean of 2.93 (0.96) for the inconsistent-first group.

To further explore the potential impact of coherence between the vignette and initial exposure to the IRAP, we analysed the individual D_{IRAP}-scores from the first pair of practice blocks. Although this analytic strategy is rarely if ever adopted in IRAP research (because the IRAP performances could not be considered relatively stable in terms of the desired stimulus control) it seemed reasonable to adopt it here to address the post-hoc question we were asking. We restricted our analysis to the first pair of practice blocks because a large number of the participants (N = 18), particularly in the consistent-first condition (N = 16), only required one pair of practice blocks before proceeding to the test blocks.

The data from the first pair of practice blocks were entered into a preliminary 2x2x4 mixed repeated measures ANOVA, and this yielded a two-way interaction for vignette and IRAP block sequence, $F(1, 60) = 62.71$, $p < .0001$, $\eta_\rho^2 = .51$ and a main

effect for vignette, $F(1, 60) = 52.36$, $p < .0001$, $\eta_p^2 = .62$. A graphical representation of the interaction for vignette and IRAP block sequence is presented in Figure 7. The graph shows that the D_{IRAP}-scores for the vignette-consistent-first conditions was marginally vignette-inconsistent for both vignette sequences (scarf-glove sequence and glove-scarf sequence). Similar, but far stronger effects, were observed for the vignette-inconsistent-first conditions. Given the highly significant two-way interaction, it was decided to conduct four follow-up unpaired *t*-tests; a Bonferroni correction ($p < .0125$) was applied. Only one of the four *t*-tests proved to be non-significant; the comparison between the two vignette-consistent-first conditions ($p > .23$; remaining $ps < .0001$). The pattern of effects for the first pair of practice-blocks suggests that when the vignette and the initial IRAP contingencies cohered, block sequence had a limited impact on IRAP performance. However, when the vignette and the IRAP contingencies did not cohere (during initial contact with the IRAP), block sequence was a dominant controlling variable.

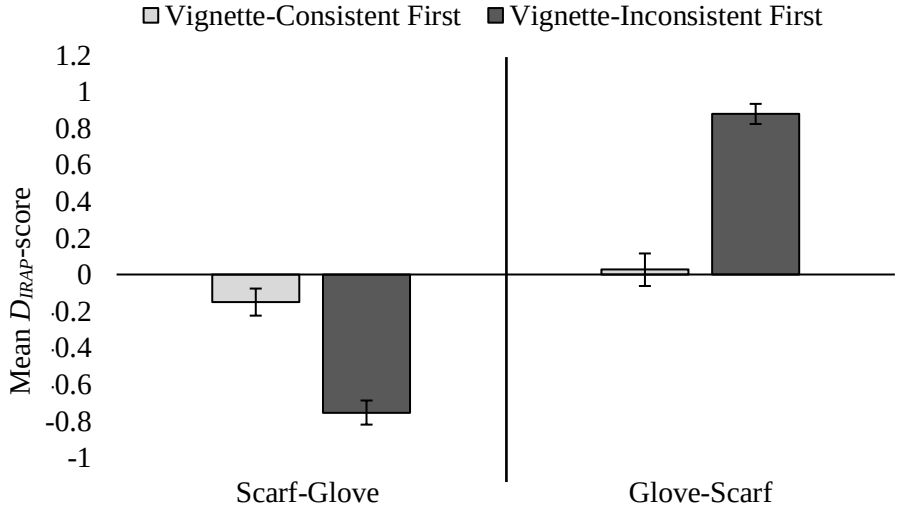

Figure 7. Mean overall *D*-scores for the first pair of practice blocks for the four conditions (scarf-glove sequence/vignette-consistent-first*;* glove-scarf sequence/vignette-consistent-first*;* scarf-glove sequence/vignette-inconsistent-first*;* glove-scarf sequence/vignette-inconsistent-first). Positive *D*-scores indicate a response bias that is consistent with the scarf-glove sequence vignette; negative scores indicate a response bias that is consistent with the glove-scarf sequence vignette.

Overall, the results of this third experiment appear to confirm that the false belief vignettes had a limited impact on the IRAP performances during the test blocks. Indeed, the primary controlling variable, at least with respect to the *Self-Correct* and *Other-Correct* trial types, was the order in which the two types of IRAP blocks were presented. In other words, the observed IRAP effects were in a direction that was consistent with the contingencies that were contacted during the first block of trials on the IRAP, rather than the content of the vignettes. On balance, the vignettes did not appear to be completely inert, as controlling variables, because the participants in the vignette-consistent-first conditions required fewer practice blocks to reach criteria than participants in the vignette-inconsistent-first conditions. Furthermore, there were large and significant differences in the actual overall IRAP effects during the initial pair of practice blocks across the two sequences (i.e. vignette-consistent-first versus vignette-inconsistent-first).

General Discussion

The primary objective of Experiment 1 was to determine if it was feasible to use matched pictures of self versus other across two IRAPs. In both IRAPs, the contrast categories were pen-related stimuli. In general, the pattern of effects indicated that pictures of faces could be incorporated into the IRAP, thus supporting the feasibility of using photographs of self versus other in IRAP perspective-taking studies. However, two effects did emerge that are worthy of further consideration, but these will be addressed in the final chapter of the thesis because the effects were also observed in studies presented in other chapters.

The primary objective of Experiment 2 was to determine if the presentation of a false belief vignette before exposure to a single IRAP would influence performance in a vignette-consistent direction. The results of this study were inconclusive for a number of reasons. First, no clear differences emerged between the active vignette and control condition (and the latter could have been interpreted as being ambiguous for the participants). Second, the use of a familiarisation IRAP, instead of requiring participants

to complete practice blocks in the belief IRAP, likely contributed to high rates of attrition thus restricting data analyses to the final pair of test blocks. Third, the order in which IRAP blocks were presented (i.e. vignette-consistent-first versus vignette-inconsistent-first) was not counterbalanced and thus it was difficult to determine if the vignette and/or the block sequence was the primary controlling variable in determining the observed IRAP effects.

The primary aim of Experiment 3 was to rectify the three main problems described above. The experiment thus involved an active control design in which the two vignettes specified the opposite states of affairs; the familiarisation IRAP was removed from the experiment, but the belief IRAP included practice blocks; and the block sequence of the IRAP was counterbalanced. The results of the third experiment were more conclusive than those of the second, in indicating that the primary controlling variable was the sequence in which the IRAP blocks were presented rather than the content of the vignettes. On balance, post-hoc analyses of performance of the first practice block pairs indicated that the block sequence and vignettes did interact with each other, but only when the vignette and the initial contingencies of the IRAP cohered with each other.

At this stage in the doctoral research programme 11 individual experiments had been completed and there was limited evidence that an IRAP had been developed that showed clear and unambiguous sensitivity to perspective-taking. On balance, the research reported in the current thesis also contributed to the on-going development of a conceptual analysis of perspective-taking and the dynamics of arbitrarily applicable relational responding, more generally. And thus, the current findings are highly valuable in that regard. A discussion of these conceptual developments, and their relationship to the current empirical work, will be presented in the final chapter of the current thesis, to which we now turn.

Chapter 5

General Discussion

The primary aim of the current thesis was to determine the potential utility of the IRAP in the domain of persepctive-taking. The format of the current chapter is as follows. The chapter begins with a summary of the main findings per thesis chapter. This is followed by a discussion of the link between the current empirical work and the conceptual analyses of perspective-taking. Specifically, the relationship between IRAP results in Chapters 2 and 4 and the the Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) model of the IRAP is explored. In addition, we grapple with a DAARRE model interpretation of the findings from Chapter 3. Finally, we suggest an alternative strategy for conceptualising perspective-taking based on a Multi-Dimensional Multi-Level (MDML) framework for RFT, and in doing so, we present a conceptual functional analysis of a classic false belief task based on the MDML framework.

Summary of the Main Thesis Findings

Chapter 2. The aim of the two experiments in Chapter 2 was to replicate previous IRAP research that successfully produced differential responding to self versus other. The initial purpose of Experiment 1 was to conduct a systematic replication and extension of the first IRAP study of perspective-taking reported by Barbero-Rubio et al. (2016) in terms of exploring potential differences in responding to self and other. Similar to the original findings, the results of our first study yielded significant effects for the two trial types *I-I* and *Other-Other* (in both the deictic and control IRAPs), although *all four* trial types yielded significant effects in the Barbero-Rubio et al. study. In addition, these two trial types differed significantly from each other in our study, which had not been the case in the original. The purpose of the second experiment was to investigate whether the introduction of a *read-aloud procedure* would produce a similar pattern of effects and whether these would more closely resemble Experiment 1 or the original study (i.e. four signficant effects versus two). The data again showed

relatively strong IRAP effects on the same two trial types, but we no longer found a significant difference between them.

In attempting to explain the different pattern of results we obtained in Experiment 1 from the original study, in terms of our significant effects in only two trial types and the fact that these differed significantly from each other, it is important to note first that some of the procedures involved in running the IRAPs differed substantively between these two studies. For example, in the study by Barbero-Rubio et al. (2016), participants received explicit perspective-taking instructions at the beginning of each IRAP block (i.e. "For the next block of trials, you have to respond as if you were you and Adrian were Adrian/as if you were Adrian and Adrian were you"). In addition, participants completed a DRT, with 12 scenarios involving reversed deictic relations (e.g. “Mario is swimming in the pool, and Ramon is sailing in a boat. If Ramon were Mario, what would he be doing?”) and eight involving double reversed deictic relations (e.g. “Luis is in Teide analysing sediments, and Maria is in Kilimanjaro searching for the source of a river. If Luis were Maria and if the Kilimanjaro were the Teide, where would Luis be?”). In stark contrast, participants in Experiment 1 of Chapter 2 were exposed to a basic familiarisation IRAP that focused on fruit and vegetables, *with no reference to perspective-taking*. And, even when participants were exposed to the deictic and control IRAPs, they received no specific instructions regarding perspective-taking, either at the beginning of the IRAPs or before each block.

As a result of these considerable procedural differeces, it is difficult to draw firm conclusions about the variables that may have been responsible for the different findings between the original study and our Experiment 1. However, there is some evidence elsewhere that the type of instructions presented before and during IRAP tasks may have quite dramatic effects on performance, in and of themselves (see Finn et al., 2016). In other words, the explicit instructions regarding, and references to, perspective-taking, provided during the original study by Barbero-Rubio et al. (2016) may have influenced their outcomes, at least to some extent, in a manner that was not operating in our Experiment 1.

The second pattern of differences from Chapter 2 that warrants exploration is the different findings recorded between our own Experiments 1 and 2, particularly in terms of the instability of the significant difference between the *I-I* and *Other-Other* trial types (this difference also differentiated our Experiment 1 from the original work). We will return to this issue later because it requires to a broader discussion of on-going conceptual developments in RFT and its approach to perspective-taking (see the section below entitled The Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) model of the IRAP and its relationship to the current empirical research.

Chapter 3. The broad aim of the empirical work in Chapter 3 was to develop an IRAP that required participants to respond to relatively complex statements regarding their own mental states and the mental states of others. Specifically, across six experiments, we employed the NL-IRAP to investigate perspective-taking with regard to the mental states of self and other. That is, the self-focused IRAP required participants to respond to positive and negative statements about themselves, while the other-focused IRAP required responding to similar statements about an other. In Experiments 1 and 2, we compared self with an *unspecified* other and explored the extent to which the two (self and other) NL-IRAPs potentially correlated with each other and with a range of relevant self-report measures. In Experiments 3-6, we compared self with a *specified* other and we manipulated the nature of this specified relationship across experiments.

In short, the data from Experiments 1 and 2 showed significant differences between the self- and other-focused IRAPs, when other was unspecified. In contrast, the data from Experiments 3-6 showed no significant differences between the self- and other-focused IRAPs, when other was specified in various ways. In order to facilitate comparisons of the findings from the six studies in Chapter 3, Table 1 provides an overview of the results from each experiment under a series of common headings. Experiment 1 compared self with an unspecified other and found a significant difference between them, with no correlations of any kind. Experiment 2 also compared self with an unspecified other, again found a significant difference between them, and only a

correlation between the IRAPs. Experiment 3 compared self with a *specified* other (person closest to) and included a focusing task. No significant difference between the IRAPs was found, nor were there any correlations. Experiment 4 compared self with a specified other (person closest to) but with *no* focusing task, and again found no significant difference between them, and no correlations. Experiment 5 compared self with a disliked specified other, found no significant difference between them, but now found a correlation between the IRAPs. Finally, Experiment 6 compared self with a specified other (different from), found no significant difference between them, and found a correlation between the IRAPs.

Table 1
Chapter 2, summary of the main findings from Experiments 1-6

Experiment	Difference in overall *D-IRAP* score between Self and Other IRAPs	IRAPs Correlated	Correlations between IRAPs and self-reports (uncorrected)	Corrected Correlations
Experiment 1 Self and Unspecified Other	Yes	No	IRAP difference score with CAPE (positive symptoms/frequency) [$r(37) = 0.34$, $p = 0.039$]	None
Experiment 2 Self and Unspecified Other	Yes	$r\ (49) = 0.315$, $p = 0.027$	Other-focused IRAP with CAPE (positive symptoms/frequency) [$r(49) = -0.3$, $p < 0.04$] Other-focused IRAP with ECR-RS attachment-related avoidance subscale for romantic partner [$r(49) = 0.34$, $p < 0.02$] and for mother [$r(49) = 0.31$, $p < 0.03$]	None
Experiment 3 Self and Specified Other (closest to), with focusing task	No	$r\ (33) = 0.414$, $p = 0.016$	Self-focused IRAP with CAPE (negative symptoms/frequency) [$r(33) = 0.38$, $p < 0.03$] and with Self-warmth Thermometer [$r(33) = 0\ .49$, $p < 0.01$]	None

			Specified other-focused IRAP with ECR-RS attachment-related avoidance subscale for one's mother [$r(33) = -0.43$, $p < 0.01$] IRAP difference score with ECR-RS attachment-related avoidance for one's mother [$r(33) = 0.41$, $p < 0.02$]	
Experiment 4 Self and Specified Other (closest to), without focusing task	No	No	Self-focused IRAP with EOSS close relationships-absent subscale [$r(30) = 0.47$, $p < 0.01$] Specified other-focused IRAP with ECR-RS attachment-related avoidance subscale for one's mother [$r(30) = -0.42$, $p < 0.05$] one's father [$r(30) = -0.391$, $p < 0.04$] with ECR-RS attachment-related anxiety subscale for best friend [$r(30) = -0.38$, $p < 0.04$] and with the overall PFI score [$r(30) = -0.42$, $p < 0.03$]	None

			IRAP difference score with ECR-RS attachment-related avoidance subscale for romantic partner [$r(30) = -0.38$, $p < 0.04$]	
Experiment 5 Self and Specified Other (dislike)	No	$r(30) = 0.410$, $p = 0.023$	Self-focused IRAP with ECR-RS attachment-related avoidance subscale for one's mother [$r(30) = -0.38$, $p < 0.04$]	None
Experiment 6 Self and Specified Other (different from)	No	No	Self-focused IRAP with ECR-RS attachment-related avoidance subscale for romantic partner [$r(30) = -0.41$, $p < 0.03$] and with the EOSS close relationships-present subscale [$r(30) = -0.43$, $p < 0.02$] Specified-other IRAP with the EOSS close relationships-present subscale [$r(30) = -0.39$, $p = 0.03$] IOS for best friend [$r(30) = -0.41$, $p = 0.03$] and with the Self Warmth Thermometer	None

[$r(30) = -0.46$, $p < 0.01$]

Overall, each experiment in Chapter 3 produced performances that would be deemed consistent with the pre-experimental histories of the participants, in that the IRAP effects were generally in the predicted direction. Furthermore, the rather robust outcomes with regard to specified versus unspecified other seem potentiually important (see also DeBernardis et al., 2014). Indeed, we failed to find strong evidence of perspective-taking when other was specified (Experiments 3-6), at least in terms of different performances across the two IRAPs, or in correlations among the IRAPs and the self-report measures. We discuss below our attempts to grapple with the observed difference between self and specified/unspecified other (see section on Grappling with the Findings from Chapter 3).

Taken together, these findings suggest some level of sensitivity in the NL-IRAPs to self and other, although the use of complex statements in this type of IRAP may have also served to undermine or reduce the impact of deictic relational responding per se. That is, in presenting such complex stimuli or relational networks in the IRAP, participants may have more or less interpreted the procedure as a sense-making or problem-solving task, in which the self versus other had little or no impact on responding, particularly when the other was specified.

The correlational analyses between the IRAPs and the self-report measures yielded very few significant effects (and none at all after Bonferroni corrections). Furthermore, the correlational analyses between the two IRAPs in each experiment were significant in some cases (Experiments 2, 3 and 5), but not others (Experiments 1, 4 and 6). Taken together, little more can be said with any conviction regarding these very limited statistical outcomes.

Chapter 4. Because of the possibility, as suggested above, that participants interpreted the NL-IRAPs as a sense-making rather than self versus other task (particularly when other was specified), we returned to using the standard IRAP in the three experiments that comprised Chapter 4, but we now tried to capture self and other using actual photographs (Experiment 1), as well as vignettes that required perspective-taking before exposure to the IRAP (Experiments 2 and 3).

The primary objective of Experiment 1 in Chapter 4 was to determine if it was feasible to use matched pictures of self versus other across two IRAPs. The results indicated that with the exception of one trial type, the effects recorded were consistent with what might be expected based on previously-established verbal relations. That is, participants tended to confirm more readily than to deny that a picture of a face was a face and that a picture of a pen was a pen. They also tended to deny more readily than

confirm that a picture of a pen was a face and that a pen-related word was a picture of a face. This pattern of effects indicated that indeed pictures of self and other can be successfully employed in an IRAP. However, the question remained regarding whether they facilitated perspective-taking per se.

At this point, we faced the challenge of deciding how to assess perspective-taking using IRAPs that presented pictures of self and other without adding complex statements that may again undermine perspective-taking (as in the NL-IRAPs). The strategy we adopted thus involved avoiding the use of complex statements within the IRAP, but instead presenting vignettes that required perspective-taking just prior to the completion of an IRAP. Specifically, the vignettes were designed to assess the attribution of false belief. The ratiobale for assessing false belief was that it is a widely used format for exploring perspective-taking in the ToM literature and its inclusion in the vignettes allowed us to more systematically depict perspective-taking than would more general self and other statemenst (as in the NL-IRAP).

