An Investigation Exploring The Effects Of Empathy On Spontaneous Visual Perspective-Taking in Humans

TOWNROW, LUKE,ALFRED,JAMES

An Investigation Exploring The Effects Of Empathy On Spontaneous Visual Perspective-Taking in Humans

Luke Townrow

Department of Psychology

Master of Science by Research

2023

Table of Contents

List Of Tables

List Of Figures

Acknowledgments

I would like to extend my deepest gratitude to my supervisors Prof. Zanna Clay (Durham University, UK), Prof. Daniel Smith (Durham University, UK) and mentor Dr. Christopher Krupenye (Johns Hopkins University, USA) for all of the support and guidance during my project. I would also like to extend my appreciation to Dr. Jeff Bowen (Johns Hopkins University, USA) for help with my statistical analyses. Thank you to Georgia Sandars, Millie Brunton and Stephanie Kordon for their volunteering to act in the video stimuli created. Thank you to Dr. Jessie Adriaense for discussions of design considerations at an early stage of the project and Dr. Robert Hepach for providing the control bubble stimuli used in experiment 2. Finally, thank you to Sam Cooper and Ella Smith for their help with data collection preparation as research assistants.

Abstract

Humans live in a dynamic social world and possess advanced Theory of Mind capabilities that facilitate our navigation of these complex environments. One subcomponent of Theory of Mind is visual perspective-taking, an individual's ability to represent the visual field of another being (Flavell, 1977). Although this process is typically considered to be explicit in its computation, perspective-taking can also occur implicitly or spontaneously (Samson et al., 2010). Far less is known about the flexibility of spontaneous perspective-taking and how it may be influenced by other fundamental social processes such as state empathy. With empathy being our ability to recognise, understand and share the emotional states of others (e.g., Cuff et al., 2016), many consider empathy and perspective-taking to be distinct but not mutually exclusive. In this thesis I set out to investigate if and how state empathy has an effect on spontaneous perspective-taking. After validating my own empathy inducing stimuli (*experiment 1*), I incorporated it within a shortened-modified *Dot Perspective Task* paradigm. As a paradigm investigating spontaneous perspective taking abilities, the Dot Perspective Task requires participants to make rapid judgments regarding the content of either their own visual field or that of a displayed avatar. Across trials, where the two perspectives are consistent or inconsistent with each other, participants typically make two types of errors, reflected in greater error rates and slower response times: egocentric intrusion (processing your own perspective hinders judgements you make regarding someone else's) and altercentric intrusion errors (processing another's perspective negatively impacts judgements you make on your own). Aiming to investigate the precise nature of empathy's influence on implicit perspective taking, I tested two hypotheses with divergent predictions. According to the *self-other distinction hypothesis*,empathy makes perception of the self and other more salient, and predicts greater egocentric and altercentric intrusion when the two perspectives do *not* differ. Alternatively, according to the *self-other merging* hypothesis,representations of the two perspectives can blur, which predicts greater intrusion effects when the two perspectives *do* differ. In study 1, I validated my empathy induction protocol by showing that stimuli elicit state empathy in participants. Then in a pilot of the main experiment 2, I found preliminary support for the self-other distinction hypothesis. However, this finding was not replicated in the main experiment, with evidence of no significant effect of condition (empathy vs baseline) on spontaneous perspective-taking. This

discrepancy thus makes it hard to draw strong conclusions regarding empathy's effect on spontaneous perspective taking. However, replicating literature employing the original paradigm, I did reliably document egocentric intrusion errors; this demonstrates that this paradigm can successfully be adapted to an online format with fewer trials. As a result, this thesis enables novel future directions with Dot Perspective Task paradigm adaptations and empathy induction methods whilst demonstrating the importance of replication in science.

Introduction

General Introduction

Humans live in a complex and dynamic social world, and compared to other species, humans are considered to be ultrasocial (Tomasello, 2014). This ultrasociability is facilitated by our advanced Theory of Mind capabilities (ToM), which refers to our ability to infer the mental states (emotions, desires, perspectives, beliefs, etc) of other people (Krupenye & Call, 2019; Premack & Woodruff, 1978). ToM capabilities are considered crucial in helping us navigate our complex social interactions, by facilitating our ability to understand, predict and even influence the behaviours of others. One subcomponent of theory of mind, called visual perspective-taking, is a considered a fundamental socio-cognitive process that involves taking the perspective of another individual by mentally representing their visual field (O'Grady et al., 2020; Surtees & Apperly, 2012).

Visual perspective-taking underlies a multitude of social behaviours, ranging from lying and deception, communication and even non-communicative interactions, such as physical navigation around other social agents. Visual perspective-taking can require the explicit computation of another's visual field (i.e., requiring conscious effort), a proposal supported by evidence that perspective-taking can be negatively influenced by executive function (Carlson & Moses, 2001) and cognitive load (Epley et al., 2004). However, perspective-taking can also be implicit: representations of others' visual perspectives can be rapidly computed without conscious effort, permitting swift judgements about another's perspective (Samson et al., 2010). The rapid nature of visual perspective taking provides a vital benefit for humans as it facilitates smooth and unconscious coordination, helping us to navigate our complex social environments and build, maintain, and repair social relationships. A critical question that will be the focus of this thesis, is how flexible this potentially automatic mechanism is and whether it interacts with - and can be manipulated by - other socio-cognitive and affective processes, like state empathy.

Samson and colleagues (2010) first identified forms of implicit perspective-taking with the *Dot Perspective Task* (DPT; Figure 1 - Samson et al., 2010). The DPT is considered as a test of

Level 1 – Visual Perspective Taking (L1-VPT), referring to an individual's understanding of *what* can or cannot be seen within another's visual field. L1-VPT is contrasted with Level 2-VPT which requires an individual to create a more complex mental representation of *how* things look from another's point of view (Flavell, 1977). As reproduced in Figure 1, the DPT displays a centrally located avatar within a computer-generated empty room, with the avatar facing either toward the left or right wall (side profile). On said walls, a range of dots (typically 0 -3) are presented. These dots are used as the content of the participant and avatar's perspective. Trials begin with a word representing whose perspective should be computed later in said trial (You vs She). A number is then displayed and refers to the dots soon to be shown in the room scene image presented shortly after. Each trial requires the participant to make rapid 'correct/incorrect' judgements about whether the number previously shown correctly represents the amount of dots that either the participant themselves (*Self* trials) or the avatar (*Other* trials) can see.

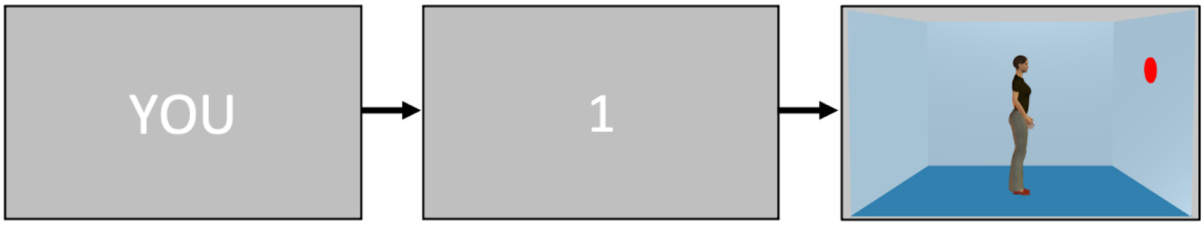


Figure 1: Stills of stimuli used in Samson et al.'s (2010) original Dot Perspective Task accessed through open access at: https://doi.org/10.6084/m9.figshare.1455943.v1. Figure displays order of slides for each trial including the computer generated room scene where participants are required to submit their response.

Samson et al (2010) were the first to demonstrate that when participants are asked to make rapid judgments about their own or another's perspective for inconsistent trials, they show evidence of two types of what are known as *intrusion error* effects: egocentric and altercentric. Egocentric intrusion refers to the tendency for participants to make slower and more erroneous judgements about another's perspective when it conflicts with their own (i.e., in inconsistent trials relative to consistent ones). Altercentric intrusion describes instances where individuals make slower and more erroneous judgements about their own perspective when processing a conflicting perspective held by another (i.e., in inconsistent trials relative to consistent ones). Altercentric intrusion provides evidence for a spontaneous perspective taking mechanism, suggesting that humans compute the perspective of another agent even when it is irrelevant to their own task (i.e., when judging their own perspective) (e.g., O'Grady et al., 2020; Samson et al., 2010; Surtees & Apperly, 2012).

According to the *mentalizing* account, the presence of both intrusion types indicate an implicit ToM system at play during the completion of the DPT. This holds that humans are able to rapidly and implicitly compute another individual's visual field ((O'Grady et al., 2020; Samson et al., 2010; Fan, 2021). As a result of an ability to reflexively infer perspectives, researchers have suggested that this capacity is driven in large part by the social nature of the avatar present within the task - that sociality directly influences human attention (Furlanetto et al., 2016). Contrary to the mentalizing approach, the *sub-mentalizing* account argues against this ToM component of DPT performance, instead attributing 'success' to lower-order domain general processes (e.g., Heyes, 2014). This account has received attention and support through findings that intrusion error types can be achieved using 'semi-social' stimuli. Semi-social objects include stimuli that include salient directional cues to elicit egocentric intrusion, such as shapes ranging from arrows (Santiesteban et al., 2014) to fans (Vestner et al., 2022). Interestingly, research supporting the sub-mentalizing account records instances of egocentric intrusion effects but questions the presence of altercentric intrusion (Cole et al., 2016; Gardner et al., 2018; Heyes, 2014; Santiesteban et al., 2014, 2017).

Although traditionally seen as mutually exclusive accounts, more recent work has focused on developing a more integrative approach. A recent meta-analysis by Shin, Holland and Phillips (2021) reviewed the DPT literature and concluded that as evidence of both accounts (with the domain-general approach accounting for more) exists, the DPT cannot simply be explained by one approach over the other. Adopting a holistic view has been considered more closely by others who have proposed an integrative way of approaching the underlying mechanisms at play during the DPT (Capozzi & Ristic, 2020; Pesimena & Soranzo, 2023). Such an integrative account entertains the idea that both directional cueing and mentalization are two processes present during the DPT task. Specifically, whilst lower-order directional cueing provides the presence of the intrusion interference itself, the more ToM-focused process determines the magnitude of these effects.

Regardless of employing a mentalizing, sub-mentalizing or integrative approach to explain DPT performances, the presence of implicit perspective-taking seems to depend on fundamental socio-cognitive processes. Recent work with a particular neurodivergent population, hinted at a possible relationship with a process with affective and cognitive components. Drayton, Santos & Baskin-Sommers (2018) compared performance on the DPT paradigm between a neurotypical sample and a sample of prisoners diagnosed with psychopathy. Broadly speaking, psychopathy refers to a mental condition characterised by atypical, antisocial emotional and behavioural symptoms, with a particularly distinct deficit in empathy and increased criminality (Blair et al., 2006a; Blair et al., 2006b; Frick et al., 1994). Interestingly, while psychopaths demonstrated comparable egocentric intrusion effects relative to a neurotypical sample, they showed diminished altercentric intrusion, they were apparently less influenced by others' perspectives if they conflicted with their own. These findings raise exciting questions about the relationship between the social processes of empathy and perspective-taking. Developing upon the results from psychopaths, Yue and colleagues (Yue et al., 2017) looked at implicit perspective taking linked with levels of oxytocin. Oxytocin is a hormone commonly referred to as the 'love' or 'bonding hormone' (Watson, 2021)thanks to its link to maternal behaviours (Lee et al., 2009), social affiliation and facilitation of empathy (Geng et al., 2018). Yue et al's research indicated that oxytocin treatment yielded an improvement in DPT performance when taking another's perspective for women, yet not for men. Researchers concluded that these results could be driven by oxytocin's effect on reducing self-bias and making individuals more attuned to other people's perspectives. My investigation seeks to develop upon the work by Drayton, Santos and Baskin-Sommers (2018) and Yue and others (2017) to focus on empathy as a cognitive process that could impact implicit, spontaneous perspective taking abilities.

To consider state empathy's influence on spontaneous perspective-taking, it is important to consider their conceptual framing and proposed relation between the two. Whilst empathy has a multitude of different definitions, most consider empathy to be a multifaceted socio-cognitive process that allows us to identify, understand and share the emotional and mental states of others (Cuff et al., 2016; Decety & Lamm, 2006; Preston & Waal, 2002). Broadly speaking, empathy can take two forms: trait empathy (referring to one's ability to express and experience empathy, associated with character traits) and state empathy (see Lyu et al., 2022). State empathy is considered an affective response that is temporary and a reaction to a concrete event (Van der Graaff et al., 2016) and is the focus of this thesis. Empathy contains multiple socio-cognitive and socio-affective components that although are distinct from one another, are not considered mutually exclusive as they commonly interact. Either independent or concurrent, the subcomponents of empathy can impact other cognitive processes like executive functions, attention and even perspective-taking (Choi & Watanuki, 2014; Vescio et al., 2003; Yan et al., 2020).

Cuff and colleagues provide a holistic definition of empathy that lays out 8 principle dimensions (see (Adriaense et al., 2020). Cuff touches on empathy's affective component (emotional responses) in addition to its cognitive element (via understanding others emotions) whilst also stressing empathic processes can be automatic as well as explicit in a top-down control process. This view of empathy suggests a differentiation between empathy and perspective-taking; whilst empathy has a cognitive and affective characterization, visual perspective-taking only includes the former. Regardless of their concurrent appearances, taking another's perspective is not sufficient for experiencing empathy as empathy can only occur when perspective-taking is experienced alongside emotional engagement (de Waal, 2008). This aligns with a dimension found in Cuff et al's definition whereby many researchers consider empathy to have both an affective and cognitive component. Whilst the former refers to the foundation of emotional states sharing that empathy is built upon, the cognitive component is considered to be perspective-taking. Additionally, as seen with perspective-taking, empathy is thought to be both under explicit control and an automatic process; researchers argue for both control and automaticity as it can be automatically elicited upon viewing stimuli (Singer et al., 2004), yet we are also able to keep the stimuli in memory and contemplate (Hodges & Wegner, 1997). Through a deeper consideration into the dimensions of empathy, (especially those of affective vs cognitive and automatic vs controlled), one can see the associations to be made between empathy and visual perspective-taking. Not only do both processes allow us to behave appropriately in different situations, empathy is considered by many to have behavioural outcomes (Eisenberg & Miller, 1987 as cited in Cuff et al, 2016) of which spontaneous perspective-taking may be one.

An investigation into implicit, spontaneous form of perspective-taking is rightly informed by an evolutionary perspective. Automaticity has been said to be a marker of foundational mechanisms (Phillips et al., 2021) in human behaviour, which are likely to be also shared with nonhuman animals. Speaking further on the close relation between empathy and perspective-taking through the lens of comparative cognition, one approach to consider is through the Russian Doll Model of Empathy (de Waal, 2003; 2008). This model stresses the role that both higher-order complex and lower-order simple mechanisms play on driving empathy. The model additionally lays out how empathy has an affective basis that becomes deeply integrated with socio-cognitive processes to contribute to empathy's mozaic-like structure of subcomponents.

