Various Ways to Understand Other Minds: Towards a Pluralistic Approach to the Explanation of Social Understanding

ANIKA FIEBICH AND MAX COLTHEART

Abstract: In this article, we propose a pluralistic approach to the explanation of social understanding that integrates literature from social psychology with the theory of mind debate. Social understanding in everyday life is achieved in various ways. As a rule of thumb we propose that individuals make use of whatever procedure is cognitively least demanding to them in a given context. Aside from theory and simulation, associations of behaviors with familiar agents play a crucial role in social understanding. This role has been neglected so far. We illustrate the roles of fluency and associations in social understanding in false belief tasks.

1. Introduction

Understanding other people as mental beings having their own inner lives that encompass emotions, intentions, beliefs and desires plays a crucial role in successfully managing the social encounters of everyday life. For example, I can explain your reaching for an umbrella by attributing to you the *belief* that it will rain and the *desire* to stay dry (call this a 'reason explanation'). Furthermore, I may create expectations about your behavior on the basis of associations. For example, I may expect you to reach for the umbrella because I know of your habit of taking an umbrella with you when going out. The focus of this article is primarily the question of what the cognitive procedures are that underlie people's understanding of another person's behavior.

In the so-called 'theory of mind debate', there are two main approaches to explaining the cognitive procedures used to achieve social understanding: Theory Theory (TT) and Simulation Theory (ST). Proponents of TT (e.g. Baron-Cohen, 1995; Perner, 1999) claim that individuals understand the behavior of others by

We are grateful for the valuable comments and suggestions by Ian Apperly during the final stages of the formulation of this article. Moreover, we would like to thank a number of people for helpful earlier comments and discussions, in particular Mike Wilby, Alexandra Zinck, Bert Timmermans, John Michael, and Kirk Michaelian. A.F. acknowledges the support provided by the Barbara-Wengeler-Stiftung during her PhD as well as ongoing support from the project *The Social Mind: Self, Other, and Interaction in Cultural Contexts*, sponsored by the Volkswagen Foundation and led by Albert Newen. M.C. would like to thank the newly founded Center of Mind, Brain and Cognitive Evolution at the Ruhr-University Bochum, Germany; he was the Center's first Visiting Fellow during the initial stages of the formulation of this paper.

Address for Correspondence: Dr. A. Fiebich, Ruhr-University Bochum, Institute of Philosophy II, Universitätsstraße 150, 44780 Bochum, Germany.
Email: anifiebich@gmail.com

Mind & Language, Vol. 30, No. 3 June 2015, pp. 235–258.

using folk psychological rules such as 'if A wants p and believes that doing q will bring about p, then ceteris paribus, A will do q' (Borg, 2007, p. 6). Basically, there are two versions of TT, an empiricist version and a nativist version. Whereas empiricist theory theorists argue that theory change can and will happen as a consequence of encountering new evidence, nativist theory theorists assume that the core of a folk psychological theory is innate and that although a mindreading brain module might need experience as a trigger, it will not be modified by experience. Proponents of ST (e.g. Gordon, 1996; Goldman, 2006), in contrast, argue that in seeking to achieve social understanding, individuals put themselves imaginatively 'into the shoes' of other people and simulate the thoughts and feelings they themselves would experience if they were in the situations of the other people. In simulation processes, individuals use their own mental experiences as an internal model for the other mind without (necessarily) referring to any (folk psychological) theory.

These accounts are, however, subject to some restrictions. Though proponents of both camps acknowledge that there are various ways of understanding other minds (i.e. a number of theory theorists allow for simulation playing a role in social cognition, and vice versa), all share the assumption that there is a default procedure that individuals apply whenever they are attempting to understand other minds. The default procedure is either theory (according to TT) or simulation (according to ST). In addition, by appealing to psychological generalizations (such as TT) or to taking one's own self as a model when running simulation routines (such as ST), neither camp provides procedures that account for the subject's knowledge about other people's person-specific features such as character traits, attitudes, habits, etc. that are characteristic for an individual's 'person-specific identity' nor for features that characterize those people's 'social identity' (i.e. the identity they have because of their belonging to a particular social group, such as being an 'Asian' or 'female'). Rather, TT and ST tend to describe social understanding as a 'depersonalized' and observational enterprise that relies on nothing but mental state attributions. Developmental psychologists from these camps have conducted a series of so-called 'theory of mind tasks' that investigated the development of mental state attribution throughout ontogeny. The ontogenetic development of false belief understanding has especially gained attention in this debate. Finally, neither TT nor ST account for the various findings from social psychology in their theories of social understanding.

In this article, we challenge the traditional assumption by introducing a pluralistic two-system account of social understanding that advocates the view that individuals expect, predict or explain the behavior of other people in everyday life in various ways depending on the contextual features of the particular situation of social understanding in which they find themselves and also depending on their cognitive competencies. Our approach does not only integrate the findings from developmental psychology and the cognitive procedures that have been proposed in framework of the theory of mind debate by referring to this literature (see above), but also integrates the various findings from social psychology that appeal to trait attribution, impression making, associations, and different systems of cognitive processing

playing a role in (social) cognition. In addition to theory and simulation, we appeal to procedures that take into account the person-specific or social characteristics of the individual whose behavior needs to be understood. We make a methodological distinction between behavior expectations, behavior predictions and behavior explanations, and we also emphasize the particular role that an interactive stance plays in behavior expectations. Moreover, like other psychologists who refer to two different systems of cognitive processing (see Section 2), we draw on Kahneman's (2011) two-system account of cognition. In line with this research, we propose—as a rule of thumb—that individuals are prone to make use of whatever procedure is cognitively least effortful in a given context of social understanding. Our approach has the advantage over other pluralistic approaches to social understanding (e.g. Andrews, 2012) that it provides a theory about which procedure individuals are likely to use in the various contexts of everyday life.