The specific objective of Experiment 2 was to determine if the presentation of a false belief vignette before exposure to a single IRAP would influence performance in a vignette-consistent direction, thus showing that the IRAP could capture complex perspective-taking. Specifically, the false belief vignette comprised of a written paragraph that described a scenario involving the participant and the other person depicted in the other-face picture. In this scenario, the participant observed that items in a box switched locations when the other person left the room. A belief IRAP was then presented that required the participant to respond to what they thought was in the box and what they thought that the other person thought was in the box. Given that the items had been switched when the other person had the left the room, the self and other should differ in terms of what they believed to be in the box. The key question was, would the observed IRAP effects be consistent with the false belief vignette? A control vignette was presented to half of the participants, in which there was no exchange of the items in the box and therefore no false belief attribution was required. However, the results

were inconclusive and there is a number of possible reasons for this outcome. (1) The vignette presented in the Control Condition could have been interpreted as ambiguous for participants. (2) High rates of attrition restricted data analyses to the final pair of test blocks. (3) The order in which the IRAP blocks were presented (i.e. vignette-consistent-first versus vignette-inconsistent-first) was not counterbalanced. Naturally, the primary aim of Experiment 3 was to rectify these three issues by: (i) including an active control design in which the two vignettes specified the opposite states of affairs; (ii) included practice blocks in the IRAP; and (iii) counterbalancing the block sequence of the IRAP. The results Experiment 3 were more conclusive, but still suggested that the primary controlling variable was the sequence in which IRAP blocks were presented, rather than the actual content of the vignettes. Indeed, the primary controlling variable, at least with respect to the *Self-Correct* and *Other-Correct* trial types, was the order in which the two types of IRAP blocks were presented. Thus, the extent to which particiapntes were respoding to self versus other appears to be limited. Nevertheless, post-hoc analyses did indicate that the vignettes did impact on performance, but only when the vignette and the initial contingencies of the IRAP cohered with each other.

The Current Empirical Research and its Relationship to Conceptual Analyses of Perspective-Taking

As noted previously, the empirical research reported in the current thesis was intertwined throughout with the on-going development of conceptual analyses within RFT itself, and thus it is important to place the empirical research here in that context. In order to achieve this objective, it will be useful to separate these conceptual analyses into two parts. The first part will focus on what has been labelled the Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) model and the second part will focus on a recent attempt to provide a systematic framework for research within RFT.

The Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) model of the IRAP and its relationship to the current empirical

research. Recently, some RFT-based research has focused on variables that appear to impact IRAP performances. In a series of studies, a persistent pattern of IRAP effects referred to as the Single-Trial-Type-Dominance-Effect (STTDE) has been recorded (see Finn, Barnes-Holnes, & McEnteggart, 2018). This pattern involves significant differences in magnitude between trial types that share the response option "True" during history-consistent blocks. These results were not previously accounted for by any existing model of IRAP performances. Thus, the authors proposed a new model, known as the Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) model. This model incorporates variables beyond the particular stimulus-stimulus relations (between label and target) presented within each trial type of the IRAP. In the paragraphs below, we use the DAARRE model to potentially explain the STTDE effect observed in the data from Experiment 1 and the Double-Trial-Type-Dominance-Effect (DTTDE) observed in the data from Experiment 2 of Chapter 2. We also use the model to explain the unexpected effect observed in Experiment 1 of Chapter 4.

Chapter 2. The findings from Experiment 1 in Chapter 2 (see Figure 2) showed a clear STTDE for the *I-I* trial type, whereas the data from Experiment 2 showed a DTTDE for the *I-I* and *Other-Other* trial types. Although the empirical work conducted in the current thesis was *not* designed to test the DAARRE model or to manipulate either of these effects, placing the current data in the context of the model and the effects observed in parallel research could be particularly instructive, especially in terms of future research.

In attempting to explain the STTDE for the *I-I* trial type in Experiment 1, we assume that self-related terms possess relatively strong orienting or recognition responses relative to other-related words (Alexopoulos, Muller, Ric, & Marendaz, 2012). We make this assumption based on the fact that, in general, most individuals engage more frequently in self-related verbal behaviour than in verbal buhaviour related to others. It is worth noting that the STTDDE was *not* observed in the Control Condition in Experiment 1, which supports the assumption that the self-related stimuli possess some functions that other-related stimuli do not. The DAARRE model as it applies to

the self and other stimuli employed in Experiments 1 and 2 of Chapter 2 is presented in Figure 1. The reader is encouraged to consult Figure 1 while reading the following text.

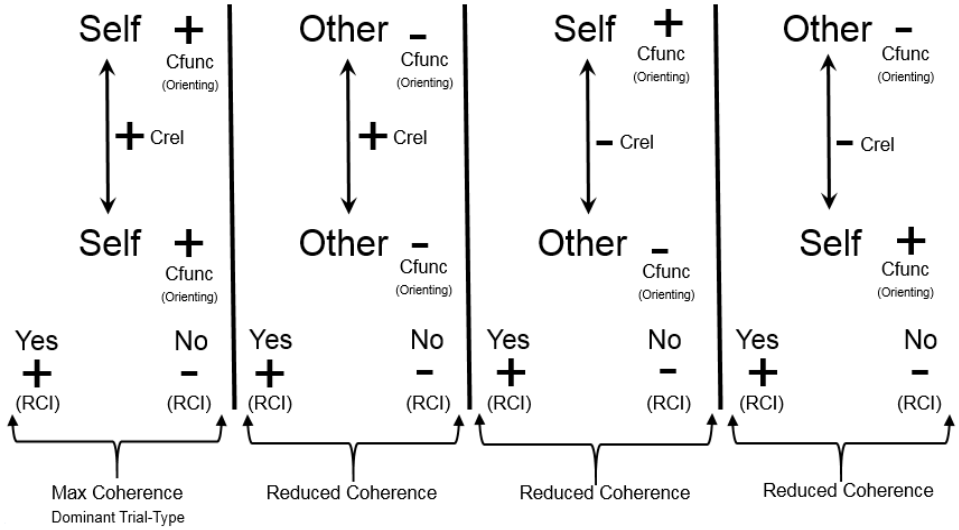

Figure 1. The DAARRE model as it applies to the deictic IRAP in Chapter 2. The positive and negative labels refer to the relative positivity of the Cfuncs, for each label and target, the relative positivity of the Crels, and the relative positivity of the RCIs in the context of the other Cfuncs, Crels, and RCIs in that stimulus set.

The model identifies three key sources of behavioural influence: (1) the relationship between the label and target stimuli (labeled as *Crels),* (2) the orienting functions of the label and target stimuli (labelled as *Cfuncs*), and (3) the coherence functions of the two response options (e.g. "Yes" and "No"). Consistent with the earlier suggestion that self-related terms likely possess stronger orienting functions relative to other-related terms, the Cfunc property for self is labelled as positive and the Cfunc property for Other is labelled as negative. The negative labelling for Other should not be taken to indicate a negative orienting function, but simply an orienting function that is weaker than that of self. The labelling of the relations between the label and target stimuli indicates the extent to which they cohere or do not cohere based on the participants' relevant verbal history. Thus, an *I-I* relation is labelled with a plus sign (i.e. coherence), whereas an I-Other relation is labelled with a minus sign (i.e. incoherence). Finally, the two response options are each labelled with a plus or minus

sign to indicate their functions as either coherence or incoherence indicators (see Maloney & Barnes-Holmes, 2016). In the current example, “Yes” (+) would typically be used in natural language to indicate coherence and “No” (-) would be used to indicate incoherence. Note, however, that these and all of the other functions labelled in Figure 6 are behaviourally determined, by the past and current verbal history of the participant, and should not be seen as absolute or inherent in the stimuli themselves.

As can be seen from Figure 1, each trial type differs in its pattern of Cfuncs and Crels, in terms of plus and minus properties, that define the trial type for the deictic IRAP. The STTDE for the *I-I* trial type in Experiment 1 may be explained, as noted above, by the DAARRE model based on the extent to which the Cfunc and Crel properties cohere with the relational coherence indicator (RCI) properties of the response options across blocks of trials. To appreciate this explanation, note that the Cfunc and Crel properties for the *I-I* trial type are all labelled with plus signs; in addition, the RCI that is deemed correct for history-consistent trials is also labelled with a plus sign (the only instance of four plus signs in the figure). In this case, therefore, according to the model this trial type may be considered maximally coherent during history-consistent trials. In contrast, during history-inconsistent trials, there is no coherence between the required RCI (minus sign) and the properties of the Cfuncs and Crel (all plus signs). According to the DAARRE model, this stark contrast in levels of coherence across blocks of trials serves to produce a relatively large IRAP effect. Now consider the *Other-Other* trial type, which requires that participants choose the same RCI as the *I-I* trial type during history-consistent trials, but here the property of the RCI (plus sign) does *not* cohere with the Cfunc properties of the label and target stimuli (both minus signs). During history-inconsistent trials, the RCI (minus sign) *does* cohere with the Cfunc properties but not with the Crel property (plus sign). Thus, the differences in coherence between history-consistent and history-inconsistent trials across these two trial types is not equal (i.e. the difference is greater for the *I-I* trial type) and thus favours the STTDE for *I-I*. Finally, as becomes apparent from inspecting the figure for the remaining two trial types (*I-Other* and *Other-I*), the differences in coherence across

history-consistent and history-inconsistent blocks is reduced relative to the *I-I* trial type (two plus signs relative to four), thus again supporting the STTDE.

At this point, the DAARRE model appears to explain the STTDDE for the deictic IRAP in Experiment 1. But how might it explain the apparent moderating influence of the read-aloud procedure which appeared to undermine the STTDE, as observed in Experiment 2? Although entirely speculative, it is possible that requiring participants to read aloud each label and target stimulus as they appeared on-screen reduced or eradicated the influence of the orienting functions of those stimuli. In other words, because every label and target was given a similar function by the read-aloud requirement, this overshadowed the differential orienting functions that were present when reading aloud was not required. Thus, the remaining controlling variable was the Crel property, which was the same across the *I-I* and *Other-Other* trial types (both plus signs). As an aside, perhaps the perspective-taking instructions and DRT training provided in the Barbero-Rubio et al. (2016) study had a functionally similar impact to the read-aloud requirement (i.e. it attenuated the Cfunc properties of the IRAP and thus a DTTDE was observed).

In presenting the foregoing model, we recognise that it is specific to the IRAP, but if the IRAP is to be developed as a method for analysing deictic relational responding, and perspective-taking more generally, it is essential that we understand as fully as possible the functional processes involved in the behavioural patterns we observe with this methodology. Imagine, for example, that a deictic IRAP was used to explore potential differences in deictic relational responding between groups with different clinical profiles. It may be important to determine if those groups differ in terms of the orienting functions for self and other, rather than any difference in their ability to relate self-to-self and other-to-other.

In considering the potential relevance of the DAARRE model to understanding IRAP performances in the clinical domain, it seems important to reflect upon the pattern of correlations we obtained between the IRAP performances and the self-report measures of psychological presented in Experiments 1 and 2 of Chapter 2. Specifically,

in Experiment 1, only the *Other-Other* trial type correlated with the CAPE. That is, increased response bias in responding to others as others predicted lower levels of psychotic-like experiences. In Experiment 2, however, only the *I-I* trial type correlated with the CAPE. That is, increases in response bias in responding to self as self predicted *higher* levels of psychotic-like experiences. Although these correlations did not remain significant following Bonferroni correction, the fact that the pattern of correlations differed in the presence versus the absence of the read-aloud procedure may indicate that manipulating the dominance of the orienting versus relational functions of the IRAP impacts upon specific features of its predictive validity. In Experiment 2, for example, the relatively strong pattern of self-self relational responding, in the absence of orienting functions (i.e. in the presence of the read-aloud procedure), predicted higher levels of psychological suffering. When relative differences in orienting functions were present in the IRAP (i.e. in the absence of the read-aloud procedure), the *Other-Other* trial type predicted lower levels of psychological suffering. In any case, the extent to which different functional properties of the IRAP predict psychological suffering will be an important avenue for future research.

Chapter 4. The STTDE was again observed in the data from Experiment 1 of Chapter 4. In the context of both the self-focused and other-focused IRAPs it is assumed that in general participants oriented toward face-related stimuli more strongly than toward pen-related stimuli (see Borji, Sihite, & Itti, 2013; Hershler, & Hochstein, 2005; Santos, Mier, Kirsh, & Meyer-Lindenberg, 2011)[1]. Specifically, the *D*-score for the *Face words-Face picture* trial type was significantly larger than for the other three trial types. The fact that the *Face words-Face picture* effect was larger than both *Face words-Pen picture* and *Pen words-Face picture* may be deemed relatively predictable because the latter two trial types do not share the same response option with the former

[1] We use the term 'face-related' stimuli based on the assumption that words co-ordinated with pictures of faces in the context of the IRAP would possess some of the functions (e.g. orienting) of those pictures, via a transfer of functions.

trial type within blocks of trials. During consistent bocks, for example, the *Face words-Face picture* trial type requires the participant to respond *Yes*, whereas the *Face words-Pen picture* and *Pen words-Face picture* trial types require the response *No.* What is more difficult to explain, however, is the difference between the *Face words-Face picture* and the *Pen words-Pen picture,* because they both share the same response option within IRAP blocks (i.e. *Yes* during consistent blocks and *No* during inconsistent blocks).

At this point, we have used the DAARRE model to explain the STTDE that emerged once again in Experiment 1 of Chpater 4 (as had been the case in Experiment 1, Chapter 2), but another differential trial type effect also emerged here in Experiment 1. which would be difficult to explain in terms of the response options alone. Specifically, participants tended to deny more readily than confirm that a picture of a pen was a face, but this was not the case (at a statistically significant level) when denying that a picture of a face was a pen. Indeed, post-hoc analyses indicated that these two trial types differed significantly from each other. How might we explain this difference, given that both trial types required the same RCI within blocks of trials? Once again, the DAARRE model may be of use here. Specifically, the *Face words-Pen pictures* trial type presents a target stimulus that coheres with the RCI in terms of its Cfunc properties, whereas the *Pen words-Face picture* trial type does not. If we assume that the spatial contiguity between the target stimulus and the response option plays a role in determining the IRAP effect, the difference in trial type effects observed here makes sense. More informally, participants may experience a '*Yes-No-No*' reaction to the *Face words-Pen pictures* trial type, but a '*No-Yes-No*' reaction to the *Pen words-Face picture* trial type, assuming that in general they read each IRAP trial from top-down. If participants find it easier to select an RCI that is functionally similar to the target stimulus they have just observed, than an RCI that is functionally *dissimilar*, the larger effect for the *Face words-Pen pictures* trial type is readily predicted.

Of course, the DAARRE model explanation presented above is post-hoc and speculative, but it seems important to present it here because it is a pattern we have

observed in other studies, in which the DAARRE model can be applied in a relatively straightforward way. Perhaps other researchers, therefore, who are using the IRAP, may find the interpretation offered here of some use in attempting to explain and explore similar effects. In any case, it seems important to continue to develop increasingly sophisticated functional analyses of the IRAP in terms of the cluster of variables that produce the patterns we observe with the measure. Indeed, this seems particularly important because the IRAP has been used widely as a measure in clinical, health, forensic and social psychological research (e.g. see Vahey et al. 2015 for a recent meta-analysis).

Grappling with the findings from Chapter 3. It is considerably more difficult to attempt to use the DAARRE model to explain the effects observed in Experiments 2-6 of Chapter 3, primarily because of the use of the NL-IRAPs. Nonethless, the unexpected pattern that was generally consistent across all six experiments, in terms of the *Negative Event-Positive Reaction* trial type yielding the weakest effect, still needs to be grappled with. The STTDE cannot readily be applied here because that effect is specific to the two trial types that require a "Yes" response during history-consistent blocks. The critical issue here is that across all six experiments the effect for the *Negative Event-Positive Reaction* trial type was weaker than for the *Positive Event-Negative Reaction* trial type, and both of these trial types require a "No" response during history-consistent blocks. Why would they differ, given that they both require the same response within blocks? This question seems particularly interesting because the NL-IRAP was employed here, and thus an explanation that appeals to the functions of separate label and target stimuli within a traditional IRAP cannot be applied. Nevertheless, given the consistency in this weak effect for the *Negative Event-Positive* trial type, across the six studies, it remains a pattern that should be targeted for systematic analyses in future research.

It is also important to note that overall the two IRAPs employed throughout the empirical work in Chapter 3 did produce performances that would be deemed consistent with the pre-experimental histories of the participants. That is, partiicpants tended to

confirm more readily than to deny that positive event was posistive and a negative event was negative. They also tended to deny more readily than confirm that a positive event was negative and a negative event was posistive. But, why did the IRAPs not distinguish in a clear and consistent way between the perspective of self versus other. One possibility is that the IRAPs simply tapped into 'sense-making.' For example, confirming, rather than denying that positive events evoke positive reactions in the self and others makes sense in natural language. The only caveat to this interpretation is that significant differences between the IRAPs were observed in Experiments 1 and 2 when other was unspecified, but not in Experiments 3-6 when other was specified. Perhaps, therefore, there was some sensitivity to self versus other, but the use of complex statements in the IRAPs potentially undermined or reduced the impact of deictic relational responding per se. That is, in presenting such complex stimuli or networks in the IRAP, participants more or less interpreted the task as a sense-making or problem-solving task, in which the self versus other had little or no impact, particularly when the other was specified. The challenge going forward, therefore, is to develop IRAPs that maintain the deictic functions of self and other, in the context of perspective-taking rather than simple sense-making.