At the centre of the Russian Doll model is the proposed Perception-Action Mechanism (PAM; Preston & de Waal, 2002), an evolutionarily present mechanism that allows certain human and nonhuman animals to access the affective state of another by representing it as our own (somewhat similar to emotion contagion). As the name suggests, it claims perspective and action share representations that allow for the experience of emotion contagion. The presence of the PAM highlights how simple processes can have an impactful outcome on behaviour, therefore closely linked to the concept of self-other merging. The presence of the PAM has been linked to human research, finding that there are physiological similarities when it comes to both observing and experiencing an emotion (Adolphs et al., 2000). De Waal's Russian Doll Model alongside the PAM (Preston & de Waal, 2002; de Waal, 2003) support the hypothesis that self-other merging effects on empathy would be driven by an evolutionary socio-affective process that simultaneously acts as the basis of empathy itself. Following the predictions of this comparative model, self-other merging effects on spontaneous perspective-taking would be rooted in a low-level, simple, evolutionarily ancient process, whereas if effects are driven by self-other distinction, it is still a fundamental process but considered higher-order in comparison.

A key characteristic of empathy is that an individual can detect that their own affective state is separate to that of its *external* source, typically in the form of an emotional state of another agent (de Vignemont & Singer, 2006; Lamm et al., 2016, 2019). This differentiation between the self and the other critically distinguishes empathy from emotional contagion, which involves a form of merging of the states of self and other (Decety & Lamm, 2006). A core focus of debate in empathy research is the degree to which empathy involves self-other distinction as compared to self-other merging . The view of empathy as involving distinction between self and other is arguably the more well-represented in the human psychological literature. Whilst self-other distinction is considered to have an important contribution to the definition of empathy within the human literature, this to some extent contrasts with certain models from an evolutionary perspective where emphasis is placed on self-other merging as a fundamental foundation.

Emotion contagion is considered to be the lowest shared process under empathy whereby an individual is affected by another's state, therefore one is not required to have an understanding that their own emotional state is causally linked to that of another (de Waal, 2008). As stated, in order to move beyond emotion contagion and toward empathy, one has to attribute another's state as being the cause of their own emotional state and this is thanks to a shift in attention away from the self and toward the other. This process is said to require self-other differentiation to allow an individual to relate to another's mental state, while detecting the difference between the two individuals (Hoffman, 1982). Self-other distinction has also been suggested to reduce someone's own personal distress after being exposed to someone else's distress. In this respect, through the process of self-other differentiation, empathy can allow the actor to recognise the needs of the other and perform appropriate prosocial behaviours to assist them (Lamm et al., 2019).

Seemingly in contrast to self-other distinction required for empathic other-orientation, is the concept of merging or *oneness*, referring to a blurring between perceptions of the self and the other. This is described to conversely also help to promote a sense of empathy and shared experience (Decety & Lamm, 2006). Much of the work supporting the presence of self-other merging stems from Batson & Shaw (1991) looking at empathic concern and altruism through prosocial behaviours and neuroscientific investigations. Such neuroscientific work has focused on shared pain experiences, studies finding that observing someone else in pain results in activation in cortical regions commonly correlated with first-person pain experiences (Singer & Lamm, 2009). As an additional polarising discussion, it is important to address sympathy alongside empathy. Whilst many of those in the psychological literature argue for a clear distinction between the two (e.g., Singer & Lamm, 2009), others consider sympathy and empathy to be strongly linked (de Waal, 2008). This study did not explore this argument as it was not central to my investigation.

There is still active debate regarding the extent to which the self-other distinction and self-other merging aspects of empathy are mutually exclusive or related (Lamm et al., 2016). The coexistence of these two sub-components suggest that although it is important for one to distinguish their own mental state from that of another, a sense of self-other merging could promote empathy through feelings of closeness. Furthermore, the argument against this mutual exclusivity of merging and distinction is also supported in the evolutionary literature with the Russian Doll model going so far as to place PAM at the basis of empathy and acknowledging that an attentional shift to the *other* is required. What is unclear however, is whether one of these sub-components plays a more prominent role when it comes to empathy's potential effect on perspective-taking. Both concepts theoretically have differing implications that help us make differing predictions on the effect empathy may have on the process of spontaneous perspective-taking.

Given the apparent shared properties and relation between implicit perspective-taking and empathy, it is important to investigate the flexibility of spontaneous perspective-taking and how it may be affected by empathy. Across two experiments I set out to investigate the relationship between these two fundamental social processes and how these effects may manifest on an implicit and automatic level of cognition. Developing on the work by Drayton, Santos & Baskin-Sommers (2018) and Yue and colleagues (2017), this study investigated whether spontaneous perspective taking effects are influenced by state empathy.

I opted for a within-subjects approach, largely to help control for between-subject variation. Individuals will vary both in terms of their response times and the computer devices and connectivity they have access to. A within-subjects design helps to control for these individual differences, therefore allowing for more direct conclusions to be drawn regarding the flexibility of spontaneous perspective-taking and the extent to which state empathy can shape it. More specifically, DPT performances for each participant can be compared following exposure to an empathy-inducing vs an emotionally neutral stimuli, allowing for the targeting of state empathy. The Dot Perspective Task is already a within-subjects design, with participants completing trials of each variable’s levels and seminal work has used the same approach when manipulating the type of avatar stimuli (Santiesteban et al., 2014). To my knowledge, the only other experiment investigating spontaneous perspective taking in relation to empathy was by Drayton, Santos & Baskin-Sommers (2018), who compared performances between neurotypical individuals and those clinically diagnosed with psychopathy using a between-subjects design. In contrast, this experiment looked to utilize a manipulation employed within-subjects to help contrast performances under state empathy and a neutral affect baseline condition. A closely related study (Simpson and Todd, 2017), using a within-subjects design considered how group membership would influence spontaneous perspective-taking intrusion effects. The authors had participants complete 2 iterations of the DPT in one sitting (within-subject design) where the avatar in one condition represented their own ingroup, whereas the next represented an outgroup. Experiment 1 used cartoon mascots that were either associated with participants’ own University college (i.e. sport team mascots) or that of a rival. Experiment 2 utilized human CGI avatars that wore colored shirts that were representative of an arbitrary group that each participant had been assigned to prior to the start of the experiment (orange vs green). Simpson and Todd ultimately found that there was no significant interaction between group membership and intrusion errors, a surprising finding to the reader. Whilst it is incredibly likely this non-significant conclusion is the consequence of group membership simply not interacting with implicit perspective-taking, it could alternatively be the result of methodological limitations. Considering points surrounding cartoon mascot stimuli, and the creation of arbitrary groupings by colour, I wish to keep these viable limitations in mind when developing new stimuli to account for possible negative influences on findings.

If state empathy is a predominantly other-oriented process, one should expect it to therefore enhance other-oriented visual perspective-taking. Despite perspective-taking traditionally being considered in its most deliberative sense, I am more specifically interested in looking at the more rapid forms of perspective-taking occurring outside conscious awareness, processing that nevertheless has a profound impact on our everyday social lives and decision making. Given the importance of such a cognitively implicit process, it is important to understand how it can be impacted, hindered or enhanced by socio-affective processes.

Overview of experimental design, hypotheses and predictions

Implementing a novel version of the DPT, I thus used a within-subjects approach to first induce a state of empathy in order to investigate (a) whether empathy influences spontaneous visual-perspective taking and (b) if so, to test two mechanistic hypotheses about the way in which it does so.

I tested two possible hypotheses for the effect that state empathy has on egocentric and on altercentric intrusion in spontaneous visual perspective taking.

Hypothesis 1: if empathy elicits *Self-Other Distinction*, it will make salient the states of both the Self and the Other, potentially producing interference even when they do not conflict.

Hypothesis 2: if empathy elicits *Self-Other Merging*, it will blur distinctions between the states of the Self and Other, potentially heightening interference when these states of perspective conflict.

Null Hypothesis: Empathy (regardless of which sub-component) does not have a significant effect on spontaneous perspective-taking performances

To test these hypotheses, I developed a variant of the DPT in which participants first viewed empathy-induction or control videos and then participated in blocks of the DPT. Critically, the empathy-induction videos were designed to induce empathy for a particular character who was then featured as the avatar in the DPT. I then measured levels of egocentric and altercentric interference based on manipulations of self and avatar perspectives being congruent or incongruent. Predictions, from these hypotheses, are provided below.

Firstly, if inducing empathy elevates self-other distinction, I predicted that participants would exhibit slower and more erroneous responses for self-consistent trials in the empathy condition compared to baseline specifically. This would be the case as I would expect self-other distinction to make the avatar’s perspective more salient, even when the two perspectives are consistent and congruent. Conversely, I predicted faster response times and greater accuracy for inconsistent trials in the empathy condition in comparison to consistent trials because the high saliency of the Other will make it easier for participants to disambiguate differing perspectives. Unlike inconsistent trials, consistent trials saw perspectives of the self and other being congruent, thus making it harder to disambiguate from this self-other distinction account. This result would suggest the presence of altercentric intrusion specifically in the empathy condition arising from faster/more correct responses in the inconsistent trials rather than the consistent ones. This contrasts expected results in the baseline condition where reports of altercentric intrusion would arise from faster/more correct responses in the consistent trials rather than the inconsistent ones.

On the other hand, if inducing empathy elicits greater experiences of self-other merging, I predicted stronger egocentric and altercentric intrusion effects in the empathy condition than in the baseline condition, regardless of perspective consistency. To elaborate, these would replicate the significant difference between the two intrusion error types seen in previous literature, however, in the empathy condition would see faster response times and greater accuracy for consistent trials (given no need to disambiguate perspectives) and/or slower, more incorrect responses to inconsistent trials (given challenges in disambiguating perspectives).

2.1 Experiment 1: Empathy Manipulation Validation

As a first step, I developed and validated a novel set of video stimuli that I designed to elicit empathy in participants. If empathy videos were effective in inducing empathy, I predicted that they would evoke a more negative emotional state as compared to the neutral control videos; evidence of this prediction will be tested through scores of self-reported affect in terms of both valence and strength.

2.2 Experiment 2: Effect of Empathy on spontaneous visual perspective taking

In the primary experiment, I implemented an empathy induction prior to applying the DPT. As a validation of my method, I expected to replicate Samson et al.'s (2010) original findings in my baseline condition (instead of an emotive video, participants were shown a short video of bubbles): that is, I expected to see faster reaction times in consistent trials compared to inconsistent trials for both *self* (evidence of altercentric intrusion) and *other* trials (evidence of egocentric intrusion).

Methods

Experiment 1: Empathy Manipulation Validation

Ethics

These experiments were approved by the Ethics Committee of the Psychology Department at Durham University, reference code: PSYCH-2022-02-08T09_37_57-zmdb44.

Participants

A sample of 41 individuals were recruited through social media platforms (women = 26, men = 15, M age = 24 years ±4.2, age range = 19-36y). Although not a focus of these analyses, for supplementary analyses, the demographic information about this participant sample were: 59% WhiteBritish, 12% Irish, 10% African, 7% Gypsy or Irish Traveller and 2% each Caribbean, Roma, Brazilian, Other African or Caribbean background and Other White. All participants currently resided in the United Kingdom, with the UK also being the country of origin of 78% of the sample. Inclusion criteria included passing two visual (identifying the actor in the video from two stills) and auditory (multiple choice question on the name of the actor's daughter/partner) attention check questions. All participants passed these attention checks.

Design

This Qualtrics-based experiment implemented a within-subjects design. In a counterbalanced order, participants watched a video designed to induce empathy and a neutral control (independent measures). They then reported their affect on a Likert scale ranging from 'extremely negative' (0) to 'extremely positive' (4). Two video sets were made each with 2 video stimuli (an empathy inducing stimuli and a video including a neutral storyline). The sample was split into two groups (Group 1: n=20; Group 2: n = 21) – each viewing the Test stimuli (emotive video) from one set and control stimuli (the neutral video) of another.

Empathy-Induction Stimuli

The stimuli of this experiment consisted of four videos designed to either induce empathy or act as a control (emotionally neutral) stimuli. Previous literature has shown that various methodologies can successfully induce empathy, including the use of vignettes (Hein et al., 2018), short films (Barlińska et al., 2015) and even music (O'Neill & Egermann, 2022). Due to the nature of the perspective-taking design I planned to use, it was important that the induction method used could directly relate to DPT. Consequently, I wanted to ensure participants felt an empathic state toward an individual that could be present at both phases of the experiment (induction and DPT). The use of short films was deemed to be most appropriate for the task. However, as profile images of the agents were required to be incorporated in the DPT trial scenes (see Figure 4), short films were made from scratch, allowing us to have a greater amount of control over the design.

I created my own video manipulations inspired by the content and format of the EmpaToM/EmpaToM-Y paradigm (Breil et al., 2021; Kanske et al., 2015) and videos created by Barraza and Zak (Barraza & Zak, 2009). Whilst the EmpaTom paradigms used short, dynamic, naturalistic videos within a measure of empathy and ToM, the empathy induction video used by Barraza and Zak was longer and involved a father telling the viewer about his young son with terminal brain cancer. Although both of these manipulations were successful in inducing empathic states in their participants, for the perspective-taking task in experiment 2 I required greater control over the videos when it came to using the focus of the videos to pose for images to be included as the avatars in the DPT empathy condition(s).

To address thisI created two sets of stimuli, each involving an empathy-inducing and control video. '*Rosie's Pond'* told a short story of a young mother at a local park talking about her daughter, whereas '*Katie's Move'* presented a young woman packing up items in an office belonging to her partner. Each stimuli-set was composed of 2 visually-identical video stimuli of approx. 2mins duration – Test and Control stimuli - depicting a visual scene with an actor expressing a neutral facial expression throughout, see Figure 2. The manipulation was in the audio narrations played over the videos and a short interview scene appearing within each. In the test conditions, actors in either video set described an emotional story of extreme familial loss, whereas videos in the control conditions touched on themes of relaxation and recollection of past events. More specifically, the two videos in the *Rosie's Pond* set portrayed a story of the loss of a child to cancer (test/empathy stimuli) and one about a young mum talking about relaxing after dropping her daughter at daycare (control/neutral stimuli). The two videos of Katie's Move portray a young woman's grief following the loss of her husband in a car accident (empathy) and the same woman packing her husband's desk because he has received a promotion as she reminisces about a trip abroad they took recently (neutral). The same actor played the role of the corresponding character for the empathy and neutral stimuli in each video set (total of two actors across four videos). In an attempt to ensure a high degree of control over the videos and minimise the extraneous differences between the videos, more basic processes or factors that could have an impact on emotionality induced from watching were considered and controlled for. Although stories, actors and locations differed between video set 1 and 2, high degree of control can be seen through the use of actresses from the same demographic, videos of the same length, comparable shot types used and keeping actress screen times consistent between videos. Given that gender was not a primary factor of interest (but can influence empathy), we kept the gender of the actors consistent across videos by using only women. See Table 1 and Figure 2 for a brief summary of each video and the differences between the videos within and between sets. All videos were shot using an iPhone XR and edited using Adobe Premiere Pro software.