In Section 2, we illustrate the role that different cognitive systems play in (social) cognition. In Section 3, we present our pluralistic approach to the explanation of social understanding and discuss a variety of cognitive procedures that are involved in behavior expectations, predictions and explanations. Here we emphasize that fluency (i.e. the subjective experience of ease during the accomplishment of social understanding) and associations play an important role in social cognition, a role which, moreover, has been much neglected in the theory of mind debate. Finally, we illustrate the implications of our approach in experimental settings such as false belief tasks in Section 4. In Section 5, we end with a summary and general discussion.

2. Two Cognitive Systems

A number of researchers have argued for a two-system approach to cognition in general (e.g. Kahneman, 2011) and social cognition in particular (e.g. Apperly and Butterfill, 2009; Baillargeon *et al.*, 2010; Carruthers, 2013), i.e. cognition that concerns the understanding of the behavior and inner life of other people. Whereas nativist theory theorists typically treat the two cognitive systems as subcomponents of a single cognitively efficient 'theory of mind' mechanism (e.g. Baillargeon *et al.*, 2010), Apperly and Butterfill (2009) recently proposed a two-system account of social cognition that involves one system that is efficient and inflexible (System 1), and another independent system that is flexible but cognitively demanding (System 2). On this account, implicit (i.e. nonlinguistic) false belief understanding is the responsibility of System 1 and enables infants to track 'belief-like states'. In contrast, explicit (i.e. linguistic) false belief understanding is the responsibility of System 2 and involves folk psychological knowledge of what mental states are and how they interrelate. System 1 is ontogenetically prior to System 2 but tracking 'belief-like states' or 'registrations' continues to play a role in adult social cognition if the situation in question features some signature limits. Unlike beliefs, 'belief-like states' must be relations to objects and properties, not to propositions, though such states resemble beliefs in having correctness conditions: for example, they fail to be correct if the

object registered is not where it is registered as being. We also defend a two-system approach to social cognition in general and false belief tasks in particular (see Section 4 for a discussion of this).

In Section 3, we introduce our pluralistic approach to the explanation of social understanding that draws on Kahneman's (2011) two-system approach to cognition, according to which the cognitive processes of System 1 are fast, relatively effortless routines that may occur without awareness (Kahneman, 2011; see also Frith and Frith, 2008) whereas the cognitive processes of System 2 are slow routines which require the expenditure of mental effort and are subject to consciousness and deliberative control. Following Kahneman (2011) we propose that 'the automatic operations of System 1 generate surprisingly complex patterns of ideas, but only the slower System 2 can construct thoughts in an orderly series of steps' (p. 21).

Reviewing a number of studies from social psychology, Kahneman pointed out that when people create judgments that involve syntactically structured thoughts, their System 2 most often adopts the impulses and associations of System 1. System 1 is modular but flexible insofar as it continuously integrates novel information. It has learned associations between ideas; some of these associations are widely shared among individuals, such as the association of particular character traits with members of social groups. Kahneman provides the example of people's widely shared association of 'shy and withdrawn' with the occupation of librarian, which requires common knowledge of the language and culture; call this 'association with social identity' (short: association_{si}). The role of associations_{si} in social understanding and interaction has been investigated intensively in research on so-called 'stereotype activations'. However, other associations are learned only by some individuals. For example, only people who know Lisa personally may associate 'being shy and withdrawn' with her; call this 'association with person-specific identity' (short: association_{pi}). In our account, not only associations_{si} but also associations_{pi} play a primary role in social understanding (also called 'person categorization' versus 'person individuation', see, e.g., Mason and Macrae, 2004). In everyday life, we are often familiar with the individuals whose behavior we attempt to understand. Here we often draw on our associations with these familiar individuals in situations of social understanding rather than engaging in (cognitively more effortful) theorizing or simulating processes.

As a rule of thumb, we propose that people typically make use of that socio-cognitive procedure that is cognitively least effortful in a given context. In typical situations of everyday social understanding, people associate ease with completing the mental task in question. That is, 'fluency', which is defined as 'the subjective experience of ease or difficulty associated with completing a mental task' (Oppenheimer, 2008, p. 237), plays a primary role in social cognition.

A number of studies in other domains have shown that fluency has impacts upon which procedure individuals use to solve a mental task; whereas individuals draw on simple procedures that involve fast, heuristic and relatively effortless reasoning processes when they experience cognitive ease, they make use of more complex, cognitively demanding and effortful strategies when a particular task appears to be difficult to them to solve, i.e. when they experience cognitive strain. Moreover, the

experience of cognitive ease depends on various variables such as repeated experience, a clear display, a primed idea, or a good mood, and is characterized by feelings of familiarity, trueness, goodness, and effortlessness (see Kahneman, 2011, pp. 60 ff. for a discussion). Findings from social psychology show that associations with familiar social groups are indeed automatically activated and hence cognitively easy to access (Macrae and Bodenhausen, 2000).

As we will argue in Section 3, nonlinguistic associations with person-specific identity or social identity can be sufficient for individuals to *expect* a particular behavior of a familiar person or a member of a familiar social group in a given context. In contrast, cognitively more 'costly' behavior expectations that rely on linguistic associations additionally involve theoretical knowledge about what mental states (e.g. beliefs and desires) or dispositions (e.g. character traits, habits etc.) are, and how they motivate agents to act in a given context. Typically, behavior expectations rely on System 1 processes. Behavior predictions or explanations, in contrast, rely (at least partially) on structured thoughts (System 2) and hence can be verbally offered. Depending on the kind of judgment, however, associations (System 1) may also come into play.