It is important at this point to draw attention to a pattern in the data that only came to light during the publication review process, and which might indicate that the two NL-IRAPs may have been sensitive, to some degree, to self versus others, when other was clearly distinct from self. Specifically, the trial type effect for *Positive Event-Negative Reaction* was significantly different from zero for the self-focused IRAPs, but not for the other-focused IRAPs, across Experiments 5 and 6. This was not the case in the four remaining experiments, in which other was unspecified or specified as similar to self. In a purely post-hoc analysis, we compared the difference between the self- and other-focused IRAPs for the *Positive Event-Negative Reaction* trial type, (Self; M = .164 SD = .324, Other; M = .032 SD = .333, $t(59) = 2.195$ p =.032. 95% CI [.024, .54], BF1 = 2.560). We raise this issue here because this pattern emerged across the two experiments in which other was clearly distinct from self. If this finding is indeed robust,

it could indicate that participants found it easier to respond ‘False’ when self, rather than a very distinct other, was co-ordinated in some way with a negative reaction. In simple terms, participants showed a bias toward denying that they are negative that was stronger than that bias for very different others. Although subtle, pursuing this potentially interesting effect could be useful in future research.

An Alternative Strategy for Conceptualising and Analysing Perspective-taking Based on a Multi-Dimensional Multi-Level (MDML) Framework for RFT

As noted in the introductory chapter to the current thesis, the approach taken to studying perspective-taking within RFT has focused heavily on the three deictic relations (I-YOU, HERE-THERE and NOW-THEN). Indeed, the development of the original perspective-taking protocol involved targeting exactly these relations. Specifically, the protocol asked participants to respond in accordance with these relations in simple and/or reversed form. As discussed in Chapter 2, more recent work on perspective-taking within RFT has employed the IRAP (see Barbero-Rubio et al., 2016) or has involved deriving spatial perspective-taking relations by exposing adult participants to a complex series of graduated ‘mental rotation’ tasks (see Guinther, 2017, 2018). In all of this work, however, there has been little if any effort to develop a technical, conceptual RFT-based analysis of tasks typically used by mainstream psychology to assess perspective-taking. For example, it remains unclear exactly what relational repertoires seem to be required to complete the types of task that aim to assess what is described as false belief, such as the Deceptive Container Task or the Sally-Anne Test, which was employed in the research reported in Chapter 4 of the current thesis. Throughout the empirical research reported herein, there was an on-going effort to systematise the increasing complexity in patterns of relational responding that have been identified within RFT and critically, in the current context, to begin to specify exactly what relational abilities are required for a child or individual to successfully complete false belief tasks. To appreciate how the systematic analysis of relational

responding leads to a potentially more precise analysis of perspective-taking it will first be necessary to outline the general framework, to which we now turn.

The Multi-Dimensional Multi-Level (MDML) Framework. In an effort to systematise the RFT account and to emphasise the relevant behavioural dynamics involved in AARR (Barnes-Holmes, Barnes-Holmes, Luciano, & McEnteggart, 2017), a multi-dimensional multi-level (MDML) framework has been proposed (see Table 2). According to this framework, AARR may be conceptualised as developing in a broad sense from; (i) mutual entailment, to (ii) simple networks involved in frames, to (iii) more complex networks involved in rules and instructions, to (iv) the relating of relations involved in analogical reasoning, and finally to (v) relating relational networks, which is typically involved in understanding and producing extended narratives, and advanced problem-solving. In identifying these as different levels, the MDML framework is not indicating that they are rigid or invariant 'stages.' Rather, lower levels are seen as containing patterns of AARR that may provide an important historical context for the patterns of AARR that occur in the levels above. In general, the different levels are based on a combination of well-established assumptions within RFT and, where possible, empirical evidence. The framework also conceptualises each of these levels as having multiple dimensions: *coherence*, *complexity*, *derivation* and *flexibility*. Each of the levels is seen as intersecting with each of the dimensions yielding a framework that consists of 20 units of analysis for conceptualising and studying the dynamics of AARR in the laboratory and in natural settings.

Table 2

Overview of the Multi-Dimensional Multi-Level (MDML) Framework

LEVELS	**DIMENSIONS**			
	Coherence	**Complexity**	**Derivation**	**Flexibility**
Mutually Entailing	Coherence/Mutual Entailment	Complexity/Mutual Entailment	Derivation/Mutual Entailment	Flexibility/Mutual Entailment
Relational Framing	Coherence/Frame	Complexity/Frame	Derivation/Frame	Flexibility/Frame
Relational Networking	Coherence/Network	Complexity/Network	Derivation/Network	Flexibility/Network
Relating Relations	Coherence/Relating Relations	Complexity/Relating Relations	Derivation/Relating Relations	Flexibility/Relating Relations
Relating Relational Networks	Coherence/Relating Relational Networks	Complexity/Relating Relational Networks	Derivation/Relating Relational Networks	Flexibility/Relating Relational Networks

A brief description of each of the four dimensions is as follows. *Coherence* refers to the extent to which specific patterns of AARR are generally consistent with other patterns of AARR. For example, the statement "A car is larger than a truck" would typically be seen as lacking coherence with the relational networks that operate in the wider verbal community. Note, however, that such a statement may be seen as coherent in certain contexts (e.g. when playing a game of 'everything is opposite'). *Complexity* refers to the level of detail or density of a particular pattern of AARR. As a simple example, the mutually entailed relation of co-ordination may be seen as less complex than the mutually entailed relation of comparison because the former involves only one type of relation (e.g. if A is the *same as* B then B is the *same as* A), but the latter involves two types of relations (if A is *bigger than* B, then B is *smaller than* A). *Derivation* refers to how well practiced a particular instance of AARR has become. Specifically, when a pattern of AARR is derived for the first time it is, by definition, highly derived (i.e. novel or emergent) and thus derivation reduces as that pattern becomes more practiced. *Flexibility* refers to the extent to which a given instance of AARR may be modified by current contextual variables. Imagine a young child who is asked to respond with the wrong answer to the question "Which is bigger, a car or a truck?" The easier this is achieved, the more flexible the AARR.

A detailed treatment of the MDML framework has been provided elsewhere (e.g. Barnes-Holmes et al., 2017) and thus there is no need to work through all the details and subtleties here. The key purpose in presenting the MDML framework in the current chapter is to show how it has been used to develop the beginnings of an RFT-based conceptual analysis of the basic false belief task (e.g. the Deceptive Container Task and Sally-Anne Test), which we will now consider.

A conceptual functional analysis of the False Belief Task based on the MDML framework. The following is 'educated guess work' at providing a functional analysis, based on the MDML framework, of the AARR that is required to 'solve' a classic false belief task. Where appropriate, the minimal levels of relational development specified within the MDML framework are indicated below. Specifically, we will argue that solving a false belief task involves the highest level of relational

development (Level 5). We recognise that one could directly train correct responding on a false belief task through explicit reinforcement, direct instruction and prompting, etc. However, obtaining correct responses that involve 'genuinely' understanding that others may believe something that is false if they have not seen what the self has seen would seem to require the relating of relational networks.

The critical relational pre-cursors. Before presenting the full MDML-based model of false belief, it seems important to identify what we see as the key relational pre-cursors and the levels of relational development at which they need to be observed. First, the three basic relational frames of co-ordination, distinction and temporality would be required, thus involving Levels 1 and 2 of the MDML framework. It would be important that relational responding involved in these frames is relatively high in coherence and complexity, but relatively low in derivation and flexibility. More informally, (1) when events are related as co-ordinate, distinct or occurring in some temporal order, these patterns of relational responding should be consistent with many other instances of previous and current relevant patterns of such responding (high coherence); (2) the three classes of relational responding should be relatively sensitive to various forms of contextual control (high complexity); (3) each of the three classes should have relatively extended or protracted behavioural histories (low derivation); and (4) the three patterns should persist in the absence of direct reinforcement, prompting, instruction and even modest levels of punishment (low flexibility).

In addition to the three basic relational frames discussed above, the core deictic relational frame of I-YOU, HERE-THERE and NOW-THEN would need to be firmly established in the behavioural repertoire. Strictly speaking, as a frame this would be located at Level 2 of the MDML framework, but in ensuring that the frame was firmly established would likely have required that the basic framing participated in larger relational networks, thereby locating it at Level 3 at least. More informally, the 'I' would have been related to many other individuals (rather than only one other) in many different times and places. Technically speaking, this would involve ensuring again relatively high levels of coherence and complexity, with low levels of derivation and flexibility. For example, if I told someone that I was on a train at noon going from

Dublin to Cork, this statement would *cohere* with the conclusion that I will arrive in Cork at approx. 3pm. If, however, I qualified that I would have to change at Mallow to complete the trip, the conclusion about the arrival time would be adjusted to approx. 3.30pm, which would require of course an increased level of *complexity* in the contextual control over the networking involved. Appropriate responding in this example would almost certainly involve low levels of derivation and flexibility. That is, there would be many broadly similar instances of informing listeners of where you are, where you are going, what time you expect to arrive, and any other qualifying conditions that would then allow both you and the listener to co-ordinate your activities.

The final critical relational pre-cursor would require causal or if-then frames and appropriate transformations of function, involving the deictic relations specified above, thus again involving relational development at Levels 1, 2 and 3, at minimum, of the MDML framework. The specific causal relation and transformations of function could be described as 'seeing leads to knowing' or in other words 'if I see an event occur, then I know that it occurred.' Ideally, this particular if-then frame would network with the deictic frame in a complex network, so that I could derive 'if you and I see an event occur, then both you and I know that it occurred.' Again, it would be necessary for this type of complex networking to be relatively high in coherence and by definition complexity, and low in derivation and flexibility. Indeed, these dimensional requirements could be seen as essential if a full-blown understanding of false belief is to be observed on relevant tasks.

Even with all of the foregoing pre-cursors in place, false belief understanding appears to require relational responding that is clearly located at both Levels 4 and 5 of the MDML framework. To appreciate why this is the case, the reader should examine Figure 2, which provides a graphical representation of the suggested functional-analytic processes involved in responding correctly to the classic Unexpected Location (false belief) Task. The bullet points below provide an additional description of the graphhical representation.

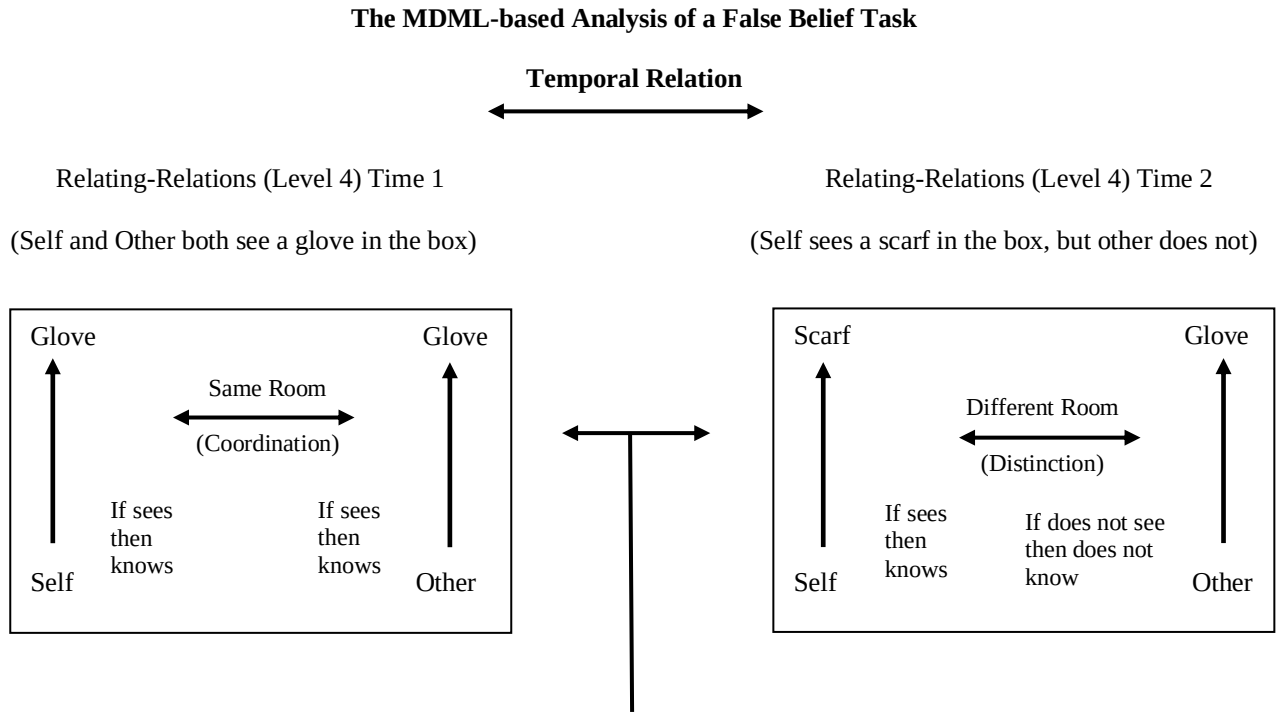

Figure 2. Graphical representation of the suggested functional-analytic processes involved in responding correctly to a classical false belief task.

- The left-hand side of the figure indicates that initially (at Time 1) both the self and other observe a glove being placed into a box; based on this and the relational pre-cursors described above, the self can conclude that both self and other know that there is a glove in the box.
- The right-hand side of the figure indicates that subsequently (at Time 2) the self observes the glove in the box being replaced with a scarf when the other is not in the room; based on this and the relational pre-cursors noted above, the self can conclude that only the self will know that there is a scarf (rather than a glove) in the box.
- The double-headed arrow linking the left and right sides of the figure indicates that responding correctly to the false belief task requires that the self relates the two relational networks as distinct in terms of what the self and other know after Time 2. The critical point here is that if the self simply reported that the other *does not know* what is in the box after Time 2, that would indicate relating relations which is best located at Level 4 of the MDML framework. If, however, the self reports that the other thinks that the box contains a glove, that requires the relating of relations at Time 2 to the relating of relations at Time 1. More informally, the self has to understand that what the other knew at Time 1 is what they still think at Time 2.

Deconstructing a classic false belief task using the MDML framework, as we have done here, clearly reveals the layers of complexity involved in this widely used task and may explain to some extent why many young children struggle to solve it correctly. The potential individual differences in levels of coherence, complexity, derivation and flexibility among the relational pre-cursors discussed above may also help explain why the literature on perspective-taking contains such wide variation in the ages at which false belief tasks can be solved correctly, and why performances vary widely depending on the precise variation of the task that is presented. At this point, we should be clear that the current MDML-based interpretation of the false belief task remains highly speculative. Nevertheless, we present it here because it seems to suggest numerous ways in which applied researchers and practitioners could approach the training and establishment of relatively robust false belief responding in individuals who find the task challenging.

If the foregoing MDML-based analysis of perspective-taking is at least partially correct, it may also help to explain why it has proven so difficult, even in our own research here, to capture perspective-taking using the IRAP. Consider two key issues in this regard. 1. Responding to the IRAP requires participants to respond under time pressure (typically < 2000ms). When considering the above MDML-based analysis it would appear unlikely that participants could engage in such complex relational responding within such a short time period. 2. The repeated presentation of similar trial types may also impact upon the likelihood of particiapnts responding with 'genuine perspective-taking'. For example, it may be the case that participants initially engage in 'genuine perspective-taking' during the first trials of the IRAP, but thereafter simply maintain 'correct' responding. In other words, participants can respond correctly across latter trials that require mutual entailment, but they are not relating relational networks. Overall therefore, it maybe the case that the IRAP as a methodology is, by its own structure, limited in its ability to capture perspective-taking 'in flight'.

Conclusions

In hindsight, the aim of the current thesis to develop an IRAP that could capture perseptvie-taking 'in flight' was extremely ambitious. The results from the eleven experiments reported here suggest that the IRAP, in its current format, is limited in its ability to capture genuine perspective-taking. However, as a journey towards understanding, in considerable technical detail, why this is the case, we have learned a great deal both about the IRAP itself and ultimately about the potential behavioural processes involved in perspective-taking as derived relational responding. We have also gained new insights into the wide inconsistencies that have marred the huge mainstream literature on persepective-taking. In these respects, this journey has been more than worthwhile, although frustrating at times. Going forward, it appears vital that we develop a clear functional analysis of what participants are required to do when presented with a perspective-taking task. The General Discussion of the current thesis has provided an initial attempt to do just that. The empirical work and discussion here also indicated that funcational analyses of the IRAP must continue, if we are to

understand the effects observed using this procedure. The on-going conceptual developments with RFT, of which the current thesis is one strand, are essential in this regard.

References

Ames, D. R. (2004). Inside the mind reader's tool kit: Projection and stereotyping. *Journal of Personality and Social Psychology, 87(3),* 340-353. DOI: 10.1037/0022-3514.87.3.340

Andrews, G., Halford, G. S., Bunch, K. M., Bowden, D., & Jones, T. (2003). Theory of mind and relational complexity. *Child Development, 74(5),* 1476-1499. DOI:10.1111/1467-8624.00618

Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition, 106,* 1093-1108. DOI:10.1016/j.cognition.2007.05.005

Apperly, I. A., Riggs, K. J., Simpson, A., Chiavarino, C., & Samson, D. (2006). Is belief reasoning automatic? *Psychological Science, 17(10),* 841-844. DOI: 10.1111/j.1467-9280.2006.01791.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63,* 596-612. DOI: 10.1037/0022-3514.63.4.596

Arriaga, X. B, & Rusbult, C. E. (1998). Standing in my partner's shoes: Partner perspective taking and reactions to accommodative dilemmas. *Personality and Social Psychology Bulletin, 24(9),* 928-948. DOI: 10.1177/0146167298249002

Arsenio, W. F. (2003). Emotional intelligence and the intelligence of emotions: A developmental perspective on mixed models, *Human Development, 46(2-3),* 97-103. DOI: 10.1159/000068582

Barbero-Rubio, A., López-López, J. C., Luciano, C., & Eisenbeck, N. (2016). Perspective-taking measured by implicit relational assessment procedure (IRAP). *The Psychological Record, 66(2),* 243-252. DOI: 10.1007/s40732-016-0166-3

Barnes-Holmes, D., Barnes-Holmes, Y., Hussey, I., & Luciano, C. (2016). Relational frame theory: Finding its historical and philosophical roots and reflecting upon its future development: An introduction to Part II. In R. D. Zettle, S. C. Hayes, P. M. D. Barnes-Holmes, & A. Biglan (Eds). *The Wiley handbook of contextual*

behavioural science (pp.117-128). West-Sussex, UK: Wiley-Blackwell. DOI:10.1002/9781118489857.ch8

Barnes-Holmes, D., Barnes-Holmes, Y., Luciano, C., & McEnteggart, C. (2017). From the IRAP and REC model to a multi-dimensional multi-level framework for analysing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science, 6(4),* 434-445. DOI: 10.1016/j.jcbs.2017.08.001

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,* 527-542. DOI:10.1007/BF03395726

Barnes-Holmes, D., Finn, M., McEnteggart, C., & Barnes-Holmes, Y. (2018). Derived stimulus relations and their role in a behavior-analytic account of human language and cognition. *Perspectives on Behavioral Science (Special Issue on Derived Relations), 41*(1), 155-173. DOI: 10.1007/S40614-017-0124-7

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,* 497–515. Available at: https://opensiuc.lib.siu.edu/tpr/vol58/iss4/1

Barnes-Holmes, Y. (2001). *Analysing relational frames: Studying language and cognition in young children* (Unpublished doctoral thesis). National University of Ireland Maynooth.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., Strand, P., & Friman, P. (2004). Establishing relational responding in accordance with more-than and less-than as generalized operant behavior in young children. *International Journal of Psychology and Psychological Therapy, 4,* 531-558.