Table 1: A table expressing the different features of the 4 created video stimuli across video sets (Rosie's Pond vs Katie's Move) and conditions (empathy vs neutral). *This was increased to match actor's screen time to the other videos

	VIDEO SET 1		VIDEO SET 2	
FEATURES	ROSIE'S POND	ROSIE'S POND	KATIE'S MOVE	KATIE'S MOVE
CONDITION	Emotive	Neutral	Emotive	Neutral
THEME	Familial Loss	Relaxation	Familial Loss	Recollection
PLOT SUMMARY	A young mother visiting a local pond reminiscing over the loss of her infant daughter to cancer	A young mother visiting a local pond after dropping her infant daughter at school, calming before work	A woman who is at her partner's work office packing his belongings following his untimely death	A woman who is at her partner's work office packing his belongings following his newly provided promotion

SCRIPT LENGTH (word count)	301	320	315	378*
VIDEO LENGTH (minutes)	1:49	1:49	1:49	1:49
SCENE CHANGES	14	14	14	14
ACTOR SCREEN TIME (minutes)	1:44	1:44	1:44	1:44
ACTOR	White, British female	White, British female	White, British female	White, British female

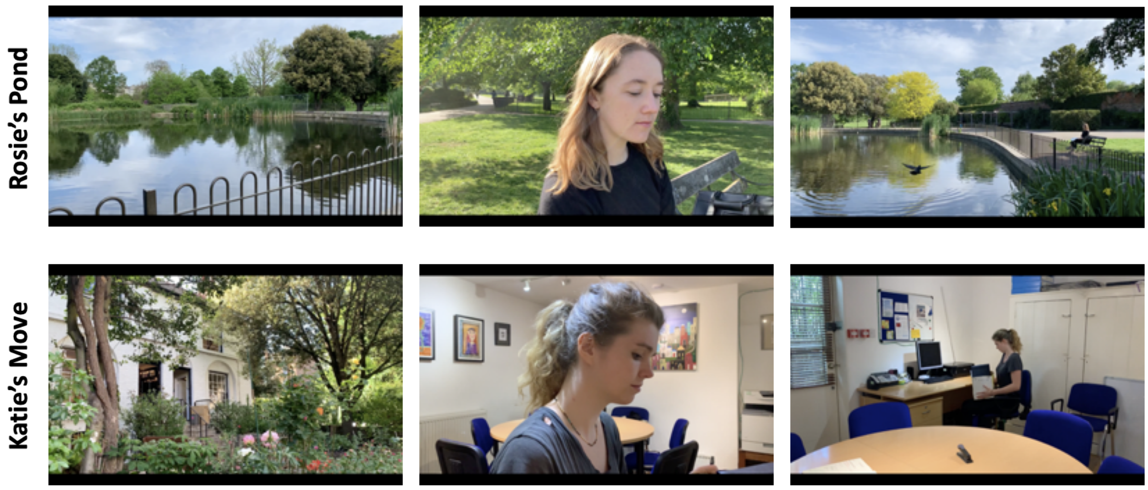


Figure 2: Figure displaying a comparison of still scenes from both video sets (Rosie's Pond and Katie's Move). Each row categorises the video set and the columns represent the same fixed time point for each video. This comparison is used to demonstrate the control attempted to achieve with minimal differences between camera shots.

To measure the reported affect induction (Kanske et al, 2015; Breil et al, 2021) (Böckler-Raettig, personal communication, 2022) resulting from the videos, participants were asked to complete an immediate self-report question of 'how do you feel' with feelings placed on a 5-point Likert scale from 'extremely negative' (0) to 'extremely positive' (4).

To ensure that participants were attending to the videos, I included two attention check questions immediately following the self-report. Each attention check was designed to account for visual and auditory attention, asking questions that can only be known from paying attention to the visual and auditory aspects of each video. For inclusion in the data, individuals were required to identify the actor in a forced choice task when presented with the congruent actor and a distractor agent (visual attention) and identify the name of the character each actor spoke about (daughter and partner – auditory). All participants passed the required attention checks.

Procedure

Following the provision of information sheet, participant consent and demographic information to use in further exploratory analyses, individuals were shown one of two the short videos. Immediately following the video shown, the empathy measure and attention check questions were provided. This procedure was then repeated for the second video. The order of the videos shown were counterbalanced across participants and group allocation of participants was randomised.

Statistical Analyses

Affect rating was recorded as the score from a 5-point Likert scale of feeling ranging from 'extremely negative' (0) to 'extremely positive' (4). Using this Likert scale, emotional valence was detected using scores of 0 and 1 as negative, 2 neutral and 3 and 4 positive. Affect strength, on the other hand, was scored in terms of a reported score's distance from the neutral 'neither positive nor negative' point on the scale (score of 2). Thus, affect strength was calculated as the absolute value of (2 minus affect rating).This approach to recording affect strength allowed us to measure the rating's distance from neutral, independent of valence directionality. For example, affect ratings of 0 (extremely negative) and 4 (extremely positive) produce the same affect strength (2).

To investigate whether the video conditions had a significant effect on participant affect strength and ratings, linear mixed model analyses were implemented using the lemeur() function in the r package "lme4". With the primary test predictor of condition (empathy vs neutral), two linear mixed models were created either predicting affect rating or affect strength with both models including participant ID as a random effect. For each model a likelihood ratio test was run against a null model (each predictor variable only predicted by the random effect (test predictor of condition removed) as a significant result signifies that each model is not producing results as it would in the null equivalent.

These two models were used to look at the effect condition had on affect rating and affect strength. As elaborated prior, it was important to investigate whether the empathy induction methods (both videos) are truly effective and successful. If so, I predicted that the empathy condition significantly results in (a) greater affect strength and (b) more negative reported valence.

Demographic information from participants were also recorded to be added as supplementary analyses. Specifically, gender and ethnicity were considered on the basis that some existing literature has suggested experienced empathy can depend on social categorization (Tarrant et al., 2009) and that there may be instances of in-group biases (Cikara et al., 2011). Demographic information was not considered to be a primary test predictor however, as more recently, sex-based differences in empathy are not universally considered (Baez et al., 2017)

Results

Descriptive Data – Affect induction

All participants reported greater negative affect following watching the emotive videos (Mean Emo=0.98 ± 0.57) than the neutral (Mean Emo=2.34 ± 0.53). This was again the case when broken down across the two video sets: Katie's Move (Mean Emo=1.05 ± SD=0.67 / Mean Neut=2.10 ± 0.45) and Rosie's Pond (MeanEmo=0.86 ± 0.48 / Mean Neut=2.57 ± 0.51). See Figure 3 for a violin plot.

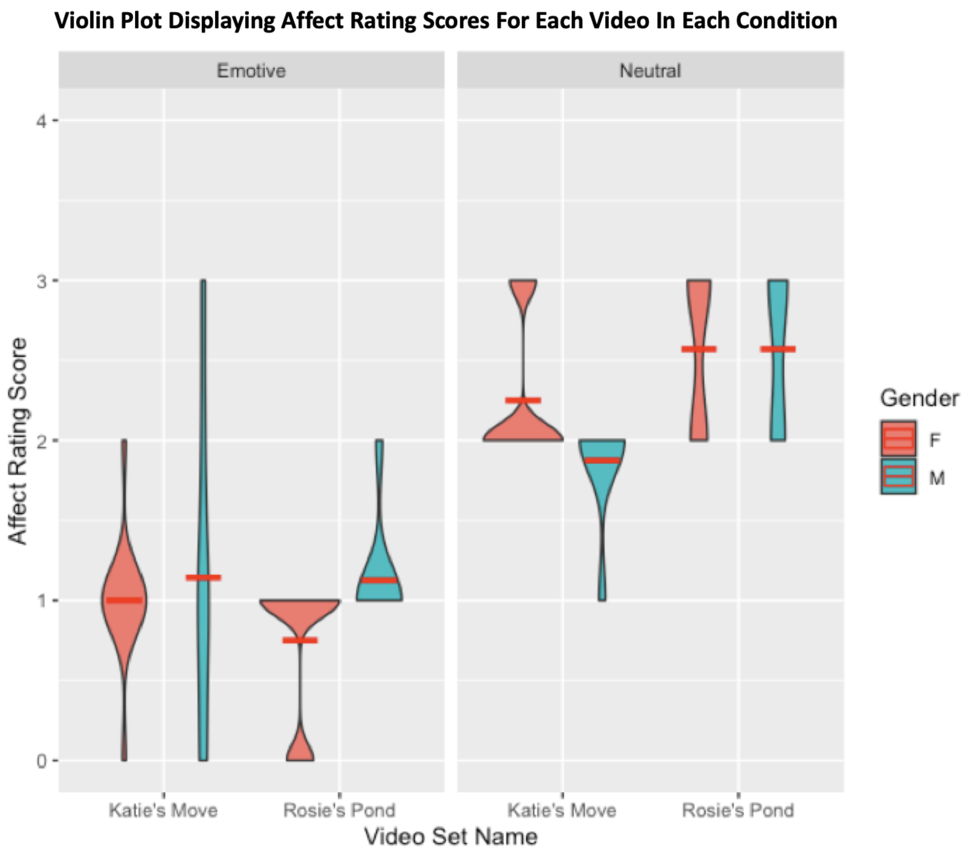


Figure 3: A violin plot representing the affect ratings participants assigned to the different short films watched. Data has been split by condition and separate plots assigned for participant gender. The red horizontal bars represent the mean values. The width of each plot represents the number of responses at that given value. With the y axis representing the 5 points on the likert scale, the lower values indicate more negative affect whereas the higher is more positive.

Controlling for valence directionality, participants reported stronger affect strength for the empathy inducing videos (M=1.07 ± 0.45) than the neutral (M=0.39 ± 0.49). This finding was found following the separation of video sets: Katie's Move (MeanEmo=1.05 ± 0.5/MeanNeut=0.2 ± 0.41) and Rosie's Pond (MeanEmo=1.1 ± 0.45/MeanNeut=0.57 ± 0.51).

<u>Inferential Data</u>

Regarding *Affect Rating*, stimuli condition significantly predicted reported affect (χ^2=77.739, d.f=1, p<0.001). A likelihood ratio test revealed the full model was significantly better at fitting the data compared to a respective null model only containing random effects (p=<0.001). Based on the above, I can conclude that video condition strongly influenced self-reported affect ratings - specifically that empathy videos resulted in more negative affect ratings. See Table 2 for test predictor breakdowns for both models of affect rating and strength.

Considering the supplementary analyses suggested, no significant interaction between gender and video condition were found when predicting affect rating scores. See Table 2 for estimate, standard errors, t values and p values of the main effect of condition on both affect strength and rating.

Similarly, the linear mixed model on *Affect Strength* found that the effect of condition was also significant. With a null model comparison demonstrating significant differences between the two models (p=<0.001), results suggest that the empathy inducing videos resulted in greater reported affect strength following viewings.

Follow-up supplementary analyses revealed no evidence of an interaction with either gender or ethnicity to condition (see appendix 1).

Table 2: Table reporting results from two linear mixed models on (a) affect rating and (b) affect strength using the lmer() function in r. In two separate models, condition was used to predict rating and strength. In these models, condition was the test predictor with the neutral stimuli as its reference group. Bold p-values signalling significance.

(a) Affect Rating				
Factor	*Estimate*	*S.E*	*t value*	*p-value*
Intercept	0.976	0.085	11.500	
Condition(Neutral)	1.366	0.120	11.380	**<0.001**

(b) Affect Strength				
Factor	*Estimate*	*S.E*	*t value*	*p-value*

Intercept	1.073	0.074	14.453	
Condition(Neutral)	-0.683	0.094	-7.259	**<0.001**

Discussion

The completion of the validation experiment returned two key conclusions. Firstly, the empathy manipulations were successful in inducing affect sharing in participants. Secondly, these results suggest that these emotive videos are not only promoting changes in affect valence (specifically negative affect), but that they are also promoting greater affect strength too. These results are discussed further below:

The analyses highlight that the emotive videos consistently resulted in lower reported affect rating scores when compared to their neutral counterparts (reflecting more negative valence). With the emotive videos inducing greater states of negative affect and the neutral videos staying at neutral or above, I can conclude that the designed manipulation was successful in putting the participant in an affect sharing state, empathising with the *victims* of the videos.

In addition to recording reported affect valence, I also examined reported affect strength of the video stimuli. Findings indicate that emotive videos induced a more intense affective response as compared to the neutral videos. This is important given that empathy is not required to be negative in nature and empathy can be positive (Morelli, Lieberman, Telzer, & Zaki as cited by Morelli et al., 2015). Considering the affect strength in responses to each video will help account for possible instances where participants may view the neutral videos as ‘somewhat positive’, as found here for the neutral *Rosie’s Pond* video. As shown in Figure 3, the affect rating score given to the neutral *Rosie’s Pond* video was seen to be diverging from a true neutral value. Considering the points raised above, it is possible that participants experienced a positive valence change through its message, however, we can control for this thanks to the record of affect strength to look for any unwanted effect.

Results from the linear mixed model highlighted no significant interaction between the test predictor of condition with gender or ethnicity (see appendix 1).

Experiment 2: Effect of Empathy on spontaneous visual perspective taking

Following the successful manipulation validation, I subsequently implemented these videos within a modified iteration of the L1-VPT, the Dot Perspective Task. This second experiment sought to directly investigate the effects of targeted empathy on spontaneous perspective-taking performance, specifically through examining rates of intrusion error.

Participants

All participants were Durham University students recruited online through the university SONA system, and tested online through Qualtrics and Psychopy/Pavlovia.

To be included in the pilot or test samples, participants needed to pass simple attention check questions following each of the trial blocks, relating to the videos watched prior to the trials themselves. Following past studies (Langton, 2018; Fan, 2021), participants were also required to perform correctly on at least 60% of trials in each condition and at least 70% of trials across both. Individual responses were removed if they were +/- 2.5 standard deviations from the grand mean (as in Fan, 2021), or if they fell beyond the 2 second maximum response window per trial (as in all dot-perspective studies, e.g., Samson et al., 2010).

Pilot Sample

We collected a pilot sample of 19 participants, three of whom were excluded (one was excluded for failing the attention check, and two others were excluded for failing to meet the accuracy criteria), leaving a final pilot sample of 16 (15 women, mean age = 20.327y ± 0.62 , age range = 20-21y).