In principle, individuals can use a theory on the basis of simple intuitive heuristics involving psychological generalizations (System 1), or they may instead engage slow and effortful reasoning processes (System 2). Likewise, simulation processes can be based on either system. For example, Goldman's (2006) conception of 'low-level simulation', which is based on mirror neuron activity, can be regarded as relying on System 1 processes, whereas his 'high-level simulation', which involves conscious creation of pretend states and a simulated decision process regarding how to act based on these states, depends upon System 2 or System 1 in a more costly manner than low-level simulation does. In general, we take 'costliness' of System 1 processes as a matter of degree, proposing that implicit theory and high-level simulation that relies on the manipulation of symbolic representations and a complex set of concepts and their interrelations is cognitively more 'costly' than low-level simulation and associations.

3. Towards a Pluralistic Approach to the Explanation of Social Understanding

As indicated above, social understanding may occur in the form of *expectations* or *predictions* when foreseeing the future behavior of another person, or *explanations* when interpreting the past behavior of another person. In behavior explanations, people answer the 'why-question', i.e. they point out *why* an individual has performed a particular action in the past (e.g. by citing the individual's reasons to act, or the dispositions that guided the individual's action). In behavior expectations and behavior predictions, in contrast, people answer the 'what-question', i.e. they point out *what* an individual will do in the future. Whereas behavior predictions and behavior explanations rely on syntactically structured thoughts (System 2) that can

be verbalized, behavior expectations draw on impressions that rely, for example, on associations (System 1).

Additionally, we follow Tomasello (1999) in distinguishing between understanding other people as *intentional agents* and as *mental agents*. Tomasello (1999, p. 179) highlights that understanding other people as intentional agents 'includes an understanding of both goal-directed behavior and the attention of others', whereas understanding other people as mental agents presupposes 'the understanding that the other persons have not just their intentions and attention as manifest in their behavior, but also thoughts and beliefs which may or may not be expressed in behavior—and which may differ from the 'real' situation' (p. 179).

The ability to understand another person as an intentional agent emerges early in ontogeny and is already present in infants just a few months old, who, via the use of person-specific dispositions and various teleological reasoning principles, form expectations about other people's pursuit of goals. In contrast, understanding other people as mental agents involves an understanding of other people's behavior as being guided by reasons. This presupposes meta-representational activities and a linguistic grasp of what mental states such as beliefs and desires are, how they interrelate, and how they motivate agents to act. Children have been found to first exhibit such understanding at around age 5, when they perform successfully in explicit (i.e. linguistic) versions of the false belief task (Wellman *et al.*, 2001, see Section 4 for a discussion).

We add a further characteristic to Tomasello's distinction that is at least not explicitly spelled out by him: when understanding another person as a mental agent, that person is deemed to act on reasons, i.e. he or she is aware of her mental states (e.g. his or her desire to eat chocolate and her belief that there is a chocolate bar in the kitchen) and acts rationally by taking these states into account. When understanding another person as an intentional agent, in contrast, the person in question is deemed to act on dispositions, i.e. not to be aware of his or her own motivational states (e.g. character traits, habits, etc.) but to act on these states in a situation that triggers them.

Understanding another person as an intentional or mental agent may rely on a variety of cognitive procedures. Theory and simulation have been convincingly proposed to be the cognitive strategies based upon which understanding other people as mental agents can be achieved. As we will see below, theory or simulation may also be involved in understanding other people as intentional agents. However, it is also the case that procedures which refer to what is associated with the person-specific identity, or the social identity of the individual whose behavior needs to be understood, may play the central role here. Moreover, cognitive procedures such as theorizing or simulation may rely on different cognitive systems (see Section 2).

3.1 Behavior Expectations

People's expectations of what another person will do in the future may be manifested in various ways, ranging from a gaze towards the goal the agent is expected to achieve to interactive responses towards what the agent is expected to do. As pointed out by

Kahneman (2011, p. 72), on the basis of associations people either *actively* expect the occurrence of events in general and aspects of other people's behavior in particular (i.e. they consciously wait for the event to occur), or else *passively* expect these (i.e. they do not consciously await the occurrence of events, but they are surprised if these events do not occur). Contrary to behavior predictions or explanations that are always conscious, behavior expectations can be conscious (active expectations) or not (passive expectations). Moreover, in contrast to behavior predictions or explanations, behavior expectations do not involve structured thoughts (System 2) and instead rely exclusively on fast cognitive processes that require little or no effort such as associations or cognitively 'cheap' theorizing or simulation processes (System 1). Here we focus on different kinds of associations.

As indicated above, we distinguish methodologically between two kinds of associations that may underlie behavior expectations and are automatically triggered in a specific context: (i) *associations with person-specific identity* (*associations*$_{pi}$), and (ii) *associations with social identity (associations*$_{si}$*)*. All of these associations are 'mentalistic' because they are restricted to animate beings and presuppose an implicit sense of the other person as an intentional agent. Associations may be nonlinguistic or linguistic; if linguistic, they may also enter into verbally uttered behavior predictions or explanations (see Section 3.2).

A number of studies have indicated that people's behavior expectations often rely on associations$_{si}$. Associations$_{si}$ are present early in ontogeny. For example, 3-month-olds are already sensitive to the social identity of an individual, i.e. social group membership, and prefer to interact with people with own-race faces (Kelly *et al.*, 2005). Later on in ontogeny, individuals categorize people into social identity groups and expect them to act on dispositions on the basis of associations$_{si}$, also called 'stereotype activation' in the literature.

Empirical studies have indicated that individuals attribute particular character traits and attitudes to a stranger on the mere basis of features of that stranger's outer appearance such as clothes, attractiveness, skin color, or gender, and this in turn might shape their expectations of the stranger's behavior. For example, individuals assume that people with stylish hair and extravagant clothing are highly extroverted (Borkenau and Liebler, 1992). By definition, extroverted individuals are primarily oriented to social settings, whereas introverts are more interested in an internal environment, preferring to listen and reflect (Lu and Hsiao, 2010, p. 151). Hence, if individuals deem a person to be highly extraverted, they will expect this person to behave in an open-minded, talkative manner towards the social environment.

