Representing other minds: Mental state reference is moderated by group membership

Jennifer Susan McClung[a,b,*], Stephen David Reicher[b]

[a] University of Neuchâtel, Centre for Cognitive Science, Institute of Biology, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland
[b] School of Psychology and Neuroscience, University of St Andrews, South Street, St Andrews, KY16 9JP, Scotland, UK

ARTICLE INFO

ABSTRACT

The ability to infer the psychological forces that drive others' behaviour is a cornerstone of human cognition. This 'theory of mind' (ToM) we have has been extensively studied in its developmental stages and non-human forms. However, how the fully developed theory of mind functions on a daily basis is still the focus of ongoing research. One capacity stemming from theory of mind involves overt linguistic mental state reference. We propose that, rather than being a capacity that those with a fully developed ToM use consistently, mental state reference is a function of our social relationship to others: specifically, whether the other is perceived as an in-group or out-group member. We therefore examined spontaneous mental state reference during casual conversation as a function of group membership. Participants were divided into 'in-group' or 'out-group' pairs using a classic minimal group paradigm. Next, they were allowed to converse casually with their partner without the experimenter present and then subsequently asked to describe their partner in a written format after interactions. We scored participants' conversations and their written descriptions of each other for frequency and complexity of mental state reference. Results showed that, when interacting with presumed out-group members, participants referenced their partners' mental states significantly less often than when interacting with presumed in-group members. This effect was found both during conversations and in subsequent descriptions of the partner. Spontaneous mental state reference is apparently not a consistent psychological process but instead subject to social constructs, specifically group membership.

1. Introduction

Theory of mind, or the ability to infer unobservable mental states and to use these mental states to predict future behaviour, has long been investigated in its incomplete or premature forms. Both from a developmental perspective (Alison & Astington, 1988; e.g. Ensink & Mayes, 2010; Wellman, Cross, & Watson, 2001) and from an inter-species comparative perspective (e.g. Call & Tomasello, 2008; Heyes, 1998; Premack & Woodruff, 1978) the 'non-normal' theory of mind has been thoroughly canvassed, somewhat to the detriment of the study of the actual mechanism itself (Apperly et al., 2010; Apperly, Riggs, Simpson, Chiavarino, & Samson, 2006). The normally functioning adult theory of mind has received attention more recently in roughly the last decade. To date, evidence suggests that the use of a normally developed theory of mind (and a host of related processes) is heavily influenced by cognitive, cultural, and social factors.

First, in terms of the impact other cognitive processes have on theory of mind, people have difficulty interpreting another person's perspective without using their own knowledge as a template (e.g.

Fussell & Krauss, 1991; Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2003; Nickerson, Baddeley, & Freeman, 1987; Royzman, Cassidy, & Baron, 2003). This process, sometimes termed epistemic egocentrism, can lead to misjudgments about other's knowledge and occurs even when people are motivated to make accurate inferences (Keysar, Ginzel, & Bazerman, 1995). People also encounter difficulties in interpreting others' visual perspective in the face of high executive demands and distractions (Epley, Morewedge, & Keysar, 2004; Lin, Keysar, & Epley, 2010), as well as with lower moods (Converse, Lin, Keysar, & Epley, 2008). Furthermore, Apperly and colleagues have shown that questions requiring theory of mind usage are answered less quickly than non-mentalistic, reality-matching questions (Apperly et al., 2006) an indication that theory of mind processes may be a function of cognitive processing demands.

Second, cultural differences also seem to play a role in how effectively people take another's perspective, as shown by a study in which Chinese and American participants were asked to infer a partner's visual perspective (Wu & Keysar, 2007). The Chinese participants inferred their partner's visual perspective more accurately than their American

---

counterparts.

Third, social factors, specifically group membership, may also alter perception, making individuals less attendant to minds perceived as 'other' (Haslam, 2006) and more likely to stereotype those perceived as less similar (Ames, 2004). On the more extreme end, people categorized as 'other', or out-group, may be infrahumanized and attributed fewer uniquely human emotions (Leyens et al., 2001). Dehumanization research shows similar effects, in that out-group members are attributed fewer human values and traits and more animalistic qualities than are in-group members (for a review see Haslam, 2006). Furthermore, Hackel and colleagues showed that out-group members are required to be more human to be perceived as having a mind, in that shared group membership impacted how participants perceived the presence of mind in not only actual humans but dolls as well (Hackel, Looser, & Van Bavel, 2014). Group membership also affects how people empathise with others: considerable research has shown that empathic responses are lowered when observing out-group compared to in-group members (for a review see Cikara, Bruneau, & Saxe, 2011).

The previously discussed evidence suggests that normal processing of others' mental states is neither automatic nor consistent, and that social factors play a role in how people attribute emotions, perceive the presence of minds, and empathically respond. If group membership can affect these processes related to theory of mind, we wanted to address whether it would also impact the mechanism in its most basic and original form, the attribution of mental states to others (Premack & Woodruff, 1978).