Main Experiment Sample

In line with sample sizes from previous research employing a within-subject Dot Perspective Task design (Simpson & Todd, 2017; Todd et al., 2011), the study was pre-registered with a target sample size of 60 for the main experiment (https://aspredicted.org/ZDX_QDB). I was able to collect a sample of 41 participations ahead of the thesis due date, 11 of whom were excluded (2 was excluded for failing the attention check, and 9 others were excluded for failing to meet the accuracy criteria), leaving a final pilot sample of 30 (26 women, mean age

= 20.69 ± 0.77 , age range = 19-23). After the pilot experiment, I ran a post hoc power analysis using the powerSim() and powerCurve() functions in r (simr package) to determine whether the pilot sample size, which was based on previous experiments (Simpson & Todd, 2017; Todd et al., 2011), had sufficient power to detect a significant three way interaction. The function compares two models; the first is a complete model, with all the possible two way interactions, main effects and three way interaction alongside fixed effects of trial number and random effect of participant ID; the second is the same model but without the three way interaction). This analysis was used to test how many participants would be required in the main study to generate sufficient power to detect the presence of a significant three way interaction between perspective, consistency and condition. An observed power analysis using powerSim() and powerCurve() functions, drew upon the obtained sample size and effect size from the current experiment and revealed that a sample size of 9 would have been required to obtain a power value of 1-β = 80%. Consequently, this power analysis showed that the general linear mixed models obtained in both the pilot (n=16) and main experiment (n=30) had sufficient sample sizes to achieve power greater than 0.8.

Design

This second experiment implemented a within-participant design conducted online. Participants completed DPT trial tasks in Pavlovia whilst all pre- and post-trial surveys (and video exposure) occurred in Qualtrics. Trials were created using Psychopy before being made online accessible through Pavlovia. My primary analyses (see section on analyses) investigated the effects of condition (empathy vs baseline) x perspective (Self vs Other) x consistency of perspective (consistent vs inconsistent) on reaction time, controlling for trial number (fixed effect) and participant ID (random effect).

Stimuli & Procedure

To explore the effects of targeted state empathy on spontaneous perspective performances, I combined the DPT paradigm with the empathy manipulations validated already in experiment 1 (see Figure 4). This resulted in participants being asked to make rapid judgments about their own perspective and that of another individual (avatar) within a virtual space in an empathy and neutral baseline condition. I created two conditions to be sat consecutively: the emotive condition saw individuals watch an emotive video from experiment 1 video set 1 or 2 (randomised assignment) before completing a block of DPT trials; conversely, the baseline condition replacing the emotive video with a newly introduced control video of bubbles used in various eye tracking and thermal imaging research (e.g., (Hepach et al., 2012). Condition order was counterbalanced across the sample. To ensure that the videos were watched closely, an attention check was included at the end of each block of trials for both conditions that were specifically relevant to the video watched in that condition.

In addition to the videos, stimuli used for the DPT trials largely reflects that of the original procedure of the room scene images and avatars, but with slight alterations. Firstly, the room scene images are a combination of the original computer-generated half-room layout designs (Samson et al, 2010; stimuli obtained via open access page: https://figshare.com/articles/dataset/Level_1_Visual_Perspective_Taking_Task/1455943) and a changing number of red dots on the left and/or right wall ranging from 1 to 3 (Surtees & Apperly, 2012). Although the room scene images themselves were kept the same as Samson and colleagues' original stimuli, I used images of real individuals as the avatar present. The avatar is a centrally located being (typically computer-generated) that is placed in the room scene images and can face either leftward or rightward in profile. To ensure effects of empathy induced by the videos are targeted specifically toward that actor/character, the actor present in the emotive condition was used as the avatar for the block of DPT trails that immediately follow. In the baseline condition, the DPT trials included a novel individual posing as the avatar. All avatar photos were taken using an iPhone XR and edited using Procreate to then be combined with room scenes in Adobe Photoshop (see Figure 4).

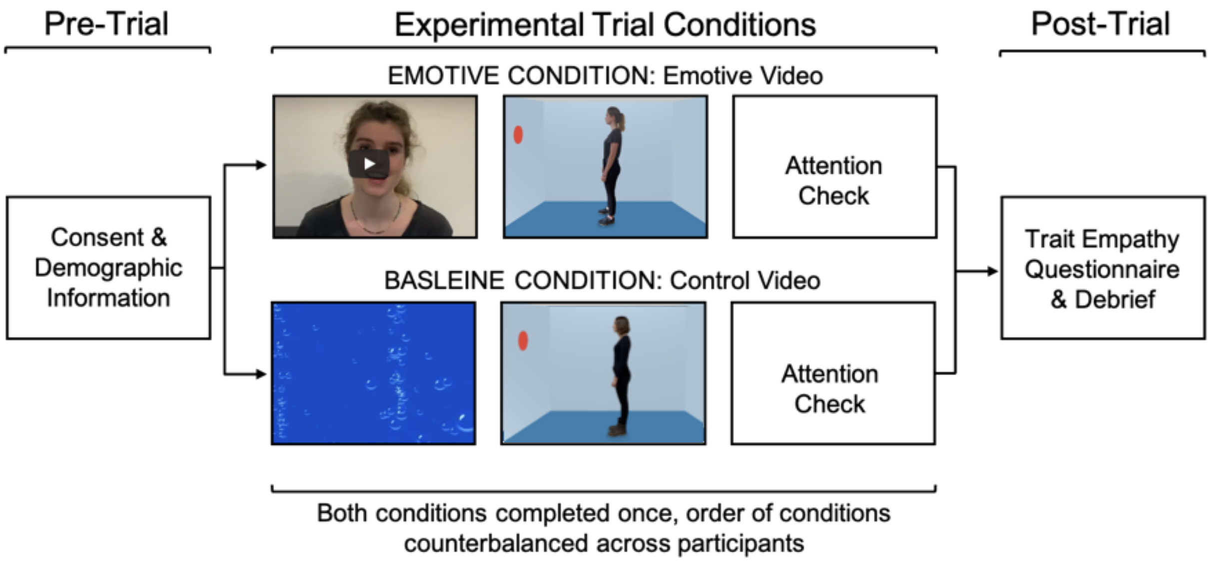


Figure 4: Flow diagram outlining the procedure for experiment 2. After completing a pre-trial demographic questionnaire, participants completed both conditions in a counterbalanced order. This involved watching a short video clip and then completing a block of DPT trials before an attention check question.

Similar to Samson and colleagues' original procedure, 50% of trials required participants to refer to their own perspective when formulating judgements (Self Trials), whereas the other 50% required to consider that of the avatar (Other Trials). Trials were further equally divided considering whether the perspectives of the participant and avatar were the same (Consistent Trials) or whether they differ (Inconsistent Trials).

Every trial followed the same procedure, starting with a fixation cross (750ms) followed 500ms break and then a slide displaying either "You" (Self Trials) or "She" (Other Trials) representing whose perspective should be taken into account for that trial (750ms). After a further 500ms, a number between 1 and 3 was shown (Surtees & Apperly, 2012) (750ms). Following a final 500ms, the room scene was shown for a maximum of 2000ms or until an answer was submitted. Participants had to provide an 'incorrect' or 'correct' response within said 2000ms to determine whether the picture matched the perspective and number shown ('yes' response – Matching Trials) or whether it is not an accurate reflection ('no' response – Mismatching Trials). See figure 5 for a visual representation of trial types.

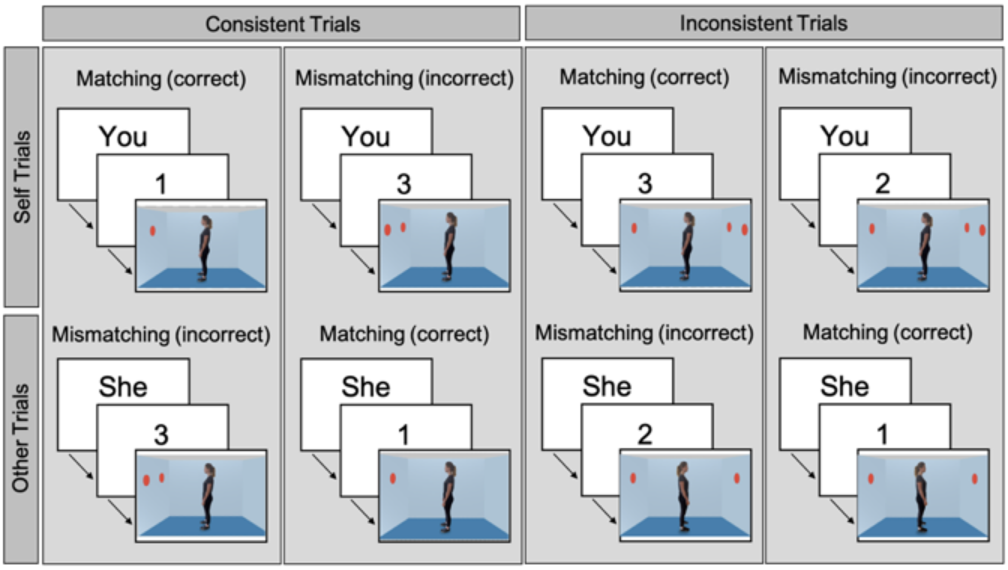


Figure 5: Dot Perspective Task trial types across trial variables: perspective (self vs other), consistency (consistent vs inconsistent) and matching (matching vs mismatching). These trial types are displayed in the style as seen in Fan (2021) to further illustrate the order of content shown within each trial; the participant is first shown a word dictating which perspective they should compute (You vs She), then a number corresponding to the number of dots visible from that perspective in the room scene (1-3). Finally, the participant is shown the room scene, and must register whether or not the previously depicted number matches the number of dots that are visible from whichever perspective they have been instructed to take.

Finally, following the completion of both conditions, participants were provided with an empathy questionnaire (Interpersonal Reactivity Index: IRI). The IRI questionnaire specifically is used to measure dispositional empathy (Davis, 1983; Krol & Bartz, 2022; Davis, 1980) and will be used within a supplementary analysis to further investigate whether trait empathy could at all exacerbate or dampen the results and effects seen across and between trial condition performances.

As mentioned earlier, I introduced some new procedural changes and design considerations not previously seen in Samson's original DPT. These can be categorised in the following subsections, with explanations provided:

Number of Trials. Due to its online administration, there were concerns that the procedure's cognitively demanding nature may result in lack of attention. Whilst there have been few studies to be successful (Vestner et al, 2022), several online DPT versions have failed to replicate original findings (Marshall et al., 2018; Rubio-Fernandez et al., 2022) which could partly be attributed to high cognitive demand. Surtees and Apperly (2012) adapted the DPT paradigm to investigate the presence of intrusion errors in adults and children. Researchers made several distinctive adaptations including the change of stimuli, avatars, positioning and timings, however, one of the main differences I drew from the design is the quantity of trials. In Samson's original procedure participants sat a total of 208 trials over 4 blocks, however, given the concerns regarding cognitive demand required for this many trials being too great for the empathic effect to be consistent for an entire block, I instead looked toward Surtees and Apperly. Researchers showed children 48 trials in total, whereas adults completed two blocks of 48 trials (96 trials total). I built on these findings from Surtees and Apperly (2012) to present participants with one block of 48 for each condition (total of 96 trials total). These two blocks of 48 trials were pseudo randomised (see Samson et al, 2010).

Combination of Human Avatars/Agents and CG Backgrounds. To maximise the effect of the empathy manipulation I wanted to keep the avatar in the DPT consistent with the individual shown in the video (directed empathy). With this in mind I could either animate the empathic video so the avatar is CG in both, or within the DPT. I estimated that incorporating human images was more likely to induce an affective state when the videos included real people. Although most iterations of the DPT have utilised CG avatars (i.e., Samson et al, 2010; Marshall, Gollwitzer & Santos, 2018; Drayton, Santos & Baskin-Sommers, 2018) or CG semi-social stimuli (Nielsen et al., 2015; Santiesteban et al., 2014), the use of real human images have produced interesting mixed results. Whilst some have found no evidence of consistency effects using this alternative design approach (Langton, 2018), others have demonstrated that the same effects can be achieved (Fan, 2021; Vestner et al, 2022). See Figure 6 for a brief table of comparison of avatar type examples.

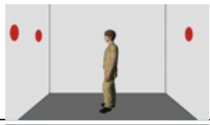
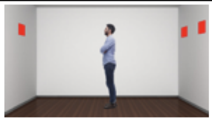
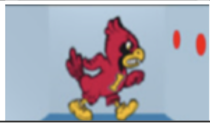
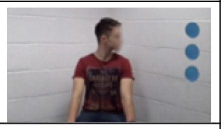
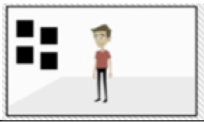
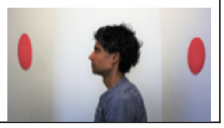
Dot Perspective Task Stimuli Examples			
Computer Generated		Real Images	
Drayton, Santos & Baskin-Sommers (2018)		Vestner et al (2022)	
Simpson & Todd (2017)		Langton (2018)	
Marshall, Gollwitzer & Santos (2018)		Rubio-Fernandez et al (2022)	

Figure 6: Figure table displaying several iterations of the Dot Perspective Task that implement computer generated avatars vs those who use real images.

Statistical Analyses

In line with previous executions of the DPT, only matching trials were used for analyses (Samson et al, 2010). Mismatching trials were excluded from analyses because some of the trial configurations including mismatching perspectives result in salient differences between some of the other trials in terms of required computations. Specifically, as noted by Simpson and Todd (2017), consistent-mismatch trials do not map on to either the participant or the avatar's perspective making it incredibly easy to compute. Whilst kept for accuracy rate analyses, erroneous responses and trials with a reaction time of greater than 2 seconds will not be included in reaction time analyses (Samson et al., 2010; Todd & Simpson, 2016).

To examine the effects of empathy on DPT performances, I ran a linear mixed effects model in r using the lmer() function in the "lme4" package. As the primary analysis, a model was run with response/reaction time (ms) as the dependent measure with participant ID set as the random effect to control for repeated contributions. For fixed effects, condition (empathy vs baseline), perspective (self vs other) and consistency (consistent vs inconsistent) were test predictors with a control predictor / fixed effect of trial number. I created a model including a three way interaction between these test predictors, however, if I obtained non-significant interactions, it would be removed and broken down into two way interactions and re-run. Since interaction terms can obscure main effects, if two way interactions returned insignificant results, these interactions would be removed entirely and a model run to include no interactions, only main effects.

A control model was then created that saw reaction times predicted by only the random and control effects. A likelihood ratio test was subsequently run to compare the full model with the null model to determine whether the combination of test predictors explained the data significantly better than a null model. Finally, the data was split into two separate data sets, one containing all *self* perspective trials and one for *other* trials. The models described above were run again on these two data sets but with 'perspective' removed as a test predictor. By splitting the data as such, I could examine each intrusion effect type - focusing on egocentric intrusion in the 'other' data set and altercentric intrusion in the 'self' trials.

Following full model analyses, all participant data was split into two distinct data sets, one containing only *other* trials and the second only *self* trials. Such a breakdown approach allowed for deeper investigation into the three way interaction obtained to see where the effect is being driven.