Other associations$_{si}$ involve categorization according to social groups such as gender (Eagly and Steffen, 1984) or race (Lin *et al.*, 2005). Associations$_{si}$ may be exhibited by people's interpersonal behavior, e.g. associating laziness or incompetence with overweight individuals (Brochu and Morrison, 2007; Puhl *et al.*, 2008) may determine how far people choose to sit from an overweight woman (Bessenoff and Sherman, 2000).

We also propose that people can form expectations that members of a particular social group will perform a particular action in a specific context by referring to past

actions in that context that have been observed to be performed by members of that group, without the need of linguistic knowledge of (conventional) rules that may underlie such behavior. Having observed Asians eating with chopsticks in restaurants where chopsticks and forks are available, you may be surprised by seeing an unknown Asian person eating with a fork because you have a (passive) expectation that Asian people eat with chopsticks on the basis of non-linguistic associations$_{si}$.

In general, associations$_{si}$ are acquired early in ontogeny and relatively resistant to modification (Macrae and Quadflieg, 2010, p. 435). Moreover, associations are activated automatically (Macrae and Bodenhausen, 2000) and are indicated by spontaneous interactive reactions towards members of a particular (social) group consistently over time.

Analogously to associations$_{si}$, we propose that people associate person-specific behaviors, habits, and character traits with familiar individuals. Developmental studies suggest that such associations$_{pi}$ are also present early in ontogeny and play a role in preverbal infants' social understanding. Associations$_{pi}$ presuppose the ability to recognize a particular individual on the basis of person-specific physical features. From birth onwards, infants are sensitive to other people's person-specific identity, as is indicated by their ability to recognize the faces of persons they frequently interact with, such as their mothers (Pascalis *et al.*, 1995).

Moreover, developmental studies have shown that infants as young as 6 months of age are already sensitive to other people's goal-directed movements, e.g. grasping gestures, and are capable of encoding an agent's goal (Woodward, 1998). Neuroscientific studies suggest that mirror neuron activity (referred to as 'low-level simulation' by Goldman, 2006[1]) plays a central role in goal-recognition. Indeed, Nyström (2008) recently found that mirror neurons are activated in the observation of goal-directed movements not just in adults but also in 6-month-olds.

But not only are 6-month-old infants capable of encoding an agent's action goal when observing his or her goal-directed movement, they also expect a familiar agent to act on his or her dispositions in a specific context, achieving such expectations by referring to a familiar agent's past purposeful actions that the infant has observed before (Woodward, 1998, 1999). On our account, such expectations rely on associations$_{pi}$ that are created during familiarization trials, as well as on a few core teleological reasoning principles. That is, preverbal infants expect a familiar agent to act on dispositions even though such infants do not yet have a concept of what 'dispositions' are. We understand 'dispositions' in an Aristotelian sense as motivational powers or forces that are triggered by contextual features. Motivational forces include person-specific preferences, character traits, habits etc. that may enter an 'other-concept' once linguistically conceptualized. In our account, an 'other-concept' entails the unity of linguistic associations$_{pi}$ and linguistic associations$_{si}$ an individual has from another person. As indicated above,

[1] Whether or not the actions of mirror neurons can be regarded as simulation routines remains controversial (see, e.g., Gallagher, 2007, for a critical discussion).

associations$_{si}$ typically get activated on the basis of bodily features that indicate another person's social identity. For example, I may associate 'is a doctor' with a stranger who is wearing doctor-style clothing, and expect him or her to operate on patients in a hospital on the basis of this association. However, associating 'is a doctor' with my brother Paul may lead me (but not you, since you do not know Paul) to the very same expectations in this situation even if Paul is not wearing doctor-style clothing. Woodward (1999) found that infants only create agent-goal-context associations in specific contexts (e.g. in which infants see an object A and an object B) when they perceive the agent acting on a particular object in an intentional (i.e. purposeful) manner; for example, when they identify the agent's action goal (e.g. getting object A) on the basis of the agent's purposeful gestures towards object A (e.g. grasping gestures). That is, it is neither merely the other person's touching of an object, nor the manual gesture of approaching the object, that makes infants recognize the person's action as goal-directed. These findings speak in favor of what Csibra and Gergely (2007) call the 'efficiency principle' as underlying infants' recognition of an agent's action goal; the infants perceive the grasping gesture as an efficient way to achieve the action goal (i.e. getting the object). This is analogous to Butterfill's and Apperly's (2013) postulated first principle of minimal theory of mind cognition, according to which goal-directed action is characterized in terms of functions in a teleological sense. In addition to what Csibra and Gergely (2007) have proposed, Butterfill and Apperly (2013) emphasized that 'representing goal-directed action […] does not require representing representation' (p. 614).

Moreover, 6-month-olds already take into account the agent's visual access in this experimental setting. As shown by Luo and Johnson (2009, p. 142), 'the infants took the agent's actions of repeatedly reaching for and grasping one of two possible objects as suggesting her preference for that object only when the agent could detect both objects, not when the agent's perceptual access to the second object was absent, either because a large screen hid the object from the agent (Experiment 1), or because the agent sat with her back toward the object (Experiment 2)'.

In general, nonlinguistic associations$_{pi}$ may underlie not only infant but also adult behavior expectations. Having observed Bill on various occasions giving large tips in restaurants, I can expect Bill to give large tips in the specific context 'restaurant' without my having a concept of 'generosity'. Nonlinguistic associations get activated in specific contexts (e.g. 'restaurants') whereas linguistic associations get activated also in unspecific contexts depending on the linguistic concept that underlies the association in question. For example, associating 'generosity' with Bill, I may expect him not just to give large tips in restaurants but also in other contexts that satisfy my linguistic concept of 'generosity' (e.g. giving a beggar money).