Baron-Cohen, S. (2001). Theory of mind in normal development and autism. *Prisme, 34,* 174-183.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *The Journal of Child Psychology and Psychiatry, 30(2),* 285-297. DOI: 10.1111/j.1469-7610.1989.tb00241.x

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"?, *Cognition, 21*, 37-46. DOI: 10.1016/0010-0277(85)90022-8

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29(5),* 407-418. DOI: 10.1023/A:1023035012436

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience (2nd ed.).* Oxford: Oxford University Press.

Barron, B. F., Verkuylen, L., Belisle, J., Paliliunas, D., & Dixon, M. R. (2018). Teaching "then-later" and "here-there" relations to children with autism: An evaluation of single reversals and transformation of stimulus function. *Behavior Analysis in Practice*. Advance online publication. DOI: 10.1007/s40617-018-0216-1

Belisle, J., Dixon, M. R., Stanley, C. R., Munoz, B., & Daar, J. H. (2016). Teaching foundational perspective-taking skills to children with autism using the PEAK-T curriculum: Single-reversal "I-you" deictic frames. *Journal of Applied Behavior Analysis. 49(4),* 965-969. DOI: 10.1002/jaba.324

Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis, 40(1),* 45-71. DOI: 10.1901/jaba.2007.7-06

Bernaerts, I., De Groot, F., & Kleen, M. (2012). De AAQ-II, een maat voor experiëntiële vermijding: Normering bij jongeren. *Gedragstherapie, 4,* 389-399.

Bernstein, A., Hadash, Y., Lichtash, Y., Tanay, G., Shepherd, K., & Fresco, D. M. (2015). Decentering and related constructs: A critical review and metacognitive process model. *Perspectives on Psychological Science, 10(5),* 599-617. DOI: 10.1177/1745691615594577

Bernstein, R. F., Laurent, S. N., Nelson, B. W., & Laurent, H. K. (2015). Perspective-taking induction mitigates the effects of partner attachment avoidance on self-partner overlap. *Personal Relationships, 22*, 356-367. DOI: 10.1111/pere.12085

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18(5),* 382-386. DOI: 10.1111/j.1467-9280.2007.01909.x

Bond, F., Hayes, S., Baer, R., Carpenter, K., Guenole, N., Orcutt, H., … Zettle, R. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*, 676-88. DOI: 10.1016/j.beth.2011.03.007

Boucher, J. (2012). Putting theory of mind in its place: Psychological explanations of the socio-emotional-communicative impairments in autistic spectrum disorder. *Autism, 16(3),* 226-246. DOI: 10.1177/1362361311430403

Borg, I., & Groenen, P. (2005). *Modernm multidimensional scaling: Theory and applications (2nd ed.).* New York: Springer-Verlag.

Bradford, E. E. F., Jentzsch, I., & Gomez, J. (2015). From self to social cognition: Theory of mind mechanisms and their relation to executive function. *Cognition, 138,* 21-34. DOI: 10.1016/j.cognition.2015.02.001

Brassil, N., Hyland, J., O'Hora, D., & Stewart, I. (2019). Reversing time and size: Mutual entailment of nonarbitrary temporal and magnitude relational responding. *The Psychological Record, 69(1),* 95-105. DOI: 10.1007/s40732-018-0323-y

Brino, A. L., Campos, R. S., Galvão, O. F., & McIlvane, W. J. (2014). Blank-comparison matching-to-sample reveals a false positive symmetry test in a capuchin monkey. *Psychology & Neuroscience, 7*(2), 193-198. DOI: 10.3922/j.psns.2014.008

Broekhof, E., Ketelaar, L., Stockman, L., van Zijp, A., Bos, M. G., & Rieffe, C. (2015). The understanding of intentions, desires and beliefs in young children with

autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45(S7)*, 2035-2045. DOI: 10.1007/s10803-015-2363-3

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language and explicit theory of mind. *Journal of Experimental Child Psychology, 130,* 67-78. DOI: 10.1016/j.jecp.2014.09.010

Bull, R., Phillips, L. H., & Conway, C. A. (2008). The role of control functions in mentalizing: Dual-task studies of theory of mind and executive function. *Cognition*, 107, 663-672. DOI:10.1016/j.cognition.2007.07.015

Call, J., & Tomasello, M. (2008). Does the Chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12(5),* 187-192. DOI: /10.1016/j.tics.2008.02.010

Carpendale, J. I., & Lewis, C. (2006). *How children develop social understanding*. Malden, UK: Blackwell Publishing.

Carr, D., Wilkinson, K. M., Blackman, D., & McIlvane, W. J. (2000). Equivalence classes in individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of Behavior, 74(1),* 101-114. DOI: 10.1901/jeab.2000.74-101

Cassidy, S., Roche, B., & Hayes, S. (2011). A relational frame training intervention to raise intelligence quotients: A pilot study. *The Psychological Record, 61(2),* 173-198. DOI: 10.1007/BF03395755

Charlop-Christy, M. J., & Daneshvar, S. (2003). Using video modelling to teach perspective taking to children with autism. *Journal of Positive Behavior Intervention, 5(1),* 12-21. DOI: 10.1177/10983007030050010101

Colonnesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology, 26*, 561-577. DOI:10.1348/026151008X285660

Corcoran, R., Mercer, G., & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: Investigating "theory of mind" in people with schizophrenia. *Schizophrenia Research, 17(1),* 5-13. DOI: 10.1016/0920-9964(95)00024-G

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70(4),* 853-865. DOI: 0009-3920/99/7004-0004

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.

Davis, M. H., Soderlund, T., Cole, J., Gadol, E., Kute, M., Myers, & Weihing, J. (2004). Cognitions associated with attempts to emphasize: How do we imagine the perspective of another? *Personality and Social Psychology Bulletin, 30(12),* 1625-1635. DOI: 10.1177/0146167204271183

Davlin, N. L., Rehfledt, R. A., & Lovett, S. (2011). A relational frame theory approach to understanding perspective-taking using children's stories in typically developing children. *European Journal of Behavior Analysis, 12(2),* 403-430. DOI: 10.1080/15021149.2011.11434392

DeBernardis, G. M., Hayes, L. J., & Fryling, M. J. (2014). Perspective taking as a continuum. *The Psychological Record, 64(1),* 123-131. DOI: 10.1007/s40732-014-0008-0

De Corte, K., Buysee, A., Verhofstadt, L., Roeyers, H., Ponnet, K., & Davis, M., (2007). Measuring empathetic tendencies: Reliability and validity of the Dutch version of the Interpersonal Reactivity Index (IRI). *Psychologica Belgica*, *47*, 235-260. DOI: 10.5334/pb-47-4-235S

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

Dixon, M. R., & Zlomke, K. M. (2005). Using the precursor to the relational evaluation procedure (PREP) to establish the relational frames of sameness, opposition, and distinction. *Revista Latinoamericana de Psicología, 37(2),* 305-316.

Dube, W. V., MacDonald, R. P. F., Mansfield, R. C., Holcomb, W. L., & Ahearn, W. H. (2004). Toward a behavioral analysis of joint attention. *The Behavior Analyst, 27(2),* 197-207. DOI: 10.1007/BF03393180

Dunn, J., Bretherton, I., & Munn, P. (1987). Conversations about feeling states between mothers and their young children. *Developmental Psychology, 23 (1),* 132-139. DOI: 10.1037/0012-1649.23.1.132

Dunn, J., & Hughes, C. (1998). Young children's understanding of emotions within close relationships. *Cognition and Emotion, 12(2),* 171-190. DOI: 10.1080/026999398379709

Dunne, S., Foody, M., Barnes-Holmes, Y., Barnes-Holmes, & Murphy, C. (2014). Facilitating repertoires of coordination, opposition, distinction, and comparison in young children with autism. *Behavioral Development Bulletin, 19(2),* 37-47. DOI: 10.1037/h0100576

Dymond, S., Roche, B., Forsyth, J. P., Whelan, R., & Rhoden, J. (2007). Transformation of avoidance response functions in accordance with same and opposite relational frames. *Journal of the Experimental Analysis of Behavior, 88(2),* 249-262. DOI: 10.1901/jeab.2007.88-249

Eisenberg, N., Spinrad, T. L., & Sadovsky, A. (2006). Empathy-related responding in children. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 517-549). Mahwah, NJ: Erlbaum.

Epley, N. (2004). A tale of tuned decks? Anchoring as accessibility and anchoring as adjustment. In. D. J., Koehler, & N. Harvey (Eds), *Handbook of judgement and decision making* (pp. 240-257). Oxford, England: Blackwell.

Epley, N., & Caruso, E. M. (2009). Perspective-taking: Misstepping into others' shoes. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds), *Handbook of imagination and mental simulation* (pp. 297-311). Hove, England: Psychology Press.

Epley, N., Caruso, E. M., & Bazerman, M. H. (2006). When perspective taking increases taking: Reactive egoism in social interaction. *Journal of Personality and Social Psychology, 91,* 872–889. DOI: 10.1037/0022-3514.91.5.872

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 17(3),* 327-339. DOI: 10.1037/0022-3514.87.3.327

Farrant, B. M., Devine, T. A., Maybery, M. T., & Fletcher, J. (2012). Empathy, perspective taking and prosocial behaviour: The importance of parenting practices. *Infant and Child Development, 21(2),* 175-188. DOI: 10.1002/icd.740

Fertuck, E. A., Jekal, A., Song, I., Wyman, B., Morris, M. C., Wilson, S. T.,...Stanley, B. (2009). Enhanced 'Reading the Mind in the Eyes' in borderline personality disorder compared to healthy controls. *Psychological Medicine, 39(12),* 1979-1988. DOI: 10.1017/S003329170900600X

Finn, M., Barnes-Holmes, D., & McEnteggart, C. (2018). Exploring the single-trial-type-dominance-effect in the IRAP: Developing a differential arbitrarily applicable relational responding effects (DAARRE) Model. *The Psychological Record, 68(1),* 11-25. DOI: 10.1007/s40732-017-0262-z

Flavell, J. H. (1977). The development of knowledge about visual perception. *Nebraska Symposium on Motivation, 25,* 43-76.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1– Level 2 distinction. *Developmental Psychology*. *17(1),* 99–103. DOI: 10.1037/0012-1649.17.1.99

Flavell, J. H. (1992). Perspectives on perspective-taking. In H. Beilin, & P. B. Pufall (Eds.), *Piaget's theory: Prospects and possibilities. The Jean Piaget symposium* series (Vol. 14, pp. 107–139). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (2004). Theory-of-mind development: Retrospect and prospect. *Merrill-Palmer Quarterly, 50(3),* 274–290. Retrieved from https://search.proquest.com/docview/230095187?accountid=11077

Flobbe, L., Verbrugge, R., Hendricks, P., & Krämer, I. (2008). Children's application of theory of mind in reasoning and language. *Journal of Logic, Language and Information, 17(4),* 417-422. DOI: 10.1007/s10849-008-9064-7

Fraley, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The Experiences in Close Relationships-Relationship Structures questionnaire: A method for assessing attachment orientations across relationships. *Psychological Assessment, 23,* 615-625. DOI:10.1037/a0022898

Franck, E., De Raedt, R., Barbez, C., & Rosseel, Y. (2008). Psychometric properties of the Dutch Rosenberg Self-esteem Scale. *Psychologica Belgica, 48*, 25-35. DOI: 10.5334/pb-48-1-25

Fretland, R. A., Andeson, S., Sundet, K., Andreassen, O. A., Melle, I., & Vaskin, A. (2015). Theory of mind in schizophrenia: Error types and associations with symptoms. *Schizophrenia Research, 162*, 42-46. DOI: 10.1016/j.schres.2015.01.024

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: development of perspective-taking abilities in 4- to 8-year-olds. *Frontiers in Psychology, 5*, 1-7. DOI: 10.3389/fpsyg.2014.00386

Frith, U., Morton, J., & Leslie, A. M. (1991). The cognitive basis of a biological disorder: Autism. *Perspectives on Disease, 14(10)*, 433-438.

Galinsky, A. D., & Ku, G. (2004). The effects of perspective-taking on prejudice: The moderating role of self-evaluation. *Personality and Social Psychology Bulletin, 30(5)*, 594-604. DOI: 10.1177/0146167203262802

Galinsky, A.D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective-taking and empathy in negotiations. *Psychological Science, 19(4)*, 378-384. DOI: 10.1111/j.1467-9280.2008.02096.x

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: The role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology, 81(4)*, 657–669. DOI: 10.1037//OO22-3514.81.4.657

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in 'theory of mind': Evidence from compromised belief-desire reasoning in old age. *Cognition, 101(1)*, 129-152. DOI: 10.1016/j.cognition.2005.05.007

Gil, E., Luciano, C., Ruiz, F. J., & Valdivia-Salas, S. (2012). A preliminary demonstration of transformation of functions through hierarchical relations. *International Journal of Psychology and Psychological Therapy, 12(1)*, 1-19.

Gil, E., Luciano, C., Ruiz, F. J., & Valdivia-Salas, S. (2014). A further experimental step in the analysis of hierarchical responding. *International Journal of Psychology and Psychological Therapy, 14(2),* 137-153.

Gilroy, S. P., Lorah, E. R., Dodge, J., & Fiorello, C. (2015). Establishing deictic repertoires in autism. *Research in Autism Spectrum Disorder, 19,* 82-92. DOI: 10.1016/j.rasd.2015.04.004

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of "sameness" and "difference" in an insect. *Nature, 410,* 930-933. DOI:10.1038/35073582

Gómez-Becerra, I., Martín, M. J., Chávez-Brown, M., & Greer, R. D. (2007). Perspective taking in children with autism. *European Journal of Behaviour Analysis, 8(1),* 13-28. DOI: 10.1080/15021149.2007.11434270

Gopnik, A., & Slaughter, V. (1991). Young children's understanding of changes in their mental states. *Child Development, 62(1*), 98–110. DOI: 10.2307/1130707

Gore, N. J., Barnes-Holmes, Y., & Murphy, G. (2010). The relationship between intellectual functioning and relational perspective-taking. *International Journal of Psychology and Psychological Therapy, 10(1),* 1-17.

Gorham, M., Barnes-Holmes, Y., & Barnes-Holmes, D. (2009). Derived comparative and transitive relations in young children with and without autism. *The Psychological Record, 59(2),* 221-246. DOI: 10.1007/BF03395660

Gould, E., Tarbox, J., O'Hora, D., Noone, S., & Bergstrom, R. (2011). Teaching children with autism a basic component skill of perspective-taking. *Behavioral Interventions, 25,* 50-66. DOI: 10.1002/bin.320

Grosse Wiesmann, C., Schreiber, J., Singer, T., Steinbesi, N., & Friederici, A. D. (2017). White matter maturation is associated with the emergence of Theory of Mind in early childhood, *Nature Communications, 8(14692*), 1-10. DOI: 10.1038/ncomms14692

Griffee, K., & Dougher, M. J. (2002). Contextual control of stimulus generalization and stimulus equivalence in hierarchical categorization. *Journal of the*

Experimental Analysis of Behavior, 78(3), 433-447. DOI: 10.1901/jeab.2002.78-433

Guinther, P. M. (2017). Contextual influence over deriving others' true beliefs using a relational triangulation perspective-taking protocol (RT-PTP-M1). *Journal of the Experimental Analysis of Behavior, 108(3),* 433-456. DOI: 10.1002/jeab.291

Guinther, P. M. (2018). Contextual influence over deriving another's false beliefs using a relational triangulation perspective taking protocol. *Journal of the Experimental Analysis of Behavior, 110(3),* 500-521. DOI: 10.1002/jeab.480

Hahs, A. D. (2015). Teaching prerequisite perspective-taking skills to children with autism. *International Journal of Psychology and Behavioral Sciences, 5(3),* 115-120. DOI: 10.5923/j.ijpbs.20150503.02

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Science, 21(6*), 803-865.

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66(3),* 843-855. DOI: 10.1111/j.1467-8624.1995.tb00909.x

Harrington, L., Siegert, R. J., & McClure, J. (2005). Theory of mind in schizophrenia: A critical review. *Cognitive Neuropsychiatry, 10(4),* 249-286. DOI: 10.1080/13546800444000056

Harwood, M. D., & Farrar, M. J. (2006). Conflicting emotions: The connection between affective perspective taking and theory of mind. *British Journal of Developmental Psychology, 24(2),* 401-418. DOI: 10.1348/026151005X50302

Hayes, J., Stewart, I., & McElwee, J. (2016). Assessing and training young children in same and different relations using the relational evaluation procedure (REP). *The Psychological Record, 66,* 547-561. DOI 10.1007/s40732-016-0191-2

Hayes, S. C. (1984). Making sense of spirituality. *Behaviorism, 12,* 99-110.

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.

Hayes, S. C., & Wilson, K. G. (1993). Some applied implications of a contemporary behavior-analytic account of verbal events, *The Behavior Analyst, 16,* 283-301. DOI: 10.1007/BF03392637

Heagle, A. I., & Rehfeldt, R. A. (2006). Teaching perspective-taking skills to typically developing children through derived relational responding. *Journal of Early Intensive Behavioral Intervention, 3,* 1-34. DOI: 10.1037/h0100321

Hendriks, A. L., Barnes-Holmes, Y., McEnteggart, C., De Mey, H. R. A., Witteman, C. J. M., Janssen, G. T. L., & Egger, J. I. M. (2016). The relationship between theory of mind and relational frame theory: Convergence of perspective-taking measures. *Clinical Neuropsychiatry, 13(2*), 17-23.