Recalling the proposed hypotheses surrounding how both self-other distinction and merging sub-components of empathy would influence perspective-taking, I used these models to investigate alter- and egocentric intrusion effect rates. If the empathy manipulation elevates self/other distinction, I predict slower responses on consistent trials involving judgements about the self in the empathy condition relative to baseline, because the empathy induction makes the avatar's perspective salient, even when perspectives do not conflict. Conversely, inconsistent trials may see faster responses in the empathy condition relative to baseline, because the salience of the avatar's perspective makes it easier to disambiguate perspectives. If the empathy manipulation instead elicits self/other merging, I predict stronger egocentric and altercentric intrusion effects in the empathy condition relative to the baseline, manifested in faster responses to consistent trials (given no need to disambiguate perspectives) and/or slower responses to inconsistent trials (given challenges in disambiguating perspectives).

In addition to the primary analyses into reaction times, I conducted supplementary analyses into accuracy rates (correct vs incorrect). The same models were used (same test predictors, interactions, fixed effects and random effects) as well as control models for comparison analyses. Given the incredibly high success rates recorded in previous literature ((Rubio-Fernandez et al., 2022; Vestner et al., 2022), intrusion effects have not been typically seen through analyses into accuracy rates, however, these models were included as supplementary tests.

Results

Pilot Results

Reaction Time

A likelihood ratio test revealed the full model explained the data significantly better than a null model containing only random effects and the intercept (p=<0.001). A linear mixed model on reaction time revealed a significant three way interaction between test predictors condition, perspective and consistency (*t*=-3.242, p=0.001) with 95% CI [-0.313,-0.077]. See Table 3 for the model's breakdown of estimate, standard error, confidence intervals and p value for the three way interaction. Figure 7 shows a simple effects plot displaying interactions between the reaction time values for each trial type factor of condition, perspective and consistency.

Table 3: Table summarising the output of the lmer() model of reaction time predicted by condition, perspective and consistency with a significant three way interaction included. For each of the test predictors, empathy, self and inconsistent were the reference groups of their respective variables. Bold p-values signalling significance.

Reaction Time - 3 way interaction						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.699	0.050	13.961		0.597	0.800
Condition(Emp)	-0.026	0.030	-0.884		-0.085	0.032
Perspective(Self)	-0.0002	0.030	-0.009		-0.060	0.060
Consistency(Incon)	0.061	0.031	1.986		0.001	0.120
Trial Number	-0.0003	0.001	-0.477	0.633	-0.001	0.001
Condition*Perspective *Consistency	-0.195	0.060	-3.242	**0.001**	-0.313	-0.077

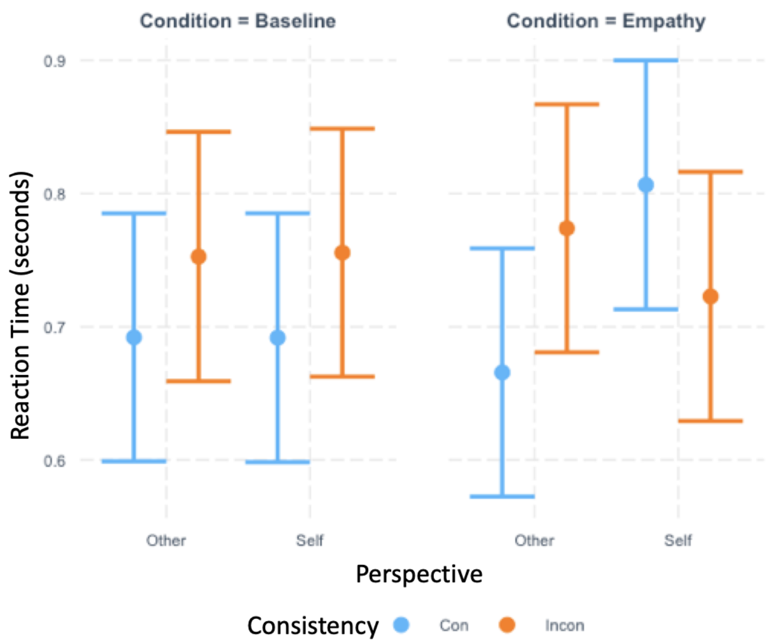


Figure 7: A simple effects plot showing the relationship between categorical variables of condition (baseline vs empathy), perspective (self vs empathy) and consistency (consistent vs inconsistent) against the numerical value of reaction time. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

In addition to the models described above, other models were run to control for different factors that may have contributed to the significant three way interaction result. Firstly, the same model was run as the first used, however it replaced condition with ‘experiment’ as a test predictor (4 experiments each using a one video in the emotion condition). Secondly, I considered ‘actress’ as a test predictor in place of condition. Results from both linear mixed models found no significant interaction (three way or two way) for any pairings, or any significant main effect findings. As a result of these non-significant findings and no changes were made to the design between the pilot and main experiment, I did not include any of these variables in the planned models for the main experiment.

Perspective Breakdown: Self

A likelihood ratio test revealed that this self model was significantly different from a null model (p=0.001). For *self* trials, a model including condition and consistency as test predictors (perspective removed as all trials are of the same perspective type) reported a significant two way interaction between the two (*t*=-3.501, p=0.001) with 95% CI [-0.238,-0.0670]. See Table 4 for a table breakdown of the two way interaction and the main effects reported. Figure 8 shows a simple effects plot displaying interactions between the reaction time values for each trial type factor of condition, perspective and consistency.

Table 4: Table summarising the output of the lmer() model of reaction time for only the *self* trials with test predictors of conditions and consistency. This model included a two way interaction with empathy as the reference group for condition and inconsistent trials as the reference group for consistency.Bold p-values signalling significance.

Reaction Time - 2 way interaction (SELF TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.695	0.050	13.763		0.593	0.797
Condition(Emp)	0.118	0.031	3.762		0.057	0.180
Consistency(Incon)	0.065	0.032	2.067		0.003	0.128
Trial Number	-0.0002	0.001	-0.284	0.777	-0.002	0.001
Perspective * Consistency	-0.153	0.044	-3.501	**0.001**	-0.238	-0.067

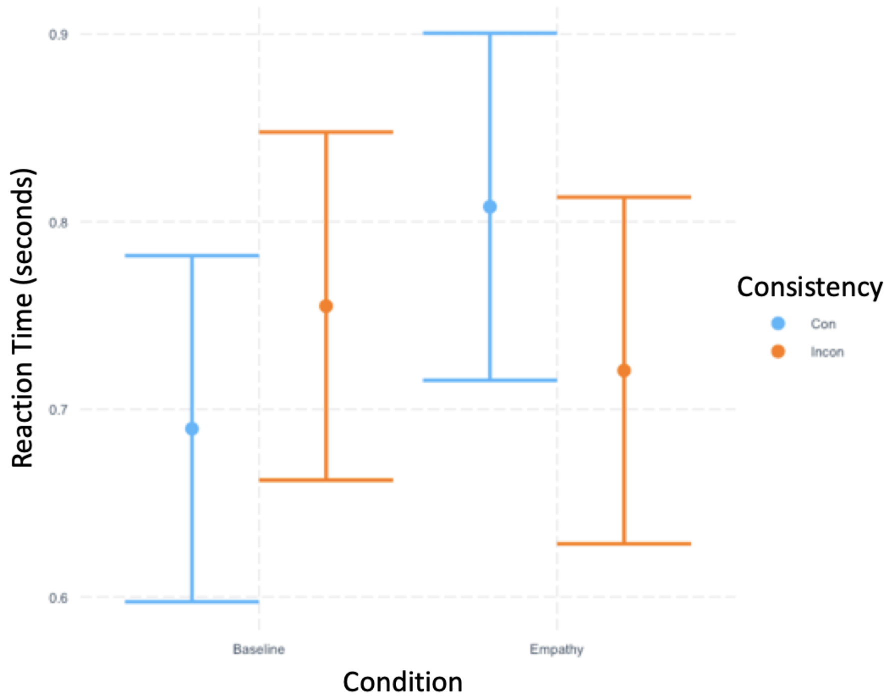


Figure 8: A simple effects plot showing the relationship between categorical variables of condition (baseline vs empathy) and consistency (consistent vs inconsistent) against the numerical value of reaction time using only the *self* trials. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

Perspective Breakdown: Other

For the *other* trials, the same approach was used - a model implemented including condition and consistency as test predictors (perspective removed as all trials are of the same perspective type). A likelihood ratio test showed the full model was significantly different to the null model (p=<0.001). Unlike the self model, the two way interaction was not significant. A subsequent reduced down model showed no significant main effect of condition (*t*=-0.219, p=0.827) with 95% CI [-0.045,0.036], but a significant result of consistency (*t*=4.008, p=<0.001) with 95% CI [0.043,0.125]. Once again, trial number was not significant (*t*=-0.440, p=0.660) with 95% CI [-0.002,0.001]. See Table 5 for a table breakdown of the main effects reported.

Table 5: Table summarising the output of the lmer() model of reaction time for only *other* trials with the predictor variables of condition and consistency. This model does not include any interactions as the two way interaction was found to be non-significant. Whilst the condition main effect was non-significant, consistency was. For each of the test predictors, empathy was the reference group for condition whereas inconsistent trials were the reference group for consistency. Bold p-values signalling significance.

Reaction Time - 2 way interaction (OTHER TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.691	0.052	13.352		0.585	0.796
Condition(Emp)	-0.005	0.021	-0.219	0.827	-0.045	0.036
Consistency(Incon)	0.084	0.021	4.008	**<0.001**	0.043	0.125
Trial Number	-0.0003	0.001	-0.440	0.660	-0.002	0.001

Accuracy Rate

Moving toward the more supplementary analyses, reaction time as a predicted variable was replaced by accuracy with its two levels: correct vs incorrect. This model was significantly different to a null model (p=<0.001). The linear mixed model reported significant three way interaction between test predictors of condition, perspective and consistency (*t*=-3.612, p=<0.001) with 95% CI [-0.435,-0.129]. See Table 6 for a breakdown of the model's interaction and a simple effects plot displaying interactions between the accuracy values for each trial type factor of condition, perspective and consistency in Figure 9.

Table 6: Table summarising the output of the lmer() model of accuracy rate predicted by condition, perspective and consistency with a significant three way interaction included. For

each of the test predictors, empathy, self and inconsistent were the reference groups of their respective variables. Bold p-values signalling significance.

Accuracy Rate - 3 way interaction						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.977	0.040	24.517		0.899	1.056
Condition(Emp)	-0.010	0.039	-0.256		-0.086	0.067
Perspective(Self)	-0.044	0.039	-1.116		-0.121	0.033
Consistency(Incon)	-0.210	0.039	-5.360		-0.287	-0.133
Trial Number	-0.0003	0.0007	-0.404	0.687	-0.002	0.001
Condition*Perspective *Consistency	-0.281	0.078	-3.612	**<0.001**	-0.434	-0.129

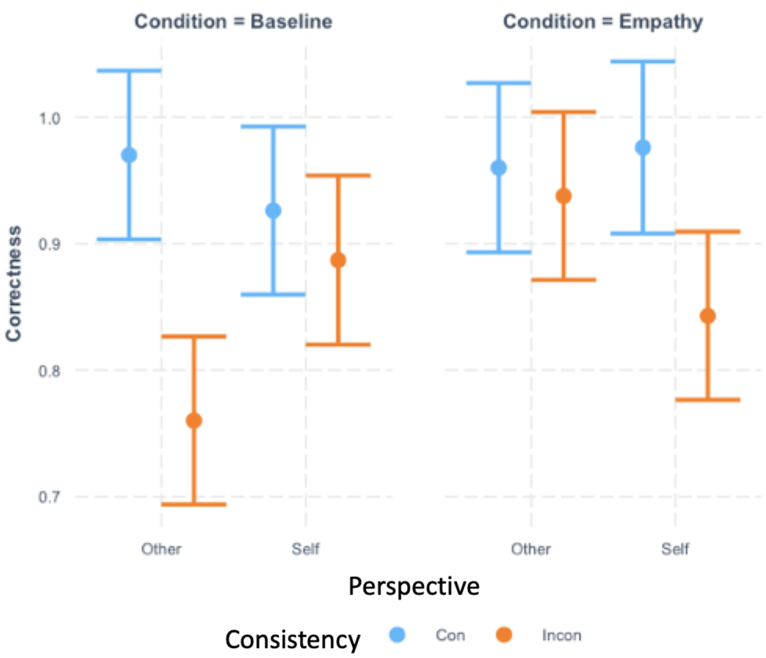


Figure 9: A simple effects plot showing the significant relationship between categorical variables of condition (baseline vs empathy), perspective (self vs empathy) and consistency (consistent vs inconsistent) against accuracy rates. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

Supplementary controls models were run to control for factors of experiment order and actress as additional test predictors in place of condition. Results from both linear mixed models returned no significant interactions (three or two way) in addition to any main effect findings.

Perspective Breakdown: Self

As seen for reaction time analyses, the data set was then split by perspective, forming self and other separate data frames. These data frames then became the basis for two separate models where accuracy is predicted by test predictors condition and consistency. The model for self was found to be significantly different to a null model without the test predictors present (p=0.022).For *self* trials, no interaction was found to be significant between condition and consistency, therefore main effects were reported from a model without the interaction included (see Table 7 for main effects breakdown of estimates, standard errors, p values and confidence intervals).

Table 7: Table summarising the output of the lmer() model of accuracy rate for only *self* trials with predictor variables of condition and consistency. This model did not include any interaction between the two predictors as such an interaction was not significant. Only the main effect of consistency was reported to be significant here. For the test predictors of condition, empathy was the reference group, whereas the reference group for consistency was inconsistent trials. Bold p-values signalling significance.

Accuracy Rate - Main effects (SELF TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.975	0.041	23.946		0.894	1.056
Condition(Emp)	-0.004	0.029	-0.130	0.896	-0.060	0.053
Consistency(Incon)	-0.079	0.029	-2.754	**0.006**	-0.136	-0.023
Trial Number	-0.001	0.001	-1.054	0.292	-0.003	0.001

Perspective Breakdown: Other

For *other* trials, the full model was significantly different from a null model made for *other* trials specifically (p=<0.001) (see table 8 for model output including the recorded significant two way interaction). A two way interaction between condition and consistency was significant (*t*=3.366, p=0.001) with 95% CI [0.078,0.297] meaning specifically that inconsistent trials in the baseline condition saw significantly more errors, driving this observed two way interaction.

Table 8: Table summarising the output of the lmer() model of accuracy rate for only *other* trials predicted by condition and consistency. This model includes a significant two way interaction between the two test predictors. For each of the test predictors, empathy was the reference group for condition, and inconsistent trials the reference group for consistency variable. Bold p-values signalling significance.