In general, behavior expectations can shape individuals' interactive responses. Expecting my sister to recoil from hedgehogs (perhaps due to a traumatic event she experienced in the past), I may cover my hedgehog's cage with a blanket when my sister is entering my living room. Phenomenologically, expectations can be experienced as a vague idea of what will happen in the future. Expectations based on associations may also have an evaluative component and play a central role for

social understanding from an interactive stance. Deeming Susi to be a person, who behaves roughly, for example, my disliking roughness may lead me to flinch from her once she enters a scene.

Finally, my association of Bill with 'being generous' might even not be tied to Bill but just to his outer appearance, since people have been found to associate particular character traits with strangers on the basis of perceived similarity to well-known persons such as good friends; White and Shapiro (1987) showed that $\text{associations}_{\text{pi}}$ may also become activated on the basis of a perceived similarity of a stranger to a well-known other person.

To sum up, behavior expectations may rely not only on theorizing or simulating processes but also on a variety of associations, including $\text{associations}_{\text{pi}}$, and $\text{associations}_{\text{si}}$. Linguistic or not, any of these associations refers to a disposition (e.g. a character trait or a preference) to perform a particular action in a specific or unspecific context. Hence the agent whose behavior is understood is deemed to be an intentional agent, i.e. one who acts on dispositions.

3.2 Behavior Predictions and Behavior Explanations

Following Andrews (2012) we caution against the 'symmetry thesis' according to which there is 'no prediction without explanation, and no explanation without prediction' (p. 38). Structured thoughts of what an individual is expected to do in the future (prediction) or why an individual has performed an action in the past (explanation) may but do not need to rely on the same cognitive mechanisms. I may be able to correctly predict *what* Bill will do, e.g. giving a large tip in a restaurant because I have observed him giving large tips in restaurants on various occasions on the basis of nonlinguistic $\text{associations}_{\text{pi}}$ (see Section 3.1) but without being capable of indicating *why* Bill gave a large tip on this particular—or indeed any—occasion; maybe because I lack the concept of 'generosity' when deeming him to act on a disposition, or maybe because I have no ideas about his beliefs and desires when conceiving him as a mental agent.

In our account, behavior predictions may be verbally encoded active expectations that rely on nonlinguistic associations (see Section 3.1) or linguistic associations, or they may instead rely on theorizing or simulating processes that involve reference to mental states or dispositions. As with behavior expectations, people answer the 'what-question' by behavior predictions. That is, they indicate *what* another person will do in the future. Determining another person's future action may draw on mental state or character trait attributions that rely on System 1 or System 2 processes. In general, however, no explicit knowledge about how particular mental states or character traits cause particular actions needs to be involved. In behavior predictions, it is sufficient to indicate *that* an agent will perform a particular action. Behavior explanations, in contrast, presuppose a linguistic grasp of how mental states or dispositions (e.g. character traits or habits) guide an agent's behavior in a given context. In behavior explanations, people answer the 'why-question' by explaining why a person has performed a particular action in the past.

In general, behavior explanations can be separated into different kinds in terms of various criteria (e.g. the length and richness of information of the narrative, the explicit mention of the person or socio-situational context etc.). In the present article, we make use of Malle *et al.*'s (2007) distinction between what they call 'reason explanations' and 'causal history explanations' which parallels the distinction we drew earlier between understanding another person as an intentional agent and understanding another person as a mental agent.

According to Malle *et al.* (2007, pp. 492–3), 'reasons can be defined as the contents of an agent's mental states (primarily beliefs and desires) in light of which and on the grounds of which the agent formed an intention to act. When people cite a reason explanation for an action, they ascribe to the agent one or more beliefs or desires that (they presume) figured in the agent's decision to so act'. In general, reason explanations rely on two assumptions; (i) the assumption of subjectivity (i.e. the agent is assumed to have been aware of his or her reasons and acted on those reasons), and (ii) the assumption of rationality (i.e. the action in question follows the rules of practical reasoning). That is, when people offer reason explanations, they deem the agent to be a *mental agent*.

Neither of these assumptions is made when people generate another kind of explanation that Malle *et al.* call 'causal history explanations'; 'causal history explanations help explain an intentional action by citing causal antecedents to the agent's reasoning and decision to act, but there is no assumption that the agent actively considered those antecedents in her reasoning process' (p. 493). For example, when an observer offers a causal history explanation such as 'Bill gave a large tip, because he is generous', the observer does not assume that Bill takes into account his generosity as a reason for giving a large tip ('I am generous. Therefore, I should give a large tip') but rather considers Bill's generosity as the causal background that motivates Bill's action. That is, Bill is conceived to be an *intentional agent*.

As we will argue below, when people create explanations of another person's behavior which require syntactically structured thoughts with semantic content (System 2), they either explicitly refer to the agent's *disposition* to perform a particular action in a given context in 'causal history explanations' (e.g. by attributing character traits such as laziness to the agent), or else they refer to the agent's *reasons* for performing an action in a given context in reason explanations (e.g. by attributing beliefs and desires to the agent). That is, people firstly need to determine the agent A's dispositions or reasons to act, which they then integrate into a behavior explanation in a second step. In either kind of explanation, such integration can rely on either using theory or running simulation routines. For example, when offering causal history explanations, people can explain an agent A's behavior by referring to A's character trait ('being c') in framework of a theory that involves folk psychological rules ('A is c and people who are c do x in context y') or simulating activities ('If I were c and would have done x in context y, why would I have done this?').

In this section, we focus on the cognitive procedures that underlie people's *determination* of dispositions or reasons to the agent. On our account, determining an agent's dispositions to act typically relies on different kinds of associations, whereas

attributing reasons to the agent is based on theory use or simulation routines. Reason explanations have been discussed extensively in the theory of mind debate. In line with this work, we propose that when people offer reason explanations in form of explanations, such as 'Kim went to the café because she believed they have the best cappuccino', theory use or simulation routines are the underlying cognitive procedures for inferring Kim's reasons for going to the café.