The present study attempts to address these gaps by investigating whether group membership affects one aspect of theory of mind usage - mental state reference - in cognitively normal adults during typical, daily interactions. Specifically, we were interested in whether group membership plays a role in how people spontaneously reference others' mental states. The aim of this study was to gather data from the most natural contexts possible: unlike previous studies, we did not want to prompt theory of mind usage but rather to gauge one of its natural manifestations. We therefore examined natural social interactions for evidence of one manifestation of theory of mind usage, spontaneous mental state reference during casual conversation. We began by categorizing pairs of people using a classic minimal group paradigm based on estimation abilities into in-group and out-group pairs (Tajfel, Billig, Bundy, & Flament, 1971). After categorization participants were allowed to freely converse with each other, after which we asked participants to describe their partner in a written format.

Our aim was to compare participants' conversations and their descriptions of each other to assess the impact of group membership on spontaneous mental state reference. Linguistic reference to mental states has long been considered an indication of developmental processes children go through as they learn to form and use mental representations of others' mental states (Piaget & Inhelder, 1966). Even the usage of simple mental state verbs, such as 'to remember' or 'to hope', requires that the user form a mental representation of the target's mental state, specifically, what is being remembered or hoped (Antonietti, Liverta-Sempio, Marchetti, & Astington, 2006). As such, this type of reference has been used to study children's developing theory of mind (e.g. Meins et al., 2002; Meins, Fernyhough, Johnson, & Lidstone, 2006). This link between language and theory of mind has been exploited by researchers to develop the 'Spontaneous Theory of Mind Protocol' (STOMP), which measures spontaneous descriptions of the mental states of characters in videos to show that it correlates with thickness of certain cortical areas of the brain (Rice & Redcay, 2014). The STOMP approach used trained coders to differentiate between physical and mental state reference, whereas our design used a finite list of mental state reference words, based on and including words used to study mental-reference in children (Jenkins, Turrell, Kogushi, Lollis, & Ross, 2003) as well as the 'state verbs' used in the linguistic category model (LCM) approach (Semin & Fiedler, 1988). We used this list to examine participants' conversations and their subsequent written descriptions of each other to compare how participants referenced the mental states of in-group and out-group partners.

2. Method

2.1. Participants

Participants were recruited using noticeboards around the University of St Andrews after the study was approved by the university's Ethics Committee. Participants from all departments except psychology were accepted to take part in the study. 86 female undergraduates (age range 17–20) took part in the study to form a total of 43 pairs. In this way we controlled for the gender of our participants in order to avoid gender effects, or the possibility that participants would use gender to categorize themselves on top of our group manipulation (Ito & Urland, 2003).

Two pairs of participants were discarded from analysis: one because one of the participants had previous experience with minimal group paradigms, and another because the recording equipment did not function during the trial. In total this produced 41 pairs: 21 in the out-group condition and 20 in the in-group condition. All participants were tested in a single session lasting approximately 30 min. All participants were naïve to the experimental hypothesis, told that their data would be treated confidentially and used anonymously in publication, gave informed consent, were fully debriefed at the end of each experiment, and received £3 for participation.

2.2. Procedure

The experiment was conducted in the Social Immersion Lab in the Psychology Department at the University of St Andrews. Prior to the experiment it was confirmed that participants did not know each other in any way. Participants arrived at the lab at the same time and were given instructions before any social chatting could take place. Participants were given information forms describing the experiment and then asked to complete a consent form.

As the minimal group paradigm, participants were told the cover story that the experiment was designed to study the link between cognitive style and social interaction. To that end, the experimenter would first assess their cognitive style and then ask them to complete a social interaction task. Their cognitive style, they were told, would be assessed using a test called the 'Dot Estimation Task' (DET), which was in reality the minimal group paradigm used to categorize participants into out-group and in-group conditions (adapted from Howard & Rothbart, 1980). Participants were told that using the DET the experimenter would be able to tell whether they were over- or under-estimators, and that this categorization was significant since estimation abilities correlated with such abilities as spatial computation and mathematical skills. The DET itself involved estimating the amount of dots present on three consecutive pictures (made using Power Point, see Fig. 1 for example below). Dot pictures were presented for 3 s each using Microsoft Power Point and a projector.

Participants were asked to write down their estimates for each picture and to do the task alone in order to 'get a clear and true read-out' of each of their cognitive styles. In actual fact this request was designed to keep participants from discussing their answers and thereby realizing that there was no actual correlation between estimates and assigned category. The experimenter then made a brief show of calculating the average of their estimations, and then arbitrarily assigned each participant to be either an over- or under-estimator. Participants were asked to wear a badge with their estimation type displayed on it (two 'over' or 'under' estimators in the in-group condition, and one of each in the out-group condition). Participants were told this was so that 'the experimenter would not forget who was what for future analysis' whereas it was in fact done to maintain the salience of the categorization. Again, this categorization was in reality arbitrary.