Accuracy Rate - 2 way interaction (OTHER TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.957	0.042	22.560		0.874	1.041
Condition(Emp)	-0.011	0.039	-0.280		-0.088	0.066
Consistency(Incon)	-0.206	0.040	-5.188		-0.284	-0.128
Trial Number	0.0004	0.001	0.409	0.682	-0.002	0.002
Condition * Consistency	0.187	0.056	3.366	**0.001**	0.078	0.297

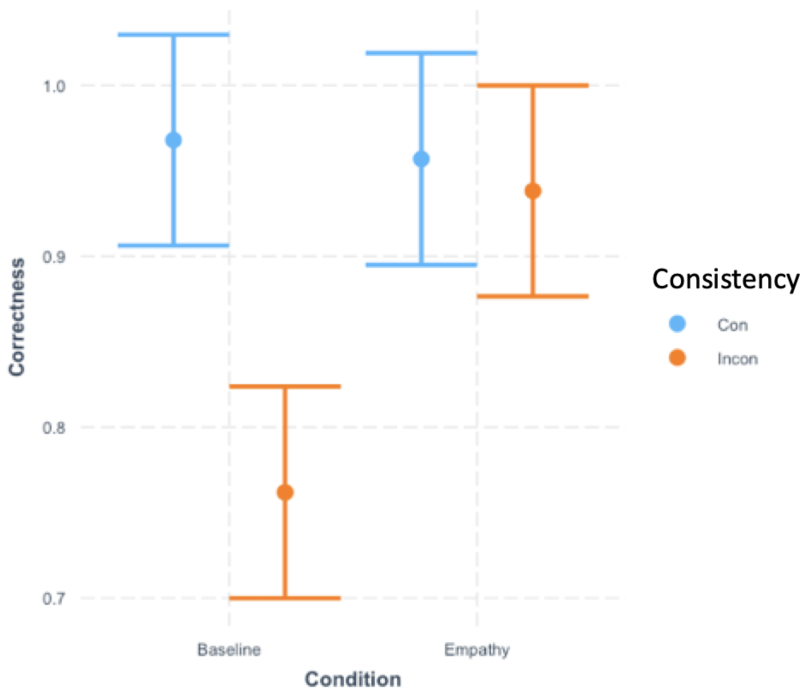


Figure 10: A simple effects plot showing the significant two-way interaction between categorical variables of condition (baseline vs empathy)and consistency (consistent vs inconsistent) against accuracy rates for only *other* trials. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

Pilot Study Discussion

Reaction Time:

Results from the pilot study revealed a significant three way interaction between condition, perspective and consistency, pointing to the empathy inductions having an impact on DPT performances. Specifically, a breakdown of the data revealed that the significant interaction is predominantly driven by *self* trials in the empathy condition (See Figure 8). The interaction alongside the catplot predominantly align with the prediction associated with *self-other distinction,* not *self-other* merging, whereby slower reaction times in the consistent trials and faster in the inconsistent trials resulted in an interesting inverse pattern compared to previous literature. This was solely in the empathy condition whilst the baseline condition replicated past literature (e.g., Samson et al, 2010; Marshall, Gollwitzer & Santos, 2018).

Extrapolating these suggestions to the perspective-taking literature in more detail, these findings can be thought of in different ways depending on each account. Firstly, from a mentalizing account (Furlanetto et al., 2016; Samson et al., 2010), these effects on reaction time could be interpreted to be the result of participant's empathic states penetrating their spontaneous computations of another's visual field. More precisely, *self-other distinction* influences implicit perspective taking performances by facilitating one's ability to respond rapidly to another's consistent perspective, whilst hindering speed of response of judgements of the self when perspectives are consistent. If a more domain-general, sub-mentalizing approach is taken, these results pose an additionally interesting proposition. As demonstrated in previous work with semi-social stimuli (e.g, Santiesteban et al, 2014), both intrusion effects can be observed without the presence of a human avatar because it is argued that the recorded effects are being driven thanks to lower-order attentional cueing processes where no mentalizing is involved (Heyes, 2014; Santiesteban et al., 2017). The experiment described in the current study cannot necessarily disambiguate between a mentalizing and sub-mentalizing account. If further research supports the sub-mentalizing account, it would be the case that the effects of empathy (and its subcomponents) may be able to influence a fundamental domain-general cognitive process whereby the orientation of perspective computation is essentially influenced.

<u>Accuracy Rate:</u>

Similarly to reaction time, the supplementary analyses into accuracy rates also yielded a significant interaction between perspective, consistency and condition. Unlike the primary model's findings however, the significant findings obtained here are somewhat misleading as they are driven by a surprising effect. When broken down by perspective, it was observed that the interaction was driven by strikingly low accuracy scores on *other* inconsistent trials in the baseline condition. This effect was not predicted by either hypothesis presented meaning that it may be an artefact of the online design. This is striking because error rate has been commonly excluded from DPT analyses as a measure (to focus on reaction time) due to frequently observed ceiling accuracy scores (Vestner et al., 2022). This is something addressed in greater detail within the general discussion. The fact that such poor accuracy was observed in the baseline condition and not in the empathy condition is suggestive of the possibility that this effect through error rate may be a result of a methodological constraint arisen from adapting the DPT paradigm for an online, within-subjects design rather than something rooted in the manipulations (addressed further in the general discussion).

<u>Summary</u>:

In conclusion, the primary analyses on reaction times demonstrate empathy's multifaceted effects on spontaneous perspective-taking performance. Through finding slower reaction times in consistent trials and faster in inconsistent trials in only the empathy condition, the results support the self-other distinction hypothesis of empathy. As a result, I obtained initial findings that support my proposal that the *self-other* distinction subcomponent of empathy may be influencing spontaneous perspective-taking. Whilst reaction time data returned fruitful, accuracy rates were not and to an unexpected degree. Not only was no significant effect of the empathy found, a three way interaction solely driven by uncharacteristically poor performance in the inconsistent *other* trials for the baseline condition (see general discussion for further elaboration).

Main Experiment Results

Reaction Time

Firstly, a likelihood ratio test revealed that the full model explained the data significantly better than a null model containing only random effects and the intercept (p=<0.001). Unlike in the pilot study, the linear mixed model on reaction time did not reveal a significant three way interaction. The model was then reduced to its two way interactions between test predictors condition, perspective and consistency with only the perspective and consistency interaction returning significant ((t=-3.976, p=<0.001) with 95% CI [-0.164,-0.056]) so condition was removed from the interactions and kept as a main effect. See Table 9 for the model's breakdown of estimate, standard error, confidence intervals and p value for each of the two way interactions, including a significant effect of trial number on results (t=-2.771, p=0.006) with 95% CI [-0.002,-0.0004]. See Figure 11 for a simple effects plot displaying interactions between the reaction time values for each trial type factor of perspective and consistency.

Table 9: Table summarising the output of the lmer() model of reaction time predicted by an interaction between perspective and consistency with condition as a main effect. The model used here incorporated a two way interaction between perspective and consistency test predictors and condition left as a main effect. For each of the test predictors, empathy, self and inconsistent were the reference groups of their respective variables. Only the two way interaction between perspective and consistency was found to be significant. This also shows a significant effect of trial number. Bold p-values signalling significance.

Reaction Time - 2 way interaction						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.727	0.032	22.551		0.663	0.7907984145
Condition(Emp)	-0.019	0.014	-1.376	0.169	-0.045	0.007
Perspective(Self)	0.087	0.020	4.393		0.048	0.126
Consistency(Incon)	0.101	0.019	5.279		0.064	0.139
Trial Number	-0.001	0.0005	-2.771	**0.006**	-0.002	-0.0004
Perspective * Consistency	-0.110	0.028	-3.976	**0.000**	-0.164	-0.056

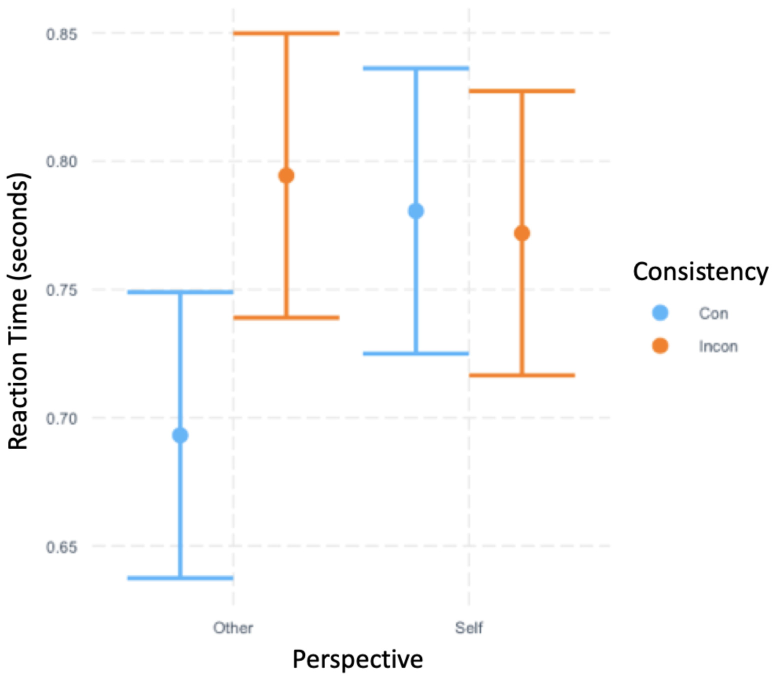


Figure 11: A simple effects plot showing the significant two-way interaction between categorical variables of perspective (self vs empathy) and consistency (consistent vs inconsistent) against the numerical value of reaction time. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

Perspective Breakdown: Self

In order to break down the results, I first ran linear models on the *self* data set. A likelihood ratio test revealed that this self model was not significantly different from a null model that predicted reaction time from only the fixed and random effect and no test predictors ($p=0.7681$). This non-significant difference was reflected through the linear mixed model run that included condition and consistency as test predictors (perspective removed as all trials are of the same perspective type) that reported a non-significant two way interaction in addition to no significant main effects when the interactions were removed. See Table 11 for a table breakdown of the two way interaction and the main effects reported.

Table 10: Table summarising the output of the lmer() model of reaction time for only *self* trials predicted by condition and consistency. This model did not include any interactions as the two way interaction was found to be non-significant. The reference groups for condition and consistency were empathy and inconsistent (respectively). No main effects were reported as significant. Bold p-values signalling significance.

Reaction Time - Main Effects (SELF TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.792	0.034	23.288		0.725	0.860
Condition(Emp)	0.009	0.020	0.461	0.645	-0.030	0.048
Consistency(Incon)	-0.012	0.020	-0.598	0.550	-0.051	0.027
Trial Number	-0.001	0.001	-1.247	0.212	-0.002	0.001

Perspective Breakdown: Other

Similar to the *self* trial breakdown, the same model was used for *other* trials. A likelihood ratio test revealed that this *other* model was significantly different from a null model (p=<0.001). Despite being significantly different from the null model, a linear mixed model conducted on the *other* trials, including condition and consistency as test predictors (perspective removed as all trials are of the same perspective type) reported no significant two way interaction between them. As a result, the model was reduced to main effects where all predictors were significant. Findings suggest reaction times were slower in the inconsistent trials relative to the consistent, and were faster for the empathy condition relative to the baseline. See Table 12 for a table breakdown of the main effects reported.

Table 11: Table summarising the output of the lmer() model of reaction time for only *other* trials predicted by condition and consistency with a significant three way interaction included. For each of the test predictors, empathy, self and inconsistent were the reference groups of their respective variables. Both main effects and the fixed effect of trial number were found to be significant.

Reaction Time - Main Effects (SELF TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.741	0.035	21.440		0.672	0.809
Condition(Emp)	-0.042	0.019	-2.229	**0.026**	-0.080	-0.005
Consistency(Incon)	0.101	0.019	5.245	**>0.001**	0.063	0.139
Trial Number	-0.001	0.0007	-2.151	**0.032**	-0.003	-0.0001

Accuracy Rate

A likelihood ratio test revealed the following full model explained the data significantly better than a null model containing only random effects and the intercept (p=<0.001). A linear mixed model on accuracy rate revealed a significant three way interaction between test predictors condition, perspective and consistency (*t*=-2.261, p=0.024) with 95% CI [-0.268,-0.019]. See Table 13 for the model's breakdown of estimate, standard error, confidence intervals and p value for the three way interaction. Figure 12 shows a simple effects plot displaying interactions between the accuracy rate values for each trial type factor of condition, perspective and consistency.

Table 12: Table summarising the output of the lmer() model of accuracy rate predicted by condition, perspective and consistency with a significant three way interaction included. For each of the test predictors, empathy, self and inconsistent were the reference groups of their respective variables. Bold p-values signalling significance.

Accuracy Analyses - 3 way interaction						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.928	0.030	30.539		0.869	0.988
Condition(Emp)	0.037	0.032	1.179		-0.025	0.100
Perspective(Self)	-0.015	0.032	-0.470		-0.078	0.048
Consistency(Incon)	-0.195	0.032	-6.104		-0.258	-0.132

Trial Number	0.001	0.001	1.575	0.115	-0.0002	0.002
Condition * Perspective * Consistency	-0.144	0.064	-2.261	**0.024**	-0.268	-0.019

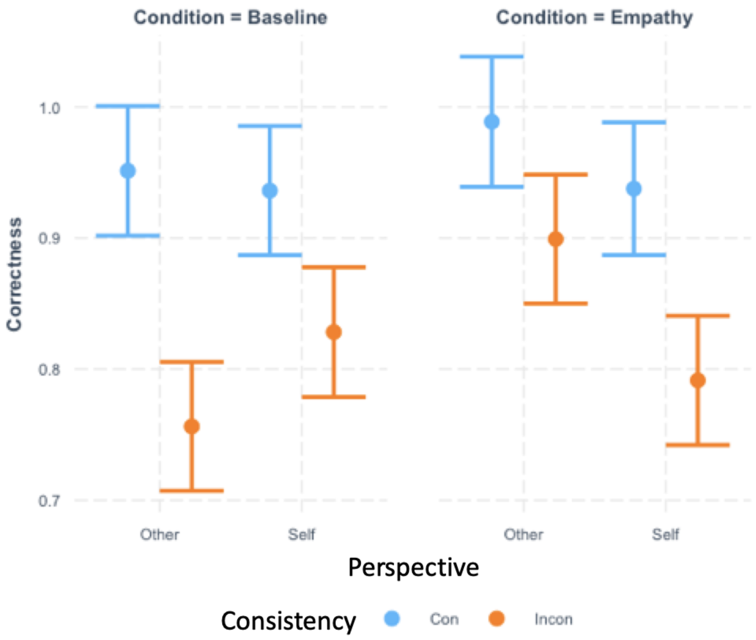


Figure 12: A simple effects plot showing the significant three-way interaction between categorical variables of condition (baseline vs empathy), perspective (self vs empathy) and consistency (consistent vs inconsistent) against accuracy rate values. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

Perspective Breakdown: Self

Breaking down the accuracy rate data set by perspective, a likelihood ratio test revealed that the self model was significantly different from a null model (p=<0.001). For these *self* trials, a model including condition and consistency as test predictors (perspective removed as all trials are of the same perspective type) reported no significant two way interaction between condition and consistency. Following analyses removing the interaction found only the main effect of consistency to be significant (*t*=-5.394, p=<0.001) with 95% CI [-0.180,-0.084].