It is necessary to clarify what theory theorists mean by the term 'theory'. Whereas the definition of 'theory' is not always spelt out explicitly, Gopnik (1998, pp. 34–36) provides a broadly defined concept of 'theory' that seems to be shared by (at least most) proponents of TT. On this account, the structural features of theories involve (i) abstractness, (ii) coherence, (iii) causality, and (iv) ontological commitment. On Gopnik's account, a folk psychological theory involves the possession of mental state concepts and hence presupposes linguistic skills. In the present investigation, we draw on this notion of 'theory'.

Theories are likely to be employed in experimental task settings such as explicit (i.e. linguistic) theory of mind tasks. Here children are capable of using folk psychological rules to attribute to an agent mental states such as beliefs and desires. In general, folk psychological knowledge of how mental states interrelate and determine an agent's decision to act enables individuals to integrate mental state attributions into a behavior explanation involving structured thoughts (System 2) when creating reason explanations.

Theory theorists agree that mental state attribution on the basis of a folk psychological theory can happen with varying degrees of mental effort and consciousness. Inferring A's mental state in a given context (as well as its interplay which constitutes A's reasons to act) can be an extremely effortful and conscious process (System 2). On other (perhaps more) occasions, where mental states are easily inferrable on the basis of simple heuristics, theories might be used with little or no conscious mental effort (System 1).

In contrast, simulation theorists 'note that the simulation process does not rely on mind readers' appeal to psychological generalizations (e.g. a generalization about human decision making), which is a crucial part of TT' (Shanton and Goldman, 2010, p. 531). Rather, simulation theorists argue that individuals are engaged in running simulation routines each time anew when it comes to understanding other minds in a given situation. In 'high-level simulation' (Goldman, 2006), for example, individuals put themselves imaginatively into another person's shoes, creating 'pretend states' that can be fed into their own decision-making mechanism, which generates a particular decision that allows them to predict or explain the other's behavior.

Like theory use, high-level simulation can require the expenditure of mental effort and structured thoughts (System 2) or rely on fast cognitive processes that may not be conscious at all (System 1); According to Goldman (2006, p. 151), 'a great deal of mindreading, even high-level mindreading, is nonconscious or minimally conscious, so we should allow simulational processes to include E-imaginative states even when

the latter are entirely nonconscious'.[2] That is, when people offer 'reason explanations', their System 2 (which creates the structured thoughts being involved in such an explanation) can draw on mental state attributions that have been created in the framework of conscious and effortful simulating activities (System 2) or simulation routines that required little or no mental effort (System 1). Again, System 1 processes need to be regarded as a matter of degree; that is, high-level simulation that relies on System 1 is still cognitively more costly than low-level simulation due to the manipulations of symbolic representations involved.

Traditionally, theory use and simulation use have been regarded as being involved in understanding another person as a mental agent when creating 'reason explanations'. Both of these procedures, however, may also play a role in creating 'causal history explanations'. In principle, offering 'reason explanations' presupposes the determination of an agent's mental states that can rely on theory use or simulation routines, whereas offering 'causal history explanations' presupposes the determination of an agent's disposition to act. We hypothesize that such determination may rely on theorizing or simulation processes but typically relies on a variety of associations given that these are available. When people offer causal history explanations, they consider the agent to be an intentional agent. That is, they take into account the causal antecedents that dispose the agent towards a particular behavior in a given context or that are antecedents that influence the agent's reasoning and decision to act (without the agent consciously knowing this). In this article, we focus on the former, i.e. on explanations that refer to an agent's dispositions to perform particular actions in specific contexts.

Associations of dispositions with a familiar person (associations_{pi}), or with membership of a familiar social group, i.e. social identity (associations_{si}), may underlie behavior expectations (see Section 3.1). On our account, in addition to theory and simulation, the linguistic counterparts of the very same associations may come into play in behavior explanations. For example, when offering a causal history explanation, I may refer to Bill's 'generosity' to explain why he gave a large tip in a restaurant.

As indicated in the previous section, individuals may refer to an agent's dispositions on the basis of linguistic associations that are part of an 'other-concept'.

[2] E-imagination (i.e. enactment imagination) is involved in our simulation processes. We E-imagine being in the other's shoes and simulate the thoughts and feelings we would experience in her situation. Goldman (2006, p. 149) distinguishes between two species of E-imagination: visual E-imagination and motor E-imagination. Visual E-imagination involves visualizing how things look from another person's perspective in order to determine whether particular objects are in her sight (i.e. her visual state) as well as any consequent beliefs about that object. Motor E-imagination, in contrast, involves an imagination of executing certain bodily movements. This needs to be distinguished from a stimuli-dependent mirroring activity since, as pointed out by Goldman (2006, p. 158), 'mirroring activity is an involuntary response to perceptual stimuli, whereas motor imagination is subject to voluntary control and not normally driven by any distinctive class of perceptual stimuli'.

Moreover, we propose that individuals have linguistic concepts of members of particular social groups ('social identity concepts') that may play a role in social understanding, i.e. a set of characteristics that individuals refer to when offering 'causal history explanations'. For example, I may explain Susi's reaching for the doll rather than the toy car by referring to her gender-specific preference to do so. Unlike linguistic associations$_{pi}$, we propose that linguistic associations$_{si}$ do not only involve group-specific dispositions such as character traits, preferences etc., but also dispositions to follow social rules. For example, I may explain your eating rice with chopsticks rather than a fork by referring to the conventional eating rules you pursue due to your Asian social identity.