Hoffman, S. G., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews,… Harris, P. L. (2016). Training children's theory-of-mind: A meta-analysis of controlled studies. *Cognition, 150,* 200-212. DOI: 10.1016/j.cognition.2016.01.006.

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development, 13(4),* 590-619. DOI: 10.1111/j.1467-9507.2004.00285.x

Hughes, S., & Barnes-Holmes, D. (2014). Associative concept learning, stimulus equivalence, and relational frame theory: Working out the similarities and differences between human and non-human behavior. *Journal of the Experimental Analysis of Behavior, 101,* 156–160. DOI:10.1002/jeab.60

Hussey, I., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). From relational frame theory to implicit attitudes and back again: Clarifying the link between RFT and IRAP research, *Current Opinion in Psychology, 2,* 11-15. DOI: 10.1016/j.copsyc.2014.12.009

Hussey, I., McEnteggart, C., Barnes-Holmes, Y., Kavanagh, D., Barnes-Holmes, D., Parling, T., & Lundgren, T. *Flexible perspective-taking: New concepts and a new behavioural measure.* ACT CBS Conference. Dublin (December, 2014).

Hyland, J. M., Smyth, S., O'Hora, D., & Leslie, J. C. (2014). The effect of before and after instructions on the speed of sequential responding. *The Psychological Record, 64 (2),* 311-319. DOI: 10.1007/s40732-014-0026-y

Idson, L. C., Chugh, D., Bereby-Meyer, Y., Moran, S., Grosskopf, B., & Bazerman, M. (2004). Overcoming focusing failures in competitive environments, *Journal of Behavioral Decision Making, 12*, 159-172. DOI: 10.1002/bdm.467

Im-Bolter, N., Agostino, A., & Owens-Jaffray, K. (2016). Theory of mind in middle childhood and early adolescence: Different from before? *Journal of Experimental Child Psychology, 149,* 98-115. DOI: 10.1016/j.jecp.2015.12.006

Izard, C. E., Fine, S., Schultz, D., Mostow, A., Ackerman, B. P., & Youngstrom, E. A. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science, 12(1),* 18-23. DOI: 10.1111/1467-9280.00304

Jackson, M. L., Mendoza, D. R., & Adams, A. N. (2014). Teaching a deictic relational repertoire to children with autism. *The Psychological Record, 64*, 791-802. DOI: 10.1007/s40732-014-0078-z

Janssen, G., De May, H., Hendriks, A., Koppers, A., Kaarsemaker, M., Witteman, C., & Egger, J. (2014). Assessing deictic relational responding in individuals with social anxiety disorder: Evidence of perspective-taking difficulties. *The Psychological Record, 64(1),* 21-29. DOI: 10.1007/s40732-014-0013-3

Kanter, J. W., Parker, C. R., & Kohlenberg, R. J. (2001). Finding the self: A behavioral measure and its clinical implications. *Psychotherapy, 38,* 198-211. DOI:10.1037/0033-3204.38.2.198

Kavanagh, D., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C., & Finn, M. (2018). Exploring differential trial type effects and the impact of a read-aloud

procedure on deictic relational responding on the IRAP. *The Psychological Record, 68,* 163-176. DOI:10.1007/s40732-018-0276-1

Kavanagh, D., Hussey, I., McEnteggart, C., Barnes-Holmes, Y., & Barnes-Holmes, D. (2016). Using the IRAP to explore natural language statements. *Journal of Contextual Behavioral Science*, *5*(4), 247–251. DOI: 10.1016/j.jcbs.2016.10.001

Kavanagh, D., Roelandt, A., Van Raemdonck, L., Barnes-Holmes, Y., Barnes-Holmes, D., & McEnteggart, C. (2019). The on-going search for perspective-taking IRAPs: Exploring the potential of the natural language IRAP. *The Psychological Record, 69(2),* 291-314. DOI: 10.1007/s40732-019-00333-w

Kent, G., Galvin, E., Barnes-Holmes, Y., Murphy, C., & Barnes-Holmes, D. (2017). Relational responding: Testing, training and sequencing effects among children with autism and typically developing children. *Behavioral Development Bulletin, 22(1),* 94-110. DOI: 10.1037/bdb0000041

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89,* 25-41. DOI: 10.1016/S0010-0277(03)00064-7

Keysar, B., & Barr, D. J. (2002). Self anchoring in conversation: Why language users do not do what they "should". In T. Gilovich, D. W. Griffin, & D. Kahneman (Eds), *Heuristics and biases: The psychology of intuitive judgement* (pp. 150-1660. New York: Cambridge University Press.

Konings, M., Bak, M., Hanssen, M., Van Os, J., & Krabbendam, L. (2006). Validity and reliability of the CAPE: A self-report instrument for the measurement of psychotic experiences in the general population. *Acta Psychiatrica Scandinavica, 114,* 55–61. DOI:10.1111/j.1600-0447.2005.00741.x

Kristen, S., Sodian, B., Thoermer, C., & Perst, H. (2011). Infants' joint attention skills predict toddlers' emerging mental state language. *Developmental Psychology, 47(5),* 1207-1219. DOI: 10.1037/a0024808

Lattal, K. A. (1975). Reinforcement contingencies as discriminative stimuli. *Journal of the Experimental Analysis of Behavior, 23,* 241-246. DOI: 10.1901/jeab.1975.23-241

LeBlanc, L. A., Coates, A. M., Daneshvar, S., Charlop-Christy, M. H., Morrus, C., & Lancaster, B. M. (2003). *Journal of Applied Behavior Analysis, 36,* 253-257. DOI: 10.1901/jaba.2003.36-253

Leech, A., Barnes-Holmes, D., & Madden, L., (2016). Using the Implicit Relational Assessment Procedure (IRAP) as a measure of spider fear, avoidance and approach. *The Psychological Record, 66*(3), 337-349. DOI: 10.1007/S40732-016-0176-1

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,* 201-239. DOI: 10.1006/jecp.1993.1032

Lovett, S., & Rehfeldt, R. A. (2014). An evaluation of multiple exemplar instruction to teach perspective-taking skills to adolescents with Asperger syndrome. *Behavioral Development Bulletin, 19(2),* 22-36. DOI: 10.1037/h0100575

Lowenkron, B. (1998). Some logical functions of joint control. *Journal of the Experimental Analysis of Behavior, 69(3),* 327-354. DOI: 10.1901/jeab.1998.69-327

Luciano, C., Gómez-Becerra, I., & Rodríguez Valverde, M. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *Journal of the Experimental Analysis of Behavior, 87(3),* 349-365. DOI: 10.1901/jeab.2007.08-06

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what object others can see when interpreting their action? *Cognition, 105(3),* 489-512. DOI: 10.1016/j.cognition.2006.10.007

Macdonald, R., Anderson, J., Dube, W. V., Geckler, A., Green, G., Holcomb, W., Mansfield, R., & Sanchez, J. (2006). Behavioral assessment of joint attention: A methodological report. *Research in Developmental Disabilities, 27,* 138-150. DOI: 10.1016/j.ridd.2004.09.006

McGreal, C., Hyland, J., O'Hora, D., & Hogan, M. (2016). Mutual entailment of temporal relations in younger and older adults: Reversing order judgements.

The Psychological Record, 66(3), 419-428. DOI: 10.1007/s40732-016-0182-3

McHugh, L., Barnes-Holmes, Y., & Barnes-Holmes, D. (2004). Perspective-taking as relational responding: A developmental profile. *The Psychological Record, 54,* 115-144. DOI: 10.1007/BF03395465

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D., & Stewart, I. (2006). Understanding false belief as generalized operant behavior. *The Psychological Record, 56(3),* 341-364. DOI: 10.1007/BF03395554

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D., Whelan, R., & Stewart, I. (2007). Knowing me, knowing you: deictic complexity in false-belief understanding. *The Psychological Record, 57,* 533-542. DOI: 10.1007/BF03395593

McHugh, L., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & Dymond, S. (2007). Deictic relational complexity and the development of deception. *The Psychological Record, 57(4),* 517-531. DOI: 10.1007/BF03395592

McKinnon, M. C., & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition, 102(2),* 179-218. DOI: 10.1016/j.cognition.2005.12.011

Mead, G. H. (1934). *Mind, self, and society*. Chicago: Chicago University Press.

Moll, H., & Meltzoff, A. N. (2011a). How does it look? Level 2 perspective-taking at 36 months of age. *Child Development, 82(2),* 661-673. DOI: 10.1111/j.146 7-8624.2010.01571.x

Moll, H., & Meltzoff, A. N. (2001b). Perspective-taking and its foundation in joint attention. In N. Eilan, H. Lerman, & J. Roessler (Eds.), *Perception, causation, and objectivity. Issues in philosophy and psychology* (pp. 286-304), Oxford, England: Oxford University Press.

Moll, H., & Tomasello, M. (2006). Level 1 perspective-taking at 24 months of age. *British Journal of Developmental Psychology, 24,* 600-613. DOI:10.1348/026151005X55370

Montag, C., Dziobek, I., Richter, I. S., Neuhaus, K., Lehmann, A., Rudolf, S., Heekeren, H. R.,…Jürgen, G. (2011). Different aspects of theory of mind in paranoid

schizophrenia: Evidence from video-based assessment. *Psychiatry Research, 186,* 203-209. DOI: 10.1016/j.psychres.2010.09.006

Montoya-Rodríguez, M. M., & Molina Cobos, F. J. (2018). Assessing perspective-taking in children through different formats of deictic framing protocol. In Hou, H. T, & Ryan, C. S., *Behavior Analysis,* IntechOpen. DOI: 10.5772/intechopne.74539

Montoya-Rodríguez, M. M., Molina, F. J., & McHugh, L. (2017). A review of relational frame theory research into deictic relational responding. *The Psychological Record, 67(4),* 569-579. DOI: 10.1007/s40732-016-0216-x

Mulhern, T., Stewart, I., & McElwee, J. (2017). Investigating relational framing of categorization in young children. *The Psychological Record, 67,* 519-536. DOI: 10.1007/s40732-017-0255-y

Mulhern, T., Stewart, I., & McElwee, J. (2018). Facilitating relational framing of classification in young children. *Journal of Contextual Behavioral Science, 8,* 55-68. DOI: 10.1016/j.jcbs.2018.04.001

Németh, N., Mátrai, P., Hegyi, P., Czéh, B., Czopf, L., Hussain, A., … Simon, M. (2018). Theory of mind disturbances in borderline personality disorder: A meta-analysis. *Psychiatry Research, 270,* 143-153. DOI: 10.1016/j.psychres.2018.08.049

Novak, G. (1996). *Developmental psychology: Dynamical systems and behavior analysis*. Reno, NV: Context Press.

Novak, G. (1998). A behavioral systems theory of development. *Mexican Journal of Behavior Analysis. 24(2),* 181-196.

Novak, G., & Pelaez, M. (2004). *Child and adolescent development: A behavioural systems approach*. London: Sage.

O'Connor, J., Rafferty, A., Barnes-Holmes, D., & Barnes-Holmes, Y. (2009). The role of verbal behavior, stimulus nameability, and familiarity on the equivalence performances of autistic and normally developing children. *The Psychological Record, 59,* 53-74. DOI: 10.1007/BF03395649

O'Hora, D., Barnes-Holmes, D., Roche, B., & Smeets, P. M. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record, 54,* 437-460. DOI: 10.1007/BF03395484

O'Hora, D., Peláez, M., Barnes-Holmes, D., Rae, G., Robinson, K., & Chaudhary, T. (2008). Temporal relations and intelligence: Correlating relational performance with performance on the Wais-III. *The Psychological Record, 58(4),* 569-584. DOI: 10.1007/BF03395638

O'Neill, J., & Weil, T. M. (2014). Training deictic relational responding in people diagnosed with schizophrenia. *The Psychological Record, 64(2),* 301-310. DOI: 10.1007/s40732-014-0005-3

O'Toole, C., & Barnes-Holmes, D., (2009). Three chronometric indices of relational responding as predictors of performances on a brief intelligence test: The importance of relational flexibility, *The Psychological Record, 59(1),* 119-132.DOI: 10.1007/BF03395652

Murphy, C., O'Connor, J., & Barnes-Holmes, Y. (2009). Relational flexibility and human intelligence: Extending the remit of Skinner's Verbal Behavior. *International Journal of Psychology & Psychological Therapy, 9*(1), 1-17.

Okuda, K., & Inoue, M. (2000). A behavior analytic view of teaching "theory of mind" to children with autism: Stimulus control and generalization on false belief tasks. *Japanese Psychological Review, 43,* 427-442.

Peláez, M. (2009). Joint attention and social referencing in infancy as precursors of derived relational responding. In R. A. Rehfeldt, & Y. Barnes-Holmes (Eds.). *Derived relational responding applications for learners with autism and other developmental disabilities: A progressive guide to change* (pp. 63-78). Oakland, CA: New Harbinger Publications.

Peláez, M., Virues-Ortega, J., & Gewritz, J. L. (2012). Acquisition of social referencing via discrimination training in infants. *Journal of Applied Behavior Analysis, 45,* 23-36. DOI: 10.1901/jaba.2012.45-23

Peláez-Nogueras, M., & Gewirtz, J. (1997). The context of stimulus control in behavior analysis. In D. M. Baer & E. M. Pinkston (Eds)., *Environment and behavior*. Boulder, CO: Westview Press.

Perner, J., Frith, U., Leslie, A. M., & Leekman, S. R. (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development, 60*, 689-700. DOI: 10.2307/1130734

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory of mind development for children with deafness or autism. *Child Development, 76(2),* 502-517. DOI: 10.1111/j.1467-8624.2005.00859.x

Preißler, S., Dziobek,I., Ritter, K., Heekeren, H. R., & Roepke, S. (2010). Social cognition in borderline personality disorder: Evidence for disturbed recognition of the emotions, thoughts, and intentions of others. *Frontiers in Behavioral Neuroscience. 4*, 182. DOI: 10.3389/fnbeh.2010.00182

Piaget, J. (1948). *The moral judgment of the child.* New York: Free Press.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge.

Pickup, G. J. (2008). Relationship between theory of mind and executive function in schizophrenia: A systematic review. *Psychopathology, 41*, 206-213. DOI: 10.1159/000125554

Pickup, G. J., & Frith, C. D. (2001). Theory of mind impairments in schizophrenia: Symptomatology, severity and specificity. *Psychological Medicine, 31*, 207-220. DOI: 10.1017/S0033291701003385

Pierce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8-13. DOI: 10.1016/j.jneumeth.2006.11.017

Premack, D., & Woodruff, G., (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences, 1(4),* 515-526. DOI: 10.1017/S0140525X00076512

Rehfeldt, R. A., Dillen, J. E., Ziomek, M. M., & Kowalchuk, R. K. (2007). Assessing relational learning deficits in perspective-taking in children with high-functional autism spectrum disorder. *The Psychological Record, 57(1),* 23-47. DOI: 10.1007/BF03395563

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale. *Personality and Social Psychology Bulletin*, *27*, 151-161. DOI: 10.1177/0146167201272002

Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and sexual categorization: A critical test of the derived difference relation. *The Psychological Record, 46(3),* 451-475. DOI: 10.1007/BF03395177

Roepke, S., Vater, A., Preißler, S., Heekeren, H. R., & Dziobek, I. (2013). Social cognition in borderline personality disorder. *Frontiers in Neuroscience, 6*, 195. DOI: 10.3389/fnins.2012.00195

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in domains of social cognition in schizophrenia: A meta-analysis of the empirical evidence. *Schizophrenia Bulletin, 39(5),* 979-992. DOI: 10.1093/schbul/sbs080

Scherzer, P., Leveille. E., Achim, A., Boisseau, E., & Stip, E. (2012). A study of theory of mind in paranoid schizophrenia. *Frontiers in Psychology, 3(423),* 1-11. DOI: 10.3389/fpsyg.2012.00432

Scheeren, A. M., de Rosnay, M., Koot, H. M., & Begeer, S.(2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *The Journal of Child Psychology and Psychiatry, 54(6),* 628-635. DOI:10.1111/jcpp.12007

Schick, B., de Villers, P., de Villers, J., & Hoffmeister, R. (2007). Language and theory of mind: A study of deaf children. *Child Development, 78(2),* 376-396. DOI: 10.1111/j.1467-8624.2007.01004.x

Schlinger, H. D. (2009). Theory of mind: An overview and behavioral perspective. *The Psychological Record, 59*, 435-448. DOI: 10.1007/BF03395673

Shamay-Tsoory, S. G., Shur, S., Barcai-Goodman, L., Medlovich, S., Harari, H., & Levkovitz, Y. (2007). Dissociation of cognitive from affective components of theory of mind in schizophrenia, *Psychiatry Research, 149*, 11-23. DOI: 10.1016/j.psychres.2005.10.018

Sharp, C., Pane, H., Ha, C., Venta, A., Patel, A. B., Sturek, J., & Fonagy, P. (2011). Theory of mind and emotion regulation difficulties in adolescents with borderline traits. *Journal of the American Academy of Child & Adolescent Psychiatry, 50(6),* 563-573. DOI: 10.1016/j.jaac.2011.01.017

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research, 14(1),* 5-13. DOI: 10.1044/jshr.1401.05

Sigman, M., & Capps, L. (1997). *Children with autism: A developmental perspective.* Cambridge: Harvard University Press.

Skinner, B. F. (1974). *About behaviorism.* New York: Random House.

Slattery, B., & Stewart, I. (2014). Hierarchical classification as relational framing. *Journal of the Experimental Analysis of Behavior, 101(1),* 61-75. DOI: 10.1002/jeab.63

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *The Journal of Genetic Psychology, 164(1),* 54-71. DOI: 10.1080/00221320309597503

Sodian, B., & Kristen-Antonow, S. (2015). Declarative joint attention as a foundation of theory of mind. *Developmental Psychology, 51(9),* 1190-1200. DOI: 10.1037/dev0000039

Sodian, B., Thoermer, C., & Metz, U. (2007). Now I see it but you don't: 14-month-olds can represent another person's visual perspective. *Developmental Science*. *10(2),* 199–204. DOI: 10.1111/j.1467-7687.2007.00580.x

Spradlin, J. E., & Brady, N. (2008). A behaviour analytic interpretation of theory of mind. *International Journal of Psychology and Psychological Therapy, 8(3),* 335-350.

Stefanis N. C., Hanssen M., Smirnis N. K., Avramopoulos D. A., Evdokimidis I. K., Stefanis C. N.,…& Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine, 32,* 347–358. https://doi.org/10.1017/S0033291701005141

Steele, D., & Hayes, S. C. (1991). Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior, 56(3),* 519-555. DOI: 10.1901/jeab.1991.56-519

Sullivan, K., Zaitchik, D., & Tager-Flusber, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*, 395-402.

Tager-Flusberg, H. (2007). Evaluating the Theory-of-mind hypothesis of autism. *Current Directions in Psychological Science. 16(6),* 311-315. DOI: 10.1111/j.1467-8721.2007.00527.x

Tager-Flusberg, H., & Anderson, M. (1991). The development of contingent discourse ability in autistic children. *Journal of Child Psychology and Psychiatry, 32(7),* 1123-1134. DOI: 1469-7610.1991.tb00353.x

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition, 110,* 124-129. DOI: 10.1016/j.cognition.2008.10.008

Uekermann, J., Kramer, M., Abdel-Hamid, M., Schimmelmann, B. G., Hebebrand, J., Daum, I.,…Kis, B. (2010). Social cognition in attention-deficit hyperactivity disorder (ADHD). *Neuroscience and Biobehavioral Reviews, 34(5),* 734-743. DOI: 10.1016/j.neubiorev.2009.10.009.

Vahey, N., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). A first test of the Implicit Relational Assessment Procedure (IRAP) as a measure of self-esteem: Irish prisoner groups and university students. *The Psychological Record, 59,* 371-388. https://doi.org/10.1007/BF03395670

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. DOI: 10.1016/j.jbtep.2015.01.004

Vaskin, A., Antonsen, B. T., Fretland, R. A., Dziobek, I., Sundet, K., & Wilberg, T. (2015). Theory of mind in women with borderline personality disorder or schizophrenia: differences in overall ability and error patterns. *Frontiers in Psychology, 6,* 1239. DOI: 10.3389/fpsyg.2015.01239

Vescio, T. K., Sechrist, G. B., & Paolucci, M. P. (2003). Perspective taking and prejudice reduction: The mediational role of empathy arousal and situational attributions. *European Journal of Social Psychology, 33(4)*, 455–472. DOI: 10.1002/ejsp.163

Vilardaga, R., Estévez, A., Levin, M. E., & Hayes, S. (2012). Deictic relational responding, empathy, and experiential avoidance as predictors of social anhedonia: Further contributions from relational frame theory. *The Psychological Record, 62(3),* 409-432. DOI: 10.1007/BF03395811

Villatte, M., Monestés, J.-L., McHugh, L., Freixa i Baqué, E. F., & Loas, G. (2008). Assessing deictic relational responding in social anhedonia: A functional approach to the development of theory of mind impairments. *Journal of Behavioral Consultation and Therapy*, *4*, 360–373. DOI: 10.1037/h0100867

Villatte, M., Monestés, J.-L., McHugh, L., Freixa i Baqué, E., & Loas, G. (2010a). Assessing perspective taking in schizophrenia using relational frame theory. *The Psychological Record, 60(3)*, 413-436. DOI: 10.1007/BF03395719

Villatte, M., Monestés, J.-L., McHugh, L., Freixa i Baqué, E., & Loas, G. (2010b). Adopting the perspective of another in belief attribution: Contribution of relational fame theory to the understanding of impairments in schizophrenia. *Journal of Behavior Therapy and Experimental Psychiatry, 41*, 125-134. DOI: 10.1016/j.jbtep.2009.11.004

Vitale, A., Barnes-Holmes, Y., Barnes-Holmes, D., & Campbell, C. (2008). Facilitating responding in accordance with the relational frame of comparison; systematic empirical analyses. *The Psychological Record, 58(3),* 365-390. DOI: 10.1007/BF03395624

Weil, T. M., Hayes, S. C., & Capurro, P. (2011). Establishing a deictic relational repertoire in young children. *The Psychological Record, 61(3),* 371-390. DOI: 10.1007/BF03395767

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72(3),* 655–684. DOI: 0009-3920/2001/7203-0001

Wellman, H. M., Phillips, A. T., & Rodriguez, T. (2000). Young children's understanding of perception, desire, and emotion. *Child Development, 71(4),* 898-912. DOI: 0009-3920/2000/7104-0011

Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behaviour modification procedures. *Journal of Child Psychology and Psychiatry, 44(3),* 456-468. DOI: 10.1111/1469-7610.00135

Whelan, R., & Barnes-Holmes, D. (2004). The transformation of consequential functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior, 82(2),* 177-195. DOI: 10.1901/jeab.2004.82-177

World Health Organisation (2017). Process of translation and adaptation of instruments. Retrieved from http://www.who.int/substance_abuse/research_tools/translation/en/

English Summary

Introduction

Perspective-taking has been broadly defined as the ability to interpret and predict the thoughts, emotions or behaviours of oneself and of others (Carpendale & Lewis, 2006) in terms of being able to 'assume an alternative perspective' where necessary. In mainstream psychology, the broad concept of perspective-taking has been sub-divided into visual perspective-taking, emotional perspective-taking and cognitive perspective-taking (usually referred to as ToM). Although these domains are often studied independently, all three appear to be involved, to varying degrees, in the ability to take the perspective of another. Visual perspective-taking has been observed as early as 12 months. Affective perspective-taking emerges later, usually around age two, and cognitive perspective-taking (ToM) comes even later around age four, with false belief understanding observed around age six. However, there have been wide variations in the ages at which these skills are first observed, with outcomes apparently sensitive to both methodological variations and broader individual differences, especially in executive functioning. There is evidence of weaker perspective-taking performances in samples with a diagnosis of ASD, BPD or schizophrenia, but these too appear to be influenced by task-specific features, and broader executive functioning abilities.

In traditional behaviour analysis, the ability to respond to one's own responding (i.e. acquire self-awareness) as an essential pre-cursor to perspective-taking, is shaped through a history of interacting with other members of the social/verbal community, although empirical support for this interpretation of self-awareness has tended to come from research with non-humans. Indeed, little behaviour-analytic research has explicitly attempted to connect the concept of self with perspective-taking skills in humans. While some researchers have offered behavioural interpretations of ToM performances in terms of equivalence relations and operant stimulus control, most research has focused on remediating deficits in perspective-taking through the provision of appropriate

learning histories. The outcomes show that video modeling interventions have enabled children with ASD to pass the Sally-Anne Test, although generalisation is more robust when training exemplars are included. Incorporating training prompts into the Sally-Anne Test has also produced positive outcomes, but appear to be necessary or less effective for participants with language deficits. Developmental behavioural researchers have devoted considerable attention to joint attention and social referencing as pre-cursors to perspective-taking. There is evidence that joint attention responding and initiation can be established in children with ASD, although the latter (initiation) is more likely to be deficient and more difficult to establish, and generalisation is not always observed. In general, as one might expect, there has been a far greater emphasis in behavioural psychology on prediction-and-influence of perspective-taking, which in a broad sense further complicates the literature in this area. That is, research from the mainstream literature has tended to produce relatively inconsistent results in attempting to tie deficits in perspective-taking to specific ages, populations and syndromes. The fact that behaviour-analytic research also shows that perspective-taking deficits are relatively amenable to change when targeted by behavioural interventions makes it even more difficult to draw firm conclusions about the very concept of perspective-taking itself.

Behavioural researchers working under the rubric of Relational Frame Theory (RFT) have proposed a distinction between verbal self-discrimination as observed in humans and non-verbal self-discrimination observed with non-humans. According to RFT, verbal self-discrimination and perspective-taking comprise repertoires of arbitrarily applicable relational responding (AARR). For RFT, perspective-taking is AARR that becomes abstracted across multiple exemplars of talking about your perspective in relation to others. The core relations involved in perspective-taking are referred to as deictic relations, and include responding from one's own perspective in terms of interpersonal relations, spatial relations and temporal relations. Most of the empirical research on deictic relational responding has employed various iterations of a developmental protocol. This research supports the distinctions among the three types of deictic relations, and the finding that these relations vary on a continuum of

complexity from simple relations to reversed relations and double reversed relations. In general, accuracies in performances on the deictic relational protocol increase as a function of age. However, even adults produce considerable error rates and some adult performances are better on double reversals than reversals. Once established via multiple-exemplar training (MET), perspective-taking repertoires can generalise to both new stimuli and contexts. Studies assessing patterns of deictic relational responding in atypical populations have found that children with ASD produce somewhat weaker patterns than their peers. Results with atypical adults have found that performances on the deictic protocol correlate with verbal ability and IQ. Participants with a diagnosis of schizophrenia have been shown to produce significantly weaker results on reversed and double reversed relations than typically-developing counterparts. Several researchers have raised concerns about using the protocol to assess deictic relational repertoires in adult samples.

A considerable number of studies have assessed the relationship between performances on the deictic protocol and traditional ToM tasks, and found evidence of a functional overlap between deictic relational responding on the protocol and ToM with numerous samples. Several studies have systematically adapted trials from the deictic protocol to create tasks that closely resemble the attribution of true and false beliefs as targeted by traditional ToM tasks. The evidence indicates that accuracy on both types of trials increases as a function of age, but there are no differences in accuracy between true and false belief attribution. Related research has compared responding to self versus other in adapted false belief tasks and found no difference in overall accuracies, but has observed significantly longer latencies on trials involving the perspective of other versus self. In studies with atypical samples, those with high social anhedonia have been found to perform significantly more poorly than controls on false belief attributions to others. Similarly, participants with a diagnosis of schizophrenia have performed significantly worse than controls on self-attributions of false belief.

Much of the work conducted in the area of deictic relational responding has focused on assessing the presence of these patterns of relational responding in a dichotomous manner. That is, participants were typically assessed for the presence of

deictic relational responding, and if found to be deficient, these relations were trained and tested to determine if they produced predicted outcomes. However, little research has focused on the relative strength of a pattern once it was observed and there have been recent calls for analyses that focus on the relative strength or persistence of derived relational responding, rather than simply its presence versus absence (e.g. Barnes-Holmes et al., 2016). In an attempt to develop methodologies for assessing the relative strength of derived relational responding, researchers have explored alternative methodologies, such as the Implicit Relational Assessment Procedure (IRAP; e.g. Barnes-Holmes, et al., 2008; Hussey, et al., 2015).

The IRAP requires participants to emit two opposing patterns of relational responding and the ease with which one pattern may be emitted relative to the other provides a measure of response strength. Specifically, the procedure typically presents label and target stimuli (e.g. the label word "puppy" with the target word "pleasant") and requires participants to respond "True" (e.g. puppy-pleasant) or "False" (e.g. puppy-unpleasant) to the stimulus pairs. An IRAP typically comprises four trial types (e.g. *Puppy-Positive*, *Puppy-Negative*, *Spider-Positive* and *Spider-Negative*) that are generally analysed independently in terms of the difference in response latencies between responding that is deemed consistent versus inconsistent with a participant's verbal history. In general, response latencies are expected to be shorter during blocks of trials that require history-consistent versus history-inconsistent responding.

The body of empirical research employing the IRAP has grown considerably, with an increasing focus on clinically relevant phenomena (Vahey et al., 2015). Using the IRAP to assess deictic relational responding, particularly in the clinical domain, would provide an alternative to the Barnes-Holmes (2001) protocol. A recent study in which the IRAP was used to target responding to self versus others is particularly relevant (Barbero-Rubio et al., 2016).

This original study presented participants with their own names and the name of the researcher as label stimuli, and statements pertaining to specific characteristics of the self versus other as targets (e.g. "is in front of the laptop"). There were two response

options (“Yes” and “No”) on each trial. The four trial types in this study were referred to as: *I-I* (participant name-participant characteristics), *Other-Other* (researcher name-researcher characteristics), *I-Other* (participant name-researcher characteristics) and *Other-I* (researcher name-participant characteristics). In general, the pattern of IRAP effects reported by Barbero-Rubio et al. (2016) indicated that participants’ response latencies showed significantly more rapid responding on the *I-I* trial type relative to the other three trial types during history-consistent blocks (i.e. responding “True” on *I-I* and *Other-Other* trial types, and responding “False” on *I-Other* and *Other-I* trial types). In addition, the difference in response latencies between consistent and inconsistent blocks for each trial type was in the predicted direction (i.e. shorter on history-consistent relative to history-inconsistent trials), and these differences were significant in terms of the normalised D_{IRAP}-scores.

Overview of the Empirical Chapters

Chapter 2. Following on a review of the literature (summarised above), Chapter 2 attempted to systematically replicate previous research by Barbero-Rubio et al. (2016). This involved returning to a traditional IRAP format that presented label stimuli at the top of the screen and target stimuli in the centre. Specifically, participants were presented with an IRAP that contained each participant’s own name (self) versus the name of the researcher (other) as label stimuli, and statements describing specific characteristics of the self (e.g. “is in front of the laptop”) versus other (e.g. “is standing up”) as targets, along with “Yes” and “No” as response options. In an extension of the research by Barbero-Rubio et al., the current work included a Control Condition in which no responding to self was involved, only to others. That is, the Control Condition involved an IRAP in which none of the trial types required responding to self, but only responding to two other separate individuals (i.e. the researcher and a picture of another unknown participant). If the comparison between self and other in a deictic IRAP is an important variable, one might expect a different pattern of results in a Control Condition

in which there is no contrast between self and other. One final way in which this research extended the Barbero-Rubio et al. study was the inclusion of self-report measures of self-esteem and the presence of psychotic-like experiences. We also retained one measure of perspective-taking employed in the original study. These were included on an exploratory basis. The results from Experiment 1 yielded significant IRAP effects for two of the trial types in both the deictic and control IRAPs, whereas four significant effects were observed in Barbero-Rubio et al. in which there was no Control Condition. A correlation matrix was calculated to determine if any of the D_{IRAP}-scores from the deictic IRAP predicted self-reported psychotic experiences (on the Community Assessment of Psychic Experiences; CAPE), self-esteem (on the Rosenberg Self-esteem Scael; RSES) or perspective-taking (on the Perspective Taking sub-scale of the Interpersonal Reactitivity Index; IRI). The only significant correlations involved the *Other-Other* trial type with the overall CAPE frequency, and the CAPE positive subscale, but neither remained significant after Bonferroni corrections.

A second experiment involved a novel method for collecting IRAP data, which had been shown to yield significant effects for all four trial types in a separate line of research being conducted by our group (Finn et al., 2018). The *read-aloud procedure* required participants to read aloud the label, target and chosen response option at the time of selection on each IRAP trial throughout the entire procedure. Given the differences between the results of Experiment 1 and those found by Barbero-Rubio et al. (2016), we investigated whether introduction of the read-aloud procedure would produce a pattern of effects similar to the original study. The data from Experiment 2 once again showed relatively strong IRAP effects on the *I-I* and *Other-Other* trial types, replicating our findings from Experiment 1 and those from Barbero-Rubio et al., although we no longer found significant differences between these two trial types. On balance, Experiment 2 now replicated the effects on *I-Other* and *Other-I* reported in the original study (i.e. they were both significantly different from zero). The additional analyses we conducted, however, indicated that they were both significantly weaker than the *I-I* and *Other-Other* trial types. The control IRAP also yielded significant effects for all four trial types, although two of the trial types (*Researcher-Researcher*

and *Other-Other*) continued to be significantly stronger than the two remaining trial types. Finally, in contrast to Experiment 1, all of the significant correlations between the deictic IRAP and the questionnaires were recorded for the *I-I* trial type (rather than *Other-Other*). Again, however, these did not remain significant after Bonferroni corrections.

Chapter 3. This chapter attempted to develop IRAPs that taregeted deictic relational responding with regard to the mental states of self and other**.** Developing such an IRAP would necessarily involve inserting complex relational networks into the procedure. This was pursed in a series of experiments that employed a novel version of the IRAP, known as the Natural Language-IRAP (NL-IRAP). The use of the NL-IRAP allowed for the presentation of relatively complex statements that required participants to infer the thoughts or beliefs of others on a trial-by-trial basis within the IRAP. Across a sequence of six experiments, a 'self-focused IRAP' required participants to respond to both positive and negative statements about themselves, whilst an 'other-focused IRAP' required participants to respond to similar statements about others. Experiments 1 and 2 investigated perspective-taking with regard to an *unspecified* other. In addition, correlational analyses were condutced to determine the extent to which the two NL-IRAPs correlated, or failed to correlate, with each other and with a range of self-report measures that were deemed to be relevant to perspective-taking. Experiments 3-6 investigated perspective-taking with regard to a *specified* other, with the specified relationship between self and other manipulated across experiments. The results from the first two experiments indicated that there were significant differences between the self- versus other-focused IRAPs, when the other remained unspecified. The remaining four experiments, however, indicated that when the other was specified there was limited evidence that performances on the two IRAPs differed significantly. The correlational analyses between the IRAPs and the self-report measures yielded very few significant effects (and none at all after Bonferroni corrections). Finally, the correlational analyses between the two IRAPs in each experiment were significant in some studies(Experiments 2, 3 and 5), but not others (Experiments 1, 4 and 6). Overall,

the findings from Chapter 3 could be seen as encouraging because each experiment produced performances that would be deemed consistent with the pre-experimental histories of the participants. That is, the IRAP effects for the most part, were in the predicted direction (e.g. participants confirmed more rapidly than they denied that when positive events occur, they react positively, but when negative events occur, they react negatively).

The results were also encouraging because we observed a possibly important distinction in perspective-taking when it applied to a general unspecified other versus someone well known to the participant (see DeBernardis et al., 2014). On balance, the results could be seen as somewhat disappointing because we failed to find strong evidence of perspective-taking when other was specified, at least in terms of different performances across the two IRAPs, or in correlations among the IRAPs and the self-report measures. Perhaps, therefore, there was some sensitivity to self versus other, but the use of complex statements in the NL-IRAPs potentially undermined or reduced the impact of deictic relational responding per se. That is, in presenting such complex stimuli or reational networks in the IRAP, participants more or less interpreted the task as a sense-making or problem-solving task, in which the self versus other had little or no impact, particularly when the other was specified.

Chapter 4. In reflecting upon the results obtained in the two previous empirical chapters, Chapter 4 attempted to develop IRAPs that maintained the deictic functions of self and other, in the context of perspective-taking rather than simple sense-making. The primary objective of Experiment 1 in this chapter was to determine if it was feasible to use matched pictures of self versus other across two IRAPs. In both IRAPs, the contrast categories were pen-related stimuli. In general, the pattern of effects indicated that pictures of faces could be incorporated into the IRAP, thus supporting the feasibility of using photographs of self versus other in IRAP perspective-taking studies. However, two effects did emerge that were worthy of further consideration, but these were thereafter addressed in the final Discussion chapter of the thesis because the effects were also observed in studies presented in other chapters.

The primary objective of Experiment 2 in Chapter 4 was to determine if the presentation of a false belief vignette before exposure to a single IRAP would influence performance in a vignette-consistent direction. The results of this study were inconclusive for a number of reasons. First, no clear differences emerged between the active vignette and the Control Condition (and the latter could have been interpreted as being ambiguous for participants). Second, the use of a familiarisation IRAP, instead of requiring participants to complete practice blocks in the belief IRAP, likely contributed to high rates of attrition, thus restricting data analyses to the final pair of test blocks. Third, the order in which IRAP blocks were presented (i.e. vignette-consistent-first versus vignette-inconsistent-first) was not counterbalanced and thus it was difficult to determine if the vignette and/or the block sequence was the primary controlling variable in determining the observed IRAP effects.

The primary aim of Experiment 3 in Chapter 4 was to rectify the three main problems described above. The experiment thus involved an active control design in which the two vignettes specified the opposite states of affairs; the familiarisation IRAP was removed from the experiment, but the belief IRAP included practice blocks; and the block sequence of the IRAP was counterbalanced. The results of this study were more conclusive than those of the second, in indicating that the primary controlling variable was the sequence in which the IRAP blocks were presented, rather than the content of the vignettes. On balance, post-hoc analyses of performance of the first practice block pairs indicated that the block sequence and vignettes did interact with each other, but only when the vignette and the initial contingencies of the IRAP cohered with each other.

Chapter 5. At this stage in the doctoral research programme, 11 individual experiments had been completed and there was limited evidence that an IRAP had been developed that showed clear and unambiguous sensitivity to perspective-taking. On balance, the research contributed considerably to the on-going development of a conceptual analysis of perspective-taking and the dynamics of AARR, more generally. And thus, the current findings are highly valuable in that regard. The General Discussion

presented a discussion of these conceptual developments and their relationship to the current empirical work.

Nederlandstalinge Samenvattig

Empirische en conceptuele analyse van perspectief-nemen:

A gedrags-analytische benadering

In een poging om de methodologieën voor het bestuderen van deictische relationele respons uit breiden, gebruikt men in recent onderzoek de Implicit Relational Assessment Procedure (IRAP) om deictische relationele respons met betrekking tot zelf ten opzichte van OTHER te meten (Barbero-Rubio et al., 2016). Het primaire doel van het huidige proefschrift was om verder te bouwen op de kleine hoeveelheid werk die hieromtrent bestaat en om het potentiële nut van de IRAP in verband met perspectief-nemen te bepalen. Met andere woorden, onderzoeken of de IRAP perspectief-nemen 'tijdens de vlucht' kan vastleggen? Het oorspronkelijke doel van hoofdstuk 2 was om het oorspronkelijke onderzoek van Barbero-Rubio et al. gedeeltelijk te repliceren en uit te breiden. Deze uitbreiding omvatte het opnemen van een controle-conditie waarbij het niet nodig was te reageren op zichzelf, maar alleen op anderen. De resultaten leverden significante IRAP-effecten op voor twee van de vier proeftypen in zowel de deictische als controle-IRAP's. Experiment 2 betrof een nieuwe methode voor het verzamelen van IRAP-gegevens (een lees-hardop procedure), waarvan werd aangetoond dat het significante effecten opleverde voor alle vier de proeftypen, en vier significante effecten werden inderdaad geregistreerd voor zowel deictic als controle IRAP's.

Hoofdstuk 3 trachtte IRAP's te ontwikkelen die deictische relationele responsen met betrekking tot de mentale toestanden van zichzelf en van anderen konden meten. Dit werd nagestreefd in een reeks experimenten met een nieuwe versie van de IRAP, bekend als de Natural Language-IRAP (NL-IRAP). In een reeks van zes experimenten vereiste een ‘zelfgerichte IRAP’ dat deelnemers moesten reageren op zowel positieve als negatieve uitspraken over zichzelf, terwijl een ‘andere gerichte IRAP’ deelnemers verplichtte te reageren op soortgelijke uitspraken over anderen. Experimenten 1 en 2 onderzochten perspectief-nemen met betrekking tot een niet-gespecificeerde andere. Experimenten 3-6 onderzochten perspectief nemen met betrekking tot een

gespecifieerde andere, waarbij de gespecificeerde relatie tussen zichzelf en de andere gemanipuleerd werd in experimenten. De resultaten van de eerste twee experimenten gaven aan dat er significante verschillen waren tussen de zelf- versus andere-gerichte IRAP's, wanneer de andere niet gespecificeerd bleef. De resterende vier experimenten gaven echter aan dat wanneer de andere gespecifieerd was, er slechts beperkte aanwijzingen waren dat de prestaties op de twee IRAP's aanzienlijk verschilden. Over het algemeen kunnen de bevindingen als bemoedigend worden beschouwd omdat elk experiment performances opleverde die als consistent werden beschouwd met de pre-experimentele geschiedenis van de deelnemers. Over het algemeen waren de resultaten evenwel enigszins teleurstellend omdat we geen sterk bewijs van perspectief-nemen konden vinden wanneer andere werd gespecificeerd, althans in termen van verschillende prestaties tussen de twee IRAP's, of in correlaties tussen de IRAP's en de zelfrapportagemaatregelen. Een mogelijke verklaring is dat er enige gevoeligheid voor zichzelf versus anderen was, maar dat het gebruik van complexe uitspraken in de IRAP's de impact van deictic relationele respons op zich ondermijnde of verminderde.

Na het reflecteren over de resultaten verkregen in de twee voorgaande empirische hoofdstukken, probeerde we in hoofdstuk 4 IRAP's te ontwikkelen die de deictische functies van zichzelf en anderen in stand hielden in de context van perspectief nemen, in plaats van eenvoudig zin-makend te zijn. Experiment 1 beoordeelde de haalbaarheid van het gebruik van gelijkaardige foto's van zichzelf versus andere over twee IRAP's. Over het algemeen gaf het patroon van effecten aan dat afbeeldingen van gezichten in de IRAP konden worden opgenomen, waardoor de haalbaarheid van het gebruik van foto's van zichzelf versus andere in perspectiefstudies met IRAP werd ondersteund. Experiment 2 probeerde te bepalen of de presentatie van een false-belief vignet vóór blootstelling aan een enkele IRAP de prestaties in een vignet-consistente richting zou beïnvloeden. De resultaten waren echter om een aantal redenen niet doorslaggevend. Ten eerste zijn er geen duidelijke verschillen naar voren gekomen tussen het actieve vignet en de controle-conditie (en de laatste kon dubbelzinnig geïnterpreteerd worden door de deelnemer). Het primaire doel van experiment 3 was om enkele problemen die in eerdere onderzoeken naar voren kwamen, weg te nemen. Het experiment omvatte dus

een actief besturingsontwerp waarin de twee vignetten de tegenovergestelde stand van zaken specificeerden; de familiarisatie IRAP werd verwijderd uit het experiment, maar oefenblokken werden toegevoegd aan de overtuiging IRAP; en de bloksequentie van de IRAP werd gecounterbalanceerd. De resultaten van het derde experiment waren meer overtuigend, omdat ze aangaven dat de primaire controlevariabele de volgorde was waarin de IRAP-blokken werden gepresenteerd in plaats van de inhoud van de vignetten. Over het algemeen gaven post-hoc analyses van de prestaties van de eerste oefenblokparen aan dat de bloksequentie en vignetten met elkaar in wisselwerking stonden, maar alleen wanneer het vignet en de initiële contingenties van de IRAP met elkaar in overeenstemming waren.

Over de 11 individuele experimenten die in dit proefschrift zijn voltooid, was er beperkt bewijs dat er succesvol een IRAP was ontwikkeld die een duidelijke en ondubbelzinnige gevoeligheid voor perspectief-nemen vertoonde. Over het algemeen heeft het onderzoek aanzienlijk bijgedragen aan de voortdurende ontwikkeling van een conceptuele analyse van perspectief-nemen en de dynamiek van willekeurig toepasbare relationele antwoorden, in het algemeen. En dus zijn de huidige bevindingen in dat opzicht zeer waardevol. De algemene discussie presenteerde een bespreking van deze conceptuele ontwikkelingen en hun relatie tot het huidige empirische werk.

Data Storage Fact Sheets

Data Storage Fact Sheet for Chapter 2

Name/identifier study: The On-going search for perspective-taking IRAPs: Exploring the Potential of the Natural Language IRAP

Author: Deirdre Kavanagh

Date: 11th February 2019

1. Contact details

1a. Main researcher

Name: Deirdre Kavanagh

Address: Henri Dunantlaan 2, B-9000 Gent

E-mail: deirdre.kavanagh@ugent.be

1b. Responsible Staff Member (ZAP)

Name: Prof. dr Dermot Barnes-Holmes

Address: Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

E-mail: Dermot.Barnes-Holmes@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

Reference of the publication in which the datasets are reported:

Kavanagh, D., Roelandt, A., Van Raemdonck, L., Barnes-Holmes, Y., Barnes-Holmes, D., McEnteggart, C. (2019). The On-going search for perspective-taking IRAPs: Exploring the potential of the natural language IRAP.The Psychological Record.

Which datasets in that publication does this sheet apply to?:

It refers to all studies that are reported in Chapter 2

3. Information about the files that have been stored

3a. Raw data

Have the raw data been stored by the main researcher? [**X**] YES / [ ] NO

If NO, please justify: N/A

On which platform are the raw data stored?

- [**X**] researcher PC
- [**X**] research group file server
- [ ] other (specify): ...

Who has direct access to the raw data (i.e., without intervention of another person)?

- [**X**] main researcher
- [**X**] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

3b. Other files

Which other files have been stored?

- [**X**] file(s) describing the transition from raw data to reported results. Specify: ...
- [**X**] file(s) containing processed data. Specify: ...
- [**X**] file(s) containing analyses. Specify: ...
- [ ] files(s) containing information about informed consent.
- [ ] a file specifying legal and ethical provisions
- [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
- [ ] other files. Specify: ...

On which platform are these other files stored?

- [**X**] individual PC
- [**X**] research group file server
- [ ] other: ...

Who has direct access to these other files (i.e., without intervention of another person)?

- [**X**] main researcher
- [**X**] responsible ZAP
- [ ] all members of the research group

- [ ] all members of UGent
- [ ] other (specify): ...

4. Reproduction

Have the results been reproduced independently?: [ ] YES / [**X**] NO

If yes, by whom (add if multiple):

Name: Address:

Affiliation: E-mail:

Data Storage Fact Sheet for Chapter 3

Name/identifier study: Exploring differntial trial-type effect and the imapct of a read-aloud proceudre on deictic relational responidng on the IRAP.

Author: Deirdre Kavanagh

Date: 24th April 2019

1. Contact details

1a. Main researcher

Name: Deirdre Kavanagh

Address: Henri Dunantlaan 2, B-9000 Gent

E-mail: deirdre.kavanagh@ugent.be

1b. Responsible Staff Member (ZAP)

Name: Prof. dr Dermot Barnes-Holmes

Address: Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

E-mail: Dermot.Barnes-Holmes@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

Reference of the publication in which the datasets are reported:

Kavanagh, D., Barnes-Holmes, Y., Barnes-Holmes, P. M. D., McEnteggart, C., & Finn, M. (2018). Exploring differential trial-type effects and the impact of a read-aloud procedure On deictic relational responding on the IRAP. *The Psychological Record, 68(2),* 163–176.

Which datasets in that publication does this sheet apply to?:

It refers to all studies that are reported in Chapter 3

3. Information about the files that have been stored

3a. Raw data

Have the raw data been stored by the main researcher? [**X**] YES / [ ] NO

If NO, please justify: N/A

On which platform are the raw data stored?

- [**X**] researcher PC
- [**X**] research group file server
- [ ] other (specify): ...

Who has direct access to the raw data (i.e., without intervention of another person)?

- [**X**] main researcher
- [**X**] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

3b. Other files

Which other files have been stored?

- [**X**] file(s) describing the transition from raw data to reported results. Specify:
- [**X**] file(s) containing processed data. Specify: ...
- [**X**] file(s) containing analyses. Specify: ...
- [ ] files(s) containing information about informed consent
- [ ] a file specifying legal and ethical provisions
- [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
- [ ] other files. Specify: ...

On which platform are these other files stored?

- [**X**] individual PC
- [**X**] research group file server
- [ ] other: ...

Who has direct access to these other files (i.e., without intervention of another person)?

- [**X**] main researcher

- [**X**] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

4. Reproduction

Have the results been reproduced independently?: [ ] YES / [**X**] NO

If yes, by whom (add if multiple):

Name: Address:

Affiliation: E-mail:

Data Storage Fact Sheet for Chapter 4

Name/identifier study: Chapter 4

Author: Deirdre Kavanagh

Date: 1st June 2019

1. Contact details

1a. Main researcher

Name: Deirdre Kavanagh

Address: Henri Dunantlaan 2, B-9000 Gent

E-mail: deirdre.kavanagh@ugent.be

1b. Responsible Staff Member (ZAP)

Name: Prof. dr Dermot Barnes-Holmes

Address: Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

E-mail: Dermot.Barnes-Holmes@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

Which datasets in that publication does this sheet apply to?:

It refers to all studies that are reported in Chapter 4

3. Information about the files that have been stored

3a. Raw data

Have the raw data been stored by the main researcher? [**X**] YES / [ ] NO

If NO, please justify: N/A

On which platform are the raw data stored?

- [**X**] researcher PC

- [**X**] research group file server
- [ ] other (specify): ...

Who has direct access to the raw data (i.e., without intervention of another person)?

- [**X**] main researcher
- [**X**] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

3b. Other files

Which other files have been stored?

- [**X**] file(s) describing the transition from raw data to reported results. Specify:
- [**X**] file(s) containing processed data. Specify: ...
- [**X**] file(s) containing analyses. Specify: ...
- [ ] files(s) containing information about informed consent
- [ ] a file specifying legal and ethical provisions
- [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...
- [ ] other files. Specify: ...

On which platform are these other files stored?

- [**X**] individual PC
- [**X**] research group file server
- [ ] other: ...

Who has direct access to these other files (i.e., without intervention of another person)?

- [**X**] main researcher
- [**X**] responsible ZAP
- [ ] all members of the research group
- [ ] all members of UGent
- [ ] other (specify): ...

4. Reproduction

Have the results been reproduced independently?: [ ] YES / [**X**] NO

If yes, by whom (add if multiple):

Name: Address:

Affiliation: E-mail:

Appendices

Appendix A

Chapter 2: Descriptive statistics for questionnaires in Experiment 1

Questionnaire	M	SD
CAPE (weighted scores)		
Overall Frequency	1.68	.26
Frequency of Positive Symptoms	1.38	.26
Frequency of Negative Symptoms	1.92	.46
Frequency of Depressive Symptoms	2.03	.46
Overall Distress	1.05	.37
Distress associated with Positive Symptoms	1.81	.51
Distress associated with Negative Symptoms	3.04	.86
Distress associated with Depressive Symptoms	2.39	.65
Rosenberg Self-esteem Scale (RSES)	13.82	4.46
Perspective-taking sub-scale of the IRI	17.96	3.13

Note: The maximum weighted score for all CAPE sub-scales is 4.00. The CAPE has no formal clinical cut-off. The maximum score for the Rosenberg Self-esteem Scale is 40 and the measure has no formal clinical cut-off. The maximum score for the Perspective-taking subscale of the IRI is 28 and the measure has no formal clinical cut-off.

Appendix B

Chapter 2: Descriptive statistics for questionnaires in Experiment 2

Questionnaire	M	SD
CAPE (weighted scores)		
Overall Frequency	1.69	.35
Frequency of Positive Symptoms	1.4	.31
Frequency of Negative Symptoms	1.75	.48
Frequency of Depressive Symptoms	2.22	.62
Overall Distress	2.05	.62
Distress associated with Positive Symptoms	1.85	1.09
Distress associated with Negative Symptoms	2.10	.51
Distress associated with Depressive Symptoms	4.87	4.15
Rosenberg Self-esteem Scale (RSES)	8.23	5.56
Perspective-taking sub-scale of the IRI	18.12	3.52

Note: See Appendix A.

Appendix C

Chapter 3: *Descriptive statistics for questionnaires in Experiment 1*

Questionnaire	**M**	**SD**
Acceptance and Action Questionnaire (AAQ-II)	17.16	7.14
CAPE (weighted scores)		
Overall Frequency	1.5	.19
Frequency of Positive Symptoms	1.32	.23
Frequency of Negative Symptoms	1.57	.34
Frequency of Depressive Symptoms	1.83	.41
Overall Distress	2.04	.45
Distress associated with Positive Symptoms	1.68	.53
Distress associated with Negative Symptoms	2.13	.86
Distress associated with Depressive Symptoms	2.33	.62

Note: The maximum weighted score for all CAPE sub-scales is 4.00. The CAPE has no formal clinical cut-off. The maximum score for the AAQ-II is 49 and the measure has a suggested clinical cut-off of ≥24.