Participants made more errors in the inconsistent condition than the consistent condition. See Table 14 for a table breakdown of the main effects reported.

Table 13: Table summarising the output of the lmer() model of accuracy rate for only *self* trials predicted by condition and consistency with no interaction included. The two way interaction was dropped from the model as it was found to be non-significant. For each of the test predictors, empathy was the reference group for condition whereas inconsistent trials were the reference group for consistency. Bold p-values signalling significance.

Accuracy Rate - Main Effects (SELF TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.911	0.034	26.557		0.843	0.978
Condition(Emp)	-0.013	0.024	-0.545	0.586	-0.061	0.034
Consistency(Incon)	-0.132	0.024	-5.394	**<0.001**	-0.180	-0.084
Trial Number	0.002	0.001	1.630	0.103	-0.0003	0.003

Perspective Breakdown: Other

Once again, a likelihood ratio test revealed that the following *other* model was significantly different from a null model (p=0.001). Setting aside *self* trials, a model on *other* trials, including condition and consistency as test predictors (perspective removed as all trials are of the same perspective type), revealed a significant two way interaction between the two (t=2.502, p=0.013) with 95% CI [0.023,0.188]. See Table 15 for a table breakdown of the two way interaction and the main effects reported. As is evident in Figure 13, both conditions yielded more errors in the inconsistent trials than consistent ones; however, this effect was more pronounced in the baseline.

Table 14: Table summarising the output of the lmer() model of accuracy rate for *other* trials predicted by condition and consistency with a significant two way interaction included. For both of the test predictors, empathy was the reference group for condition and inconsistent trials were the reference group for consistency. Bold p-values signalling significance.

Accuracy Rate - Main Effects (OTHER TRIALS)						
Factor	*Estimate*	*S.E*	*t value*	*p-value*	*CI: 2.5%*	*CI: 97.5%*
(Intercept)	0.941	0.031	30.781		0.881	1.001
Condition(Emp)	0.038	0.030	1.279		-0.020	0.097
Consistency(Incon)	-0.197	0.030	-6.560		-0.256	-0.138
Trial Number	0.0005	0.001	0.659	0.510	-0.001	0.002
Condition * Consistency	0.105	0.042	2.502	**0.013**	0.023	0.188

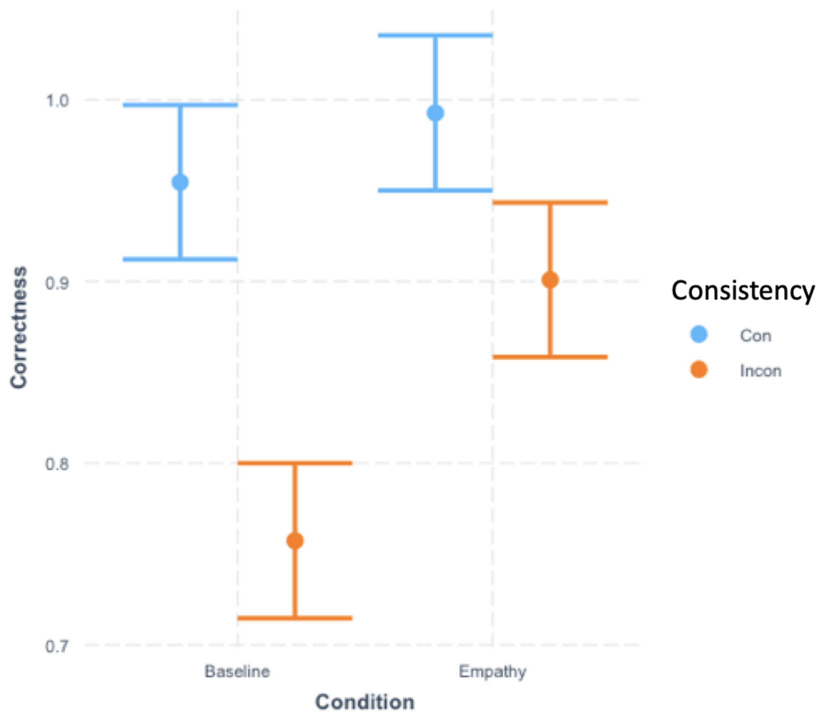


Figure 13: A simple effects plot showing the significant two-way interaction between categorical variables of perspective (self vs empathy) and consistency (consistent vs inconsistent) against accuracy rates for only *other* trials. Average values are represented with a dot, plotted alongside 95% confidence intervals. All plots were created using the catplot() function in r package

Main Experiment Discussion

Reaction Time:

Despite the promising results from the pilot study and no changes to the design, the main experiment analyses did not replicate the significant three way interaction between condition, consistency and perspective that was found in the pilot study. My results partially replicate findings in previous literature: I found evidence of egocentric intrusion (other trials) but not altercentric intrusion (self trials) when condition was collapsed. This effect was driven by *other* trials, suggesting the presence of egocentric intrusion, but not altercentric intrusion, regardless of condition. Furthermore, a significant effect of trial number was found, suggesting that reaction time responses became faster as the trials progressed. Contrary to the reaction time findings from the pilot experiment, the presence of egocentrism and absence of altercentric effects is not fully consistent with either the sub-mentalizing or mentalizing accounts of spontaneous perspective-taking.

Accuracy Rate:

My models examining accuracy rates subsequently revealed a significant three way interaction between condition, consistency and perspective. Overall, I found higher accuracy in the consistent trials than inconsistent trials across both *self* and *other* iterations - a finding replicated in numerous previous studies. However, for *other* trials, the error rate was greater in inconsistent baseline trials, resulting in more exaggerated egocentric intrusion in the baseline than the empathy condition. When compared to baseline for *other* trials, the accuracy of inconsistent trials was better in the empathy condition (whilst still being far lower at approx. 77% than numbers reported in past lab studies: e.g., 92% - Samson et al., 2010). Overall these results suggest a presence of both egocentric intrusion in the *other* trials and altercentric in the *self* trials, replicating to some extent previous literature (e.g., Samson et al., 2010).

Summary:

Overall, these findings provide evidence of egocentric intrusion in both reaction time and error rates. I also found evidence of altercentric intrusion but only in error rates. However, unlike the pilot study,in the main experiment, I did not find interaction effects involving condition that would provide evidence in support of either the self-other distinction or self-other merging hypotheses.

General Discussion

This study was designed to investigate the effects of empathy on human spontaneous perspective-taking, themselves being two fundamental social processes. I hypothesized that empathy's effects on perspective-taking could be driven by either a process of *self-other* merging or *self-other* distinction. In order to test these hypotheses, I first set out to create and validate a set of empathy-inducing stimuli in the form of novel videos (experiment 1). Following the manipulation validation, the stimuli were integrated within a modified online version of the *Dot Perspective Task* (Samson et al., 2010). Firstly, effects driven by self-other merging hypothesis would predict blurred distinctions between the self and the other, thus resulting in greater rates of both egocentric and altercentric intrusion errors in the empathy conditions as compared to baseline. Greater rates of both intrusion effects would manifest in terms of slower and more erroneous responses for inconsistent vs consistent trials for both self and other trial types when exposed to the empathy stimuli. Results for both the main experiment and its pilot failed to provide support for the self-other merging hypothesis. Statistical analyses in the main experiment revealed that empathy did not have a significant effect on either reaction time nor accuracy rate. Alternatively, the *self-other distinction hypothesis*, would predict that the differences between the self and the other would be more salient, thus resulting in slower and more erroneous responses for trials focused on the self where perspectives were consistent (self-consistent trials) specifically in the empathy condition. This was a novel prediction as it countered the findings from existing literature that reported altercentric intrusion occurring because of slower/poorer performances on inconsistent trials rather than consistent ones. This was the opposite to the self-other distinction hypothesis that suggested the empathy condition would see altercentric intrusion driven by slower/poorer performance in consistent trials compared to inconsistent ones. As previously mentioned, the main experiment reported no significant differences in performance between the empathy and baseline condition, therefore providing no clear evidence that empathy penetrates spontaneous perspective-taking abilities. Having said this, it is important to note that the pilot study for the main experiment indicated a different pattern of results. Within the pilot study, consistent with the self-other distinction hypothesis, reaction time analyses revealed that participants were significantly slower when making judgements about their own perspective when it was consistent to that of the avatar's in the empathy condition. Additionally, results from the baseline condition found the existence of both egocentric and altercentric intrusion errors, replicating existing findings in the literature.

Following the null results obtained from the main experiment but results supporting the self-other distinction hypothesis in the pilot, these experiments together can be argued to suggest that it is unlikely there is a strong relationship between two social processes, empathy and implicit perspective-taking. However, more work is required to make claims with greater internal validity. Although it would not be permissible to formulate a grand conclusion about empathy's influence on implicit perspective-taking performance from the findings obtained here, this study provides some important and novel considerations stemming from its limitations. These aspects are discussed further below.

After obtaining data from the pilot and main experiment two, three different conclusions can be drawn. Firstly, findings from the main experiment's pilot study supported the self-other distinction hypothesis, finding altercentric intrusion in the empathy condition driven by poor performance specifically on self-consistent trials. Secondly, the main experiment reported null results, concluding here that empathy had no significant effect on spontaneous perspective taking performance. Finally, I found evidence of both egocentric and altercentric intrusion effects in error rate analyses for the main experiment, replicating previous literature and therefore providing validation that this shortened, modified DPT was successful. Regarding the last, most prior studies report such low error rates that such effects are not found, however I may have found more pronounced error rates in this study due to its online format. Although error rates are not as typically documented, they are in line with the predictions of both egocentric and altercentric intrusion effects. One way to possibly account for finding variations is to look to possible methodological issues to investigate. An important factor to consider when it comes to the data I obtained links to the stimuli used in both experiments. Although experiment 1 strongly suggested that the stimuli were able to elicit more negative affect, there are three possibilities as to why no significant effect could be seen (relating to the stimuli). Firstly, the empathy manipulation I created may not have been effective in inducing empathy for the avatar but instead elicited a self-oriented affective state like emotion contagion or state matching, two self-oriented processes rather than an other-oriented one. As elaborated on in the empathy manipulation validation experiment, the decision to include measures of affect strength and ratings were informed through their prior and repeated application within the EmpaTom / EmpaToM-Y paradigms (Breil et al., 2021; Kanske et al., 2015). Despite the use of established measures, the nature of the EmpaToM and the videos used in my experiments differ as the former employs a battery of emotive videos. One possibility is that although the video I implemented was longer and focused on a single character, the use of multiple videos may be a better method of inducing a stronger (less individual-targeted) sense of empathy in participants. Moving forward, to strengthen my claim of induced empathy and the success of the inductions, more explicit measures of state empathy, i.e., the State Empathy Scale (Shen, 2010) could be provided to participants. If the additional measure consolidated the affect strength and ratings scores reported, conclusions can be made about the validity of the empathy induction with greater confidence. Secondly, no significant effect of condition in experiment 2 may be linked to the stimuli's lack of strength. A possible critique of the stimuli is that the approach did not elicit a pronounced enough effect to reliably penetrate the mechanisms underlying DPT performances (mentalizing, domain-general or both). A means of modifying the stimuli in such a way could be focused more at the group level. A possible shortcoming of the empathy inducing videos in the study could be that the stories were not relatable to the participants in the sample. With a mean age of 20, it is possible that participants may not have fully related to a mother's grief or the loss of a partner; in a follow-up, more emphasis should be applied to the relatability of the storylines to a study's target sample. Such an approach was utilised in the EmpaToM-Y paradigm for adolescents (Breil et al., 2021) where the researchers replaced traditional stories used in the EmpaToM with more age-appropriate plots like school life and peer drama. Finally it is important to mention that the stimuli may not have led to an effect on spontaneous perspective-taking, not through the fault of the videos themselves, but instead because empathy does not have an effect on perspective-taking on this level.

One of the largest possible sources of contention is associated with the study's methodologies. To begin, unlike most of the DPT iterations in the past, I converted experiment 2 into an online format. Some previous studies that have implemented an online approach have failed to replicate the findings of their lab-based counterparts (Marshall et al., 2018; Rubio-Fernandez et al., 2022) - failures in replication I predominantly attributed to the high cognitive demands of the paradigm. With this prior in mind, I adapted the design to a shortened version. However, it is possible that the issues remained. Unlike studies run in a lab setting, no experimenter was present to ensure participants completed the task. Moreover, there was notable evidence of lack of attention or application in my results, evident through several responses with 0% or 50% accuracy (rapidly pressing one key continuously, n=3). Nevertheless, following exclusions of inattention, the main experiment recorded 90% accuracy of those whose data went forward to be included in analyses. The surprisingly high accuracy rates are somewhat in line with other online DPT studies reporting 96-98% (e.g., Rubio-Fernandez et al, 2022; Vestner et al., 2022). To understand potential sources of variation, it would be important to run the experiment in its current state in a lab setting to consolidate whether poor reaction time performances are the result of online testing or simply the adapted design of the task.

Despite the study's limitations, the results from the pilot experiment (experiment 2) suggest that the adapted DPT design included here could be promising. Whilst designing the main experiment, I had concerns about the paradigm's demanding layout when it comes to testing online, but also when taken with the empathy induction. Typically asking participants to complete 208 total trials, the traditional DPT design presented concerns that any effect of state empathy introduced by watching the videos could deteriorate as the trial block(s) progressed. This led to considerations to use an already established altered approach by Surtees and Apperly (2012) where intrusion effects were found using only 96 trials across 2 blocks (for adults). Using this approach would have been sufficient to address the concerns of online testing (I had thought); however, I was concerned about the use of two blocks. If trials were split into 2 blocks the effect of empathy could have greatly reduced or even extinguished between the break. Using an in-person approach could be fruitful to explore the application of this paradigm alongside possibly stronger stimuli to observe if this is true of the effect. In my DPT experiment, I wanted to reduce this further to see if the effects could be observed and replicated using only 48 trials (in each condition), instead of 96. As the reaction time pilot experiment was successful and the main experiment replicated previous egocentric and to a partial extent altercentric findings, I believe that this design approach shows promise, perhaps especially in a lab setting. Future work should explore this possible success with more direct comparisons between the three DPT designs (208, 96 and 48 trials) to explicitly compare findings under the same conditions. This would be an important venture as if successful, a shortened DPT layout could allow more exploration into the relationship other processes (cognitive, social and affective) can have with DPT performances without having the concern of such effects reducing within the trial blocks.