Of course, like mental states, causal antecedents may also be determined by theorizing or simulating processes. That is, people may infer via theory or simulation what another person's causal antecedent (e.g. the character trait 'being generous') is, based upon the behavior perceived (e.g. giving large tips in a restaurant), and cite this causal antecedent when offering a causal history explanation. Malle *et al.* (2007) found that people typically offer reason explanations more often than causal history explanations. Intriguingly, however, people offer causal history explanations more often than reason explanations when they are familiar with the agent whose behavior they are attempting to understand. On our account, this is due to people's tendency to provide behavior explanations that rely on cognitive procedures of least cognitive effort. If available, associations$_{pi}$ or association$_{si}$ are automatically activated and people may easily refer to such associations when offering causal history explanations. In contrast, belief attribution does not rely on an automatic process (Apperly *et al.*, 2006).

Following Kahneman (2011) we point out that although only System 2 is capable of forming verbal judgments and structured thoughts, people's own reasoning processes as well as their judgments about other people's behavior most often rely on fast and automatic associations and 'intuitive heuristics' (System 1); as Kahneman (2011, p. 24) says, 'System 1 continuously generates suggestions for System 2: impressions, intuitions, intentions, and feelings. If endorsed by System 2, impressions and intuitions turn into beliefs […]. When System 1 runs into difficulty, it calls on System 2 to support more detailed and specific processing that may solve the problem of the moment'. The use of theory and simulation routines (by means of which reasons or causal antecedents can be inferred) can be relatively 'cheap' methods for inferring mental states or causal antecedents. However, we propose that even cheap theorizing or simulating processes are more costly than any kind of association. If individuals are tempted to offer a causal history explanation but run into trouble because no associations are available that help to explain the agent's behavior, they may try more effortful cognitive processes; for example, they may try to remember other dispositions of the agent that explain his or her behavior, or theorize or simulate what the 'causal antecedent' is that underlies the agent's action, or they may conceive the agent's behavior no longer as unintentional but as intentional and offer reason explanations rather than causal history explanations. As we have seen above, theory use or use of simulation routines can serve as a

means of determining an agent's reasons to act. This speaks in favor of Gallagher's (2001) proposal that individuals often engage in cognitively effortful theorizing or simulating activities as a consequence of being puzzled about another person's behavior.

4. Research Implications

Our pluralistic approach to the explanation of social understanding abandons the traditional assumption, made both by proponents of ST and by proponents of TT, that there is a default procedure for achieving social understanding. Instead, we propose that understanding other minds can be achieved in various ways, none of which is a default procedure. In general, we have integrated the findings of the theory of mind literature with the findings of the social psychology literature in our approach to the explanation of how social understanding is achieved, which is that people typically make use of that cognitive procedure which is cognitively least effortful in a given context. Which procedure is the cognitively least effortful one for an individual in a given context depends on that individual's cognitive capacities. As already indicated, studies in other domains have shown that fluency plays an important role in determining which cognitive procedure individuals use to solve a mental task, and we appeal to future research on fluency in the domain of social cognition. Moreover, we appeal to future research on the role what we have called 'associations with person-specific identity' in social cognition, i.e. associations of particular person-specific characteristics (character traits, habits etc.) with familiar individuals that dispose them to particular behaviors in specific contexts. On our view, these associations play an important and widely neglected role for social understanding in everyday life where people are often familiar with the people whose behavior they attempt to understand.

In general, given that reason explanations capture just one part of social cognition (Malle *et al.*, 2007), a novel focus of future research might be on how the offering of causal history explanations develops ontogenetically. Finally, fine-grained experimental designs might contribute to determining under which circumstances people offer which kinds of explanation, and which cognitive procedures are most likely being used to generate the explanation in question.

Future research may investigate the role of associations$_{pi}$ in behavior expectations, predictions, and explanations. Our pluralistic approach to the explanation of social understanding predicts that people will tend to refer to an agent's dispositions on the basis of linguistic associations$_{pi}$ in causal history explanations if such associations are available (this presupposes that the agent whose behavior needs to be explained is familiar to the observer), instead of offering a reason explanation by means of theorizing or simulating processes. Future research may show whether people prefer to consider dispositions on the basis of associations$_{pi}$ rather than associations$_{si}$ in behavior explanations when people are familiar to the agent and whether this is due to a hierarchy of cognitive accessibility among the different kinds of associations. In

general, we hypothesize that nonlinguistic associations are cognitively 'cheaper' than linguistic ones and that using any kind of association is 'cheaper' than theorizing or simulating processes. We appeal to future research that could test this hypothesized taxonomy.

Whereas linguistic associations that are part of an 'other-concept' may be methodologically determined by questionnaire research, the non-linguistic associations that an individual has concerning another person are observable (and hence measurable) in the spontaneous behavioral interactive responses the individual exhibits towards that other person. Notably, studies on stereotype activation have shown that what we have called non-linguistic mentalistic associations$_{si}$ and linguistic mentalistic associations$_{si}$ (being part of a social identity concept, short: concept$_{si}$) may dissociate. That is, interactive responses relying on automatically activated stereotyping processes (i.e. non-linguistic associations$_{si}$) may diverge from the perceiver's social identity concept (i.e. linguistic associations$_{si}$). If perceivers become aware of that divergence, they may be able to override undesirable implicit processes. For example, implicit race prejudice such as a fear response when facing a black person, indicated by increased neural activity in the amygdala, does not correlate with measures of explicit, conscious race prejudice (Phelps *et al.*, 2000) and is reduced the longer the black face is presented to the participant (Cunningham *et al.*, 2004). Since in the latter study the magnitude of activity in right dorsolateral prefrontal cortex and anterior cingulate cortex predicted how much the amygdala activity would be reduced for the long presentations, these brain regions have been associated with explicit, deliberate processes aimed at reducing the undesirable prejudicial response. Such a conclusion is consistent with the proposal (Coltheart, 2010) that right dorsolateral prefrontal cortex is specifically associated with cognitive hypothesis evaluation, an explicit, deliberate cognitive operation. In the two-factor model of delusional belief (Coltheart *et al.*, 2011), the second factor needed for delusions to arise is an impairment of cognitive hypothesis evaluation, and lesions of right dorsolateral cortex have often been identified in delusional patients, supporting this claim about one function of right dorsolateral prefrontal cortex. These findings suggest that associations$_{si}$ typically become part of a behavior explanation just when the perceiver's recognition of the agent's character traits does *not* diverge from the perceiver's concept$_{si}$. We assume that the same holds true for associations$_{pi}$.

Finally, we propose that fluency and associations$_{pi}$ are not only worth investigating in future research but that what has been found in social psychology concerning trait association and fluency may also allow for alternative interpretations of the findings of contemporary research such as different versions of so-called 'false belief tasks'. More precisely, we propose that fluency plays a central role in 4- to 5-year-olds theorizing processes in explicit (i.e. linguistic) versions of the false belief task, whereas associations$_{pi}$ come into play when preverbal infants perform successfully in implicit (i.e. nonlinguistic) versions of the false belief task.

4.1 Fluency in Explicit False Belief Tasks

A widely shared assumption is that the best indicator of understanding other people's beliefs is the understanding of other people's *false* beliefs. As pointed out by Wellman *et al.* (2001, p. 655), 'mental-state understanding requires realizing that such states may reflect reality and may be manifest in overt behavior, but are nonetheless internal and mental, and thus distinct from real-world events, situations, or behaviors'. Belief Reasoning (BR) involves the capacity to build meta-representations, i.e. BR enables children to understand other people's epistemic states as representations that may diverge from the own representations and include the risk of failing. Moreover, children can predict or explain other people's behavior by BR on the basis of a variety of folk psychological rules.

Since the early 1980s a variety of explicit (i.e. linguistic) versions of the so-called 'false belief task' have been conducted to investigate at which age BR is acquired. In one version of the false location task, children were presented with a story in which the protagonist Maxi puts a chocolate bar into cupboard x. Then Maxi leaves the room. In his absence, Maxi's mother moves the chocolate bar from cupboard x to cupboard y. Then Maxi returns. The children were asked to indicate the cupboard where Maxi will look for the chocolate bar. Whereas 3-year-olds nominate cupboard y where they themselves believe the chocolate bar to be located, 5-year-olds typically point correctly to cupboard x, exhibiting an explicit (i.e. verbal) understanding of Maxi's false belief (see Wellman *et al.*, 2001 for a review).

In contrast, children's use of BR in the true belief task has been regarded as being ontogenetically prior to children's use of BR in the false belief task,

However, recent findings from developmental studies challenge this view and indicate that in true belief tasks children younger than age 6 make use of procedures other than BR: whereas 3-year-old children have been found to simply refer to their own (rather than the other person's) belief when predicting another person's behavior (call this 'Reality Reasoning' (RR)), 5-year-old children make use of a reasoning procedure, called Perceptual Access Reasoning (PAR), that relies on two rules of thumb and only requires to account for the perceptual access of the agent in the given situation, which is cognitively less demanding than BR (Fabricius *et al.*, 2010).

In line with research on fluency and cognitive dissonance, we propose that 5-year-old children experience cognitive dissonance (i.e. a feeling of discomfort when simultaneously holding conflicting beliefs) in reasoning processes in false belief tasks in which their own belief (i.e. persistent epistemic state) differs from that of the agent, but not in true belief tasks, in which there is no such difference. We postulate that this experience of cognitive dissonance is accompanied by an experience of cognitive strain that results in 5-year-old children having recourse to higher cognitive levels of reasoning. That is, 5-year-old children may understand other people's behavior by means of BR if they experience cognitive strain (such as in false belief tasks) but they revert to simpler heuristics (PAR) when such an experience is missing, such as in true belief tasks (see Fiebich, 2014, for a discussion of this hypothesis). Notably, this hypothesis is empirically testable by replicating a modified version of Fabricius *et al.*'s true belief task that includes additional factors

that induce cognitive strain. If 5-year-olds pass such modified versions, this would indicate that cognitive strain places them indeed on higher-cognitive levels such as BR.

4.2 Associations$_{pi}$ and Teleological Reasoning in Implicit False Belief Tasks

In a landmark study, Onishi and Baillargeon (2005) simplified the explicit (i.e. linguistic) version of the false location task by using looking time as a measure of preverbal infants' implicit (i.e. nonlinguistic) false-belief understanding. Onishi and Baillargeon regarded infants' longer looking-times on trials in which the agent searches for the object where he or she has not seen it hidden before as a sign of infants' surprise and as evidence that infants' expectations about the agent's behavior based upon an implicit false belief understanding were violated ('violation-of-expectation task'). This study was followed by a large number of related studies using various measurements (e.g. gaze direction in anticipatory looking tasks) to access preverbal infants' implicit understanding of another person's false belief about an object's location, content, or identity (e.g. He *et al.*, 2011; Scott and Baillargeon, 2009; Southgate *et al.*, 2007).

Following Apperly and Butterfill (2009), we assume that the cognitive processes that allow preverbal infants to pass implicit versions of the false belief task rely on System 1, whereas System 2 comes into play in explicit false belief tasks in which children need to generate a verbal behavior prediction. As indicated in Section 2, we propose that System 1 is modular but flexible insofar as it continuously integrates novel information from System 2. That is, what may have been inferred once via an effortful reasoning process (System 2) may be automatically accessible (System 1) on a later occasion. For example, 4- to 5-year-olds may initially only engage in BR when they experience cognitive strain (such as in explicit false belief tasks). 6-year-olds, however, who became experienced in BR, may use BR with ease and employ this procedure also in situations that do not induce cognitive strain (such as in explicit true belief tasks) (see Fiebich, 2014, for a discussion).