Appendix D

Chapter 3: Descriptive statistics for questionnaires in Experiment 2

Questionnaire	**M**	**SD**
Self-warmth Thermometer	56.65	22.02
Psychological Flexibility Index (PFI)	355.31	27.19
CAPE (weighted scores)		
Overall Frequency	1.72	.32
Frequency of Positive Symptoms	1.43	.34
Frequency of Negative Symptoms	1.94	.44
Frequency of Depressive Symptoms	2.08	.52
Overall Distress	2.15	.54
Distress associated with Positive Symptoms	1.66	.44
Distress associated with Negative Symptoms	2.07	.58
Distress associated with Depressive Symptoms	2.54	.73
ECR-RS		
Attachment-related avoidance (Mother)	19.82	9.47
Attachment-related anxiety (Mother)	5.33	3.51
Attachment-related avoidance (Father)	23.65	9.55
Attachment-related anxiety (Father)	6.41	4.61
Attachment-related avoidance (Partner)	11.45	5.55
Attachment-related anxiety (Partner)	9.02	4.79
Attachment-related avoidance (Best Friend)	14.12	6.33
Attachment-related anxiety (Best Friend)	7.69	4.6
EOSS		
Overall EOSS	72.82	15.97
Casual acquaintances-absent	17.47	5.3
Casual acquaintances-present	23.98	4.54
Close relationships-absent	11	6.19
Close relationships-present	20.38	6.02

IOS		
Best friend	4.71	1.34
Other people	2.98	1.13

Note: The maximum score is 100 for the Self-warmth Thermometer. The maximum score for the PFI is 492. The maximum weighted score for all CAPE sub-scales is 4.00. The maximum score for each of the EC-RS attachment related avoidance sub-scales is 42 and the Attachment related anxiety sub-scale is 21. The maximum overall EOSS score is 140 with the maximum score for each sub-scale at 35. Finally, the maximum score for each of the IOS scales is 7. None of the scales have formal clinical cut-offs

Appendix E

Chapter 3: Descriptive statistics for questionnaires in Experiment 3

Questionnaires	**M**	**SD**
Self-warmth Thermometer	53.54	22.77
Psychological Flexibility Index (PFI)	323.42	23.99
Overall Other-focusing Task Score	24.94	3.18
CAPE (weighted scores)		
Overall Frequency	1.71	.44
Frequency of Positive Symptoms	1.39	.4
Frequency of Negative Symptoms	1.93	.54
Frequency of Depressive Symptoms	2.09	.62
Overall Distress	2.09	.58
Distress associated with Positive Symptoms	1.66	.51
Distress associated with Negative Symptoms	2.06	.55
Distress associated with Depressive Symptoms	2.35	.93
ECR-RS		
Attachment-related avoidance (Mother)	18.24	9.75
Attachment-related anxiety (Mother)	4.56	2.75
Attachment-related avoidance (Father)	22.52	8.64
Attachment-related anxiety (Father)	5.52	3.89
Attachment-related avoidance (Partner)	11.61	4.61
Attachment-related anxiety (Partner)	9.21	5.18
Attachment-related avoidance (Best Friend)	13.88	4.61
Attachment-related anxiety (Best Friend)	6.91	4.03
EOSS		
Overall EOSS	75.7	13.28
Casual acquaintances-absent	17.33	5.53
Casual acquaintances-present	24.97	4.83
Close relationships-absent	10.82	5.57
Close relationships–present	22.58	5.04

IOS			
	Best friend	5.06	1.14
	Other people	3.24	1.0

Note: See note for Appendix D.

Appendix F

Chapter 3: Descriptive statistics for questionnaires in Experiment 4

Questionnaire	***M***	***SD***
Self-warmth Thermometer	49.73	24.41
Psychological Flexibility Index (PFI)	346.73	24.64
CAPE (weighted scores)		
Overall Frequency	1.53	.38
Frequency of Positive Symptoms	1.33	.35
Frequency of Negative Symptoms	1.81	.49
Frequency of Depressive Symptoms	1.75	.52
Overall Distress	1.83	.55
Distress associated with Positive Symptoms	1.48	.65
Distress associated with Negative Symptoms	1.81	.61
Distress associated with Depressive Symptoms	2.16	.69
ECR-RS		
Attachment-related avoidance (Mother)	16.57	7.99
Attachment-related anxiety (Mother)	4.5	2.7
Attachment-related avoidance (Father)	23.93	10.3
Attachment-related anxiety (Father)	6.97	5.8
Attachment-related avoidance (Partner)	11.1	4.71
Attachment-related anxiety (Partner)	8.23	4.88
Attachment-related avoidance (Best Friend)	12.83	5.32
Attachment-related anxiety (Best Friend)	6.1	3.47
EOSS		
Overall EOSS	72.37	16.65
Casual acquaintances-absent	16.6	4.97
Casual acquaintances-present	23.33	4.72
Close relationships-absent	12.53	6.46
Close relationships-present	19.9	6.49
IOS		
Best friend	4.3	1.09

Other people	2.7	1.08

Note: See note for Appendix D.

Appendix G

Chapter 3: Descriptive statistics for questionnaires in Experiment 5

Questionnaire	***M***	***SD***
Self-warmth Thermometer	47.43	27.18
Psychological Flexibility Index (PFI)	358.1	23.93
CAPE (weighted scores)		
Overall Frequency	1.76	.36
Frequency of Positive Symptoms	1.46	.28
Frequency of Negative Symptoms	1.97	.53
Frequency of Depressive Symptoms	2.13	.59
Overall Distress	2.15	.47
Distress associated with Positive Symptoms	1.7	.5
Distress associated with Negative Symptoms	2.45	.77
Distress associated with Depressive Symptoms	2.01	.56
ECR-RS		
Attachment-related avoidance (Mother)	17.9	7.9
Attachment-related anxiety (Mother)	6.13	3.5
Attachment-related avoidance (Father)	24.03	9.26
Attachment-related anxiety (Father)	8.3	5.37
Attachment-related avoidance (Partner)	12.03	5.51
Attachment-related anxiety (Partner)	10.33	5.2
Attachment-related avoidance (Best Friend)	14.07	6.03
Attachment-related anxiety (Best Friend)	7.73	4.17
EOSS		
Overall EOSS	74.93	14.96
Casual acquaintances-absent	17.27	6.25
Casual acquaintances-present	24.63	4.25
Close relationships-absent	12.03	5.73
Close relationships-present	21	6.26
IOS		

Best friend	4.43	1.25
Other people	2.7	1.06

Note: See note for Appendix D.

Appendix H

Chapter 3: Descriptive statistics for questionnaires in Experiment 6

Questionnaire	***M***	***SD***
Self-warmth Thermometer	49.91	26.2
Psychological Flexibility Index (PFI)	363.13	25.44
CAPE (weighted scores)		
Overall Frequency	1.86	.37
Frequency of Positive Symptoms	1.62	.41
Frequency of Negative Symptoms	1.98	.45
Frequency of Depressive Symptoms	2.23	.49
Overall Distress	2.24	.46
Distress associated with Positive Symptoms	1.84	.49
Distress associated with Negative Symptoms	1.99	.52
Distress associated with Depressive Symptoms	2.72	.57
ECR-RS		
Attachment-related avoidance (Mother)	16.07	8.93
Attachment-related anxiety (Mother)	4.83	3.13
Attachment-related avoidance (Father)	23.07	10.70
Attachment-related anxiety (Father)	6.17	4.69
Attachment-related avoidance (Partner)	11.07	4.28
Attachment-related anxiety (Partner)	9.2	4.77
Attachment-related avoidance (Best Friend)	11.77	4.94
Attachment-related anxiety (Best Friend)	6.1	4.11
EOSS		
Overall EOSS	76.6	15.07
Casual acquaintances-absent	18.23	5.35
Casual acquaintances-present	25.77	3.53
Close relationships-absent	11.67	6.42
Close relationships–present	20.93	6.38
IOS		
Best friend	5.10	1.40

Other people	3.13	1.22

Note: See note for Appendix D.

Appendix I

Chapter 4: Descriptive statistics for questionnaires in Experiment 1

Questionnaire	***M***	***SD***
Self-warmth Thermometer	56.72	22.82
Psychological Flexibility Index (PFI)	353.59	27.15
CAPE (weighted scores)		
Overall Frequency	1.76	.35
Frequency of Positive Symptoms	1.49	.42
Frequency of Negative Symptoms	1.93	.39
Frequency of Depressive Symptoms	2.14	.51
Overall Distress	2.26	.51
Distress associated with Positive Symptoms	1.8	.54
Distress associated with Negative Symptoms	2.15	.58
Distress associated with Depressive Symptoms	2.73	.68
ECR-RS		
Attachment-related avoidance (Mother)	17.94	9.15
Attachment-related anxiety (Mother)	5.313	3.92
Attachment-related avoidance (Father)	21.84	9.61
Attachment-related anxiety (Father)	6.38	4.26
Attachment-related avoidance (Partner)	10.34	4.48
Attachment-related anxiety (Partner)	8.53	4.64
Attachment-related avoidance (Best Friend)	12.13	5.72
Attachment-related anxiety (Best Friend)	6.34	3.34
EOSS		
Overall EOSS	74.63	15.08
Casual acquaintances-absent	18.13	5.64
Casual acquaintances-present	24.25	4.26
Close relationships-absent	11.19	6.08
Close relationships–present	21.06	4.59
IOS		
Best friend	5.06	1.31
Other people	3.22	1.34

Note: See note for Appendix D.

Appendix J

Strategy Questionnaire:

1. **How successful were you at taking the perspective of the other person? ***

1	2	3	4	5
Not successful				Very successful

2. **How much of your responding on the computer task was influenced by the scenario that you read before and throughout the task?**

1	2	3	4	5
Not much				A lot

3. **Did you have a strategy to complete the task?**

*Question 1 was only presented to participants in the False Belief Condition

Similarity Questionnaire:

The following questions are about the two pictures that you brought with you today.

1. **How similar do you think you look the other person in the second picture that you brought with you today?**

1	2	3	4	5
Not similar at all				Very similar

How similar are each of these characteristics?

	Not similar at all				Very similar
Hair color	1	2	3	4	5
Age	1	2	3	4	5
Eye color	1	2	3	4	5
Skin color	1	2	3	4	5
Facial expression	1	2	3	4	5

How attractive would you rate yourself?

1	2	3	4	5
not attractive				very attractive

How attractive would you rate the person in the second picture?

1	2	3	4	5
not attractive				very attractive

Appendix K

Chapter 4: Descriptive statistics for questionnaires in the False Belief Condition for Experiment 2

Questionnaire	***M***	***SD***
Self-warmth Thermometer	58.17	18.76
Psychological Flexibility Index (PFI)	343.72	31.18
CAPE (weighted scores)		
Overall Frequency	1.75	.28
Frequency of Positive Symptoms	1.52	.31
Frequency of Negative Symptoms	1.9	.37
Frequency of Depressive Symptoms	2.02	.33
Overall Distress	2.10	.38
Distress associated with Positive Symptoms	1.77	.43
Distress associated with Negative Symptoms	2.06	.65
Distress associated with Depressive Symptoms	2.47	.55
ECR-RS		
Attachment-related avoidance (Mother)	19.69	8.5
Attachment-related anxiety (Mother)	6.31	4.34
Attachment-related avoidance (Father)	24.41	9.79
Attachment-related anxiety (Father)	6.59	4.21
Attachment-related avoidance (Partner)	11.24	4.73
Attachment-related anxiety (Partner)	8.41	4.10
Attachment-related avoidance (Best Friend)	12.86	4.22
Attachment-related anxiety (Best Friend)	7.03	4.03
EOSS		
Overall EOSS	72.69	20.05
Casual acquaintances-absent	18.93	15.85
Casual acquaintances-present	22.24	4.55
Close relationships-absent	11.14	5.57
Close relationships–present	20.38	6.28
IOS		

Best friend	4.55	1.40
Other people	3.07	1.39
Strategy Questionnaire		
How successful were you at taking the perspective of the other person? (only asked in False Belief Condition)	3.83	.65
How much of your responding on the computer task was controlled by the vignette?	3.87	1.11
Similarity Questionnaire		
How similar do you think you look to the other person?	3.2	.71
Similar hair colour	4.07	.87
Similarity in age	3.57	.94
Similar eye colour	3.37	.99
Similar skin colour	4.33	.88
Similar facial expression	3.43	1.10
How attractive are you?	3.37	.77
How attractive is the other person?	3.9	.65

Note: Minimum possible score for the Similarity and Strategy questions was 1 and the maximum score was 5. See Appendix D for all other minimun and maximum scores.

Chapter 4: Descriptive statistics for questionnaires in the Control Condition for Experiment 2

Questionnaire	***M***	***SD***
Self-warmth Thermometer	56.55	19.69
Psychological Flexibility Index (PFI)	350.17	26.51
CAPE (weighted scores)		
Overall Frequency	1.75	.29
Frequency of Positive Symptoms	1.52	.28
Frequency of Negative Symptoms	1.88	.44
Frequency of Depressive Symptoms	2.09	.43
Overall Distress	2.03	.49
Distress associated with Positive Symptoms	1.73	.52
Distress associated with Negative Symptoms	1.91	.76
Distress associated with Depressive Symptoms	2.40	.59
ECR-RS		
Attachment-related avoidance (Mother)	20.72	10.02
Attachment-related anxiety (Mother)	6.76	5.55
Attachment-related avoidance (Father)	22.55	10.04
Attachment-related anxiety (Father)	6.1	4.97
Attachment-related avoidance (Partner)	11.03	3.95
Attachment-related anxiety (Partner)	8.38	4.44
Attachment-related avoidance (Best Friend)	14.35	6.77
Attachment-related anxiety (Best Friend)	6.89	3.9
EOSS		
Overall EOSS	73.17	11.94
Casual acquaintances-absent	15.86	5.104
Casual acquaintances-present	24.66	3.22
Close relationships-absent	12.07	6.65
Close relationships–present	20.59	5.07
IOS		
Best friend	4.52	1.43

Other people	3.03	1.38
Strategy Questionnaire		
How successful were you at taking the perspective of the other person? (only asked in False Belief Condition)		
How much of your responding on the computer task was controlled by the vignette?	3.63	1.35
Similarity Questionnaire		
How similar do you think you look to the other person?	3.04	.69
Similar hair color	4.13	.9
Similarity in age	3.79	.69
Similar eye color	3.67	1.24
Similar skin color	4.5	.59
Similar facial expression	3.42	.97
How attractive are you?	3.21	.59
How attractive is the other person?	3.52	.73

Note: See Appendix D.

Appendix L

Chapter 4: Themes that emerged from the open-ended strategy question in Experiment 2

Theme	***False Belief Condition***	***Control Condition***
Linked the object words and pictures together	15	10
Focused only on the object word and didn't read the full sentence	4	6
Rehearsed the link between the object and person before the IRAP	4	1
Focused on the color of the picture	2	1
Relied on the IRAP feedback to aid their performance	3	0
Focused on the scenario	0	2
Initially relied on the scenario but then switched	0	1
Trial and error to respond	1	0
Ignored the Scenario	0	1
Other	3	1
No details provided	1	1

Note: The responses from some participants included several categories.

Appendix M

Chapter 4: Descriptive statistics for the scarf-glove sequence/vignette-consistent first condition in Experiment 3

Questionnaire	***M***	***SD***
Strategy Questionnaire		
How successful were you at taking the perspective of the other person?	3.71	.92
How much of your responding on the computer task was controlled by the vignette?	3.18	1.07
Similarity Questionnaire		
How similar do you think you look to the other person?	3.06	.66
Similar hair colour	3.69	1.08
Similarity in age	3.77	.90
Similar eye colour	3.18	.95
Similar skin colour	4.18	.81
Similar facial expression	3.29	.99
How attractive are you?	3.24	.66
How attractive is the other person?	3.71	.99

Note: See Appendix K for the possible maximum and minimum scores.

Chapter 4: Descriptive statistics for the scarf-glove sequence/vignette-inconsistent first condition in Experiment 3

Questionnaire	***M***	***SD***
Strategy Questionnaire		
How successful were you at taking the perspective of the other person?	3.93	.59
How much of your responding on the computer task was controlled by the vignette?	3.67	1.18
Similarity Questionnaire		
How similar do you think you look to the other person?	3.43	.51
Similar hair colour	4.14	1.03
Similarity in age	3.21	.98
Similar eye colour	3.36	1.28
Similar skin colour	4.43	.76
Similar facial expression	3.29	.91
How attractive are you?	3.07	.48
How attractive is the other person?	4.07	.62

Note: See Appendix K for the possible maximum and minimum scores.

Chapter 4: Descriptive statistics for the glove-scarf sequence/vignette-consistent first condition in Experiment 3

Questionnaire	***M***	***SD***
Strategy Questionnaire		
How successful were you at taking the perspective of the other person?	4.06	.85
How much of your responding on the computer task was controlled by the vignette?	2.75	1.18
Similarity Questionnaire		
How similar do you think you look to the other person?	3.2	.41
Similar hair colour	4.13	.81
Similarity in age	3.88	1.09
Similar eye colour	3.38	1.2
Similar skin colour	4.13	.81
Similar facial expression	3.25	.78
How attractive are you?	3.18	.40
How attractive is the other person?	4.06	.68

Note: See Appendix K for the possible maximum and minimum scores.

Chapter 4: Descriptive statistics for the glove-scarf sequence/vignette-inconsistent first condition in Exoeriment 3

Questionnaire	***M***	***SD***
Strategy Questionnaire		
How successful were you at taking the perspective of the other person?	3.56	.96
How much of your responding on the computer task was controlled by the vignette?	2.75	1.53
Similarity Questionnaire		
How similar do you think you look to the other person?	3.06	.57
Similar hair colour	4.31	.70
Similarity in age	4.13	.81
Similar eye colour	3.06	1.34
Similar skin colour	4.31	.70
Similar facial expression	3.07	1.58
How attractive are you?	3.06	.68
How attractive is the other person?	3.56	1.03

Note: See Appendix K for the possible maximum and minimum scores.

Appendix N

Chapter 4: Themes that emerged from the open-ended strategy question in Experiment 3

Theme	***Consistent Rule First***	***Inconsistent Rule First***
Linked the object words and pictures together	7	9
Focused only on the object word and didn't read the full sentence	7	4
Rehearsed the link between the object and person before the IRAP	8	1
Focused on the colour of the picture	1	0
Relied on the IRAP feedback to aid their performance	5	9
Focused on the scenario	6	10
Initially relied on the scenario but then switched	1	1
Trial and error to respond	2	3
Ignored the scenario	5	3
Other	5	1
No details provided	0	0
Focused on what they had	1	0

Note: The responses from some participants included several categories.