In addition to discussion of these potential limitations, there are two further conceptual ideas that could be explored. Firstly, for the main experiment participants were asked to complete the Interpersonal Reactivity Index (IRI; Davis, 1983), a standardised measure of trait empathy. Despite not being able to use these results in this study, in future work, I will incorporate these scores into additional analyses to see if trait empathy had a mediating effect on the stimuli's influence on reaction time and/or accuracy rates. It may be the case that only those with high trait empathy are susceptible to the effects of empathy on spontaneous perspective-taking performance, suggesting that individual differences could have a highly influential impact on the results I found in the data I obtained. This finding would be consistent with work by Drayton, Santos and Baskin-Sommers (2018), showing individual differences in the manifestation of altercentric intrusion. A final possible future direction is centred around employing a comparative approach to this study's research questions. Following further investigation into whether empathy does have a direct effect on implicit perspective-taking, if found to be present, there would be reason to also study this question in nonhuman animals. From an evolutionary perspective, the potential automaticity of these foundational cognitive mechanisms points to the possibility that they are more evolutionarily ancient and could be shared with other nonhuman primates. Future work on implicit perspective taking capacities in nonhuman primates could be incorporated with the comparative work on empathy to further our understanding when it comes to the evolutionary path these processes may have developed down to be observed as they are in humans.

In summary, although results from this study ultimately did not support the hypotheses under investigation, the study presents several interesting considerations and contributes novel opportunities for future adapted approaches to implement the Dot Perspective Task to investigate implicit perspective-taking and its relation to empathy. Future recommendations should focus on the further validation and exploration into the empathy induction method, in-person testing of this shortened DPT paradigm and direct comparisons between each DPT methodology iteration. Finally, results highlight the crucial importance of replication in the scientific process. In this study, although the pilot obtained a very promising set of findings that supported the pre-registered predictions and hypotheses, this was not replicated in the main experiment despite no changes to the design and participant pool. The DPT as a paradigm investigates implicit mechanisms and although it is incredibly robust through replications, when it comes to adapting the design we have to be careful to test any adaptation's durability and validity.

Statement of Copyright

Bibliography

Adriaense, J. E. C., Koski, S. E., Huber, L., & Lamm, C. (2020). Challenges in the comparative study of empathy and related phenomena in animals. *Neuroscience & Biobehavioral Reviews*, *112*, 62–82. https://doi.org/10.1016/j.neubiorev.2020.01.021

Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A. M., Cetkovich, M., & Ibáñez, A. (2017). Men, women…who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. *PLOS ONE*, *12*(6), e0179336. https://doi.org/10.1371/journal.pone.0179336

Barlińska, J., Szuster, A., & Winiewski, M. (2015). The Role of Short- and Long-Term Cognitive Empathy Activation in Preventing Cyberbystander Reinforcing Cyberbullying Behavior. *Cyberpsychology, Behavior, and Social Networking*, *18*(4), 241–244. https://doi.org/10.1089/cyber.2014.0412

Barraza, J. A., & Zak, P. J. (2009). Empathy toward Strangers Triggers Oxytocin Release and Subsequent Generosity. *Annals of the New York Academy of Sciences*, *1167*(1), 182–189. https://doi.org/10.1111/j.1749-6632.2009.04504.x

Batson, C. D., & Shaw, L. L. (1991). Evidence for Altruism: Toward a Pluralism of Prosocial Motives. *Psychological Inquiry*, *2*(2), 107–122. https://doi.org/10.1207/s15327965pli0202_1

Blair, K. S., Newman, C., Mitchell, D. G. V., Richell, R. A., Leonard, A., Morton, J., & Blair, R. J. R. (2006). Differentiating among prefrontal substrates in psychopathy: Neuropsychological test findings. *Neuropsychology*, *20*, 153–165. https://doi.org/10.1037/0894-4105.20.2.153

Blair, R. j. r., Peschardt, K. s., Budhani, S., Mitchell, D. g. v., & Pine, D. s. (2006). The development of psychopathy. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 262–276. https://doi.org/10.1111/j.1469-7610.2006.01596.x

Breil, C., Kanske, P., Pittig, R., & Böckler, A. (2021). A revised instrument for the assessment of empathy and Theory of Mind in adolescents: Introducing the EmpaToM-Y. *Behavior Research Methods*, *53*(6), 2487–2501. https://doi.org/10.3758/s13428-021-01589-3

Capozzi, F., & Ristic, J. (2020). Attention AND mentalizing? Reframing a debate on social orienting of attention. *Visual Cognition*, *28*(2), 97–105. https://doi.org/10.1080/13506285.2020.1725206

Carlson, S. M., & Moses, L. J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development*, *72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

Choi, D., & Watanuki, S. (2014). Effect of empathy trait on attention to faces: An event-related potential (ERP) study. *Journal of Physiological Anthropology*, *33*(1), 4. https://doi.org/10.1186/1880-6805-33-4

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Cole, G. G., Atkinson, M., Le, A. T. D., & Smith, D. T. (2016). Do humans spontaneously take the perspective of others? *Acta Psychologica*, *164*, 165–168. https://doi.org/10.1016/j.actpsy.2016.01.007

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A Review of the Concept. *Emotion Review*, *8*(2), 144–153. https://doi.org/10.1177/1754073914558466 de Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, *10*(10), 435–441. https://doi.org/10.1016/j.tics.2006.08.008

de Waal, F. B. M. (2008). Putting the Altruism Back into Altruism: The Evolution of Empathy. *Annual Review of Psychology*, *59*(1), 279–300. https://doi.org/10.1146/annurev.psych.59.103006.093625

Decety, J., & Lamm, C. (2006). Human Empathy Through the Lens of Social Neuroscience. *The Scientific World Journal*, *6*, 1146–1163. https://doi.org/10.1100/tsw.2006.221

Drayton, L. A., & Santos, L. R. (2018). What do monkeys know about others' knowledge? *Cognition*, *170*, 201–208. https://doi.org/10.1016/j.cognition.2017.10.004

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective Taking as Egocentric Anchoring and Adjustment. *Journal of Personality and Social Psychology*, *87*, 327–339. https://doi.org/10.1037/0022-3514.87.3.327

Fan, C. (2021). Processing Mechanisms Of Eye-Head Cues And Eye-Finger-Pointing Cues In The Dot-Perspective Task (Doctoral dissertation, Open Access Te Herenga Waka-Victoria University of Wellington).

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

Frick, P. J., O'Brien, B. S., Wootton, J. M., & McBurnett, K. (1994). Psychopathy and conduct problems in children. *Journal of Abnormal Psychology*, *103*(4), 700–707. https://doi.org/10.1037/0021-843X.103.4.700

Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 158–163. https://doi.org/10.1037/xhp0000138

Gardner, M. R., Bileviciute, A. P., & Edmonds, C. J. (2018). Implicit Mentalising during Level-1 Visual Perspective-Taking Indicated by Dissociation with Attention Orienting. *Vision*, *2*(1), Article 1. https://doi.org/10.3390/vision2010003

Geng, Y., Zhao, W., Zhou, F., Ma, X., Yao, S., Becker, B., & Kendrick, K. M. (2018). Oxytocin Facilitates Empathic- and Self-embarrassment Ratings by Attenuating Amygdala and Anterior Insula Responses. *Frontiers in Endocrinology*, *9*, 572. https://doi.org/10.3389/fendo.2018.00572

Hein, S., Röder, M., & Fingerle, M. (2018). The role of emotion regulation in situational empathy-related responding and prosocial behaviour in the presence of negative affect. *International Journal of Psychology*, *53*(6), 477–485. https://doi.org/10.1002/ijop.12405

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young Children Are Intrinsically Motivated to See Others Helped. *Psychological Science*, *23*(9), 967–972. https://doi.org/10.1177/0956797612440571

Heyes, C. (2014). Submentalizing: I Am Not Really Reading Your Mind. *Perspectives on Psychological Science*, *9*(2), 131–143. https://doi.org/10.1177/1745691613518076

Hodges, S. D., & Wegner, D. M. (1997). Automatic and controlled empathy. In *Empathic accuracy* (pp. 311–339). The Guilford Press.

Holland, C., Shin, S. M., & Phillips, J. (2021). Do you see what I see? A meta-analysis of the Dot Perspective Task. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*(43). https://escholarship.org/uc/item/7cs5r2xq

Kanske, P., Böckler, A., Trautwein, F.-M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and Theory of Mind. *NeuroImage*, *122*, 6–19.

https://doi.org/10.1016/j.neuroimage.2015.07.082

Krupenye, C., & Call, J. (2019). Theory of mind in animals: Current and future directions. *WIREs Cognitive Science*, *10*(6), e1503. https://doi.org/10.1002/wcs.1503

Lamm, C., Bukowski, H., & Silani, G. (2016). From shared to distinct self–other representations in empathy: Evidence from neurotypical function and socio-cognitive disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686), 20150083. https://doi.org/10.1098/rstb.2015.0083

Lamm, C., Rütgen, M., & Wagner, I. C. (2019). Imaging empathy and prosocial emotions. *Neuroscience Letters*, *693*, 49–53. https://doi.org/10.1016/j.neulet.2017.06.054

Langton, S. R. H. (2018). I Don't See It Your Way: The Dot Perspective Task Does Not Gauge Spontaneous Perspective Taking. *Vision*, *2*(1), Article 1. https://doi.org/10.3390/vision2010006

Lee, H.-J., Macbeth, A. H., Pagani, J., & Young, W. S. (2009). Oxytocin: The Great Facilitator of Life. *Progress in Neurobiology*, *88*(2), 127–151. https://doi.org/10.1016/j.pneurobio.2009.04.001

Marshall, J., Gollwitzer, A., & Santos, L. R. (2018). Does altercentric interference rely on mentalizing?: Results from two level-1 perspective-taking tasks. *PLOS ONE*, *13*(3), e0194101. https://doi.org/10.1371/journal.pone.0194101

Morelli, S. A., Lieberman, M. D., & Zaki, J. (2015). The Emerging Study of Positive Empathy. *Social and Personality Psychology Compass*, *9*(2), 57–68. https://doi.org/10.1111/spc3.12157

Nielsen, M. K., Slade, L., Levy, J. P., & Holmes, A. (2015). Inclined to see it your way: Do altercentric intrusion effects in visual perspective taking reflect an intrinsically social process? *Quarterly Journal of Experimental Psychology*, *68*(10), 1931–1951.

https://doi.org/10.1080/17470218.2015.1023206

O'Grady, C., Scott-Phillips, T., Lavelle, S., & Smith, K. (2020). Perspective-taking is spontaneous but not automatic. *Quarterly Journal of Experimental Psychology*, *73*(10), 1605–1628. https://doi.org/10.1177/1747021820942479

O'Neill, K., & Egermann, H. (2022). Induced empathy moderates emotional responses to expressive qualities in music. *Musicae Scientiae*, *26*(3), 476–496. https://doi.org/10.1177/1029864920974729

Pesimena, G., & Soranzo, A. (2023). Both the domain-general and the mentalising processes affect visual perspective taking. *Quarterly Journal of Experimental Psychology*, *76*(3), 469–484. https://doi.org/10.1177/17470218221094310

Premack, D., & Woodruff, G. (1978). Does a chimpanzee have a theory of mind. *Behavioral and Brain Sciences*, *1*, 515–526. https://doi.org/10.1017/S0140525X00076512

Preston, S. D., & Waal, F. B. M. de. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/S0140525X02000018

Rubio-Fernandez, P., Long, M., Shukla, V., Bhatia, V., & Sinha, P. (2022). Visual perspective taking is not automatic in a simplified Dot task: Evidence from newly sighted children, primary school children and adults. *Neuropsychologia*, *172*, 108256. https://doi.org/10.1016/j.neuropsychologia.2022.108256

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266. https://doi.org/10.1037/a0018729

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows:

Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 929–937. https://doi.org/10.1037/a0035175

Santiesteban, I., Kaur, S., Bird, G., & Catmur, C. (2017). Attentional processes, not implicit mentalizing, mediate performance in a perspective-taking task: Evidence from stimulation of the temporoparietal junction. *NeuroImage*, *155*, 305–311. https://doi.org/10.1016/j.neuroimage.2017.04.055

Shen, L. (2010). *State Empathy Scale* [Database record]. APA PsycTests. https://doi.org/10.1037/t81376-000

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Singer, T., & Lamm, C. (2009). The Social Neuroscience of Empathy. *Annals of the New York Academy of Sciences*, *1156*(1), 81–96. https://doi.org/10.1111/j.1749-6632.2009.04418.x

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, *303*(5661), 1157–1162. https://doi.org/10.1126/science.1093535

Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and Automatic Perspective Taking in Children and Adults. *Child Development*, *83*(2), 452–460. https://doi.org/10.1111/j.1467-8624.2011.01730.x

Tarrant, M., Dazeley, S., & Cottom, T. (2009). Social categorization and empathy for outgroup members. *British Journal of Social Psychology*, *48*(3), 427–446.

https://doi.org/10.1348/014466608X373589

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, *100*, 1027–1042. https://doi.org/10.1037/a0022308

Tomasello, M. (2014). The ultra-social animal. *European Journal of Social Psychology*, *44*(3), 187–194. https://doi.org/10.1002/ejsp.2015

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. https://doi.org/10.1002/ejsp.163

Vestner, T., Balsys, E., Over, H., & Cook, R. (2022). The self-consistency effect seen on the Dot Perspective Task is a product of domain-general attention cueing, not automatic perspective taking. *Cognition*, *224*, 105056. https://doi.org/10.1016/j.cognition.2022.105056

Watson, S. (2021, July 20). *Oxytocin: The love hormone*. Harvard Health. https://www.health.harvard.edu/mind-and-mood/oxytocin-the-love-hormone

Yan, Z., Hong, S., Liu, F., & Su, Y. (2020). A meta-analysis of the relationship between empathy and executive function. *PsyCh Journal*, *9*(1), 34–43. https://doi.org/10.1002/pchj.311

Yue, T., Jiang, Y., Yue, C., & Huang, X. (2017). Differential Effects of Oxytocin on Visual Perspective Taking for Men and Women. *Frontiers in Behavioral Neuroscience*, *11*. https://www.frontiersin.org/articles/10.3389/fnbeh.2017.00228

APPENDIX

<u>Appendix 1:</u> Table from experiment 1 reporting main effect results from two linear mixed models on (a) affect rating and (b) affect strength using the lmer() function in r. In two separate models, condition was used to predict rating and strength. For these models we included condition, gender and ethnicity with neutral, gender and white-british as their respective reference groups. A three way interaction was first included, however as it was non-significant, a two way was run with the same result until the model included no interaction at all. Bold p-values signalling significance.

(a) Affect Rating				
Factor	*Estimate*	*S.E*	*t value*	*p-value*
Intercept	0.979	0.147	6.646	
Condition(Neutral)	1.366	0.120	11.386	**<0.001**
Gender(Male)	0.013	0.125	0.106	0.916
Ethnicity(White-British)	-0.011	0.145	-0.076	0.940

(b) Affect Strength				
Factor	*Estimate*	*S.E*	*t value*	*p-value*
Intercept	1.043	0.136	7.697	
Condition(Neutral)	-0.683	0.094	-7.259	**<0.001**
Gender(Male)	-0.105	0.118	-0.894	0.373
Ethnicity(White-British)	0.088	0.137	0.641	0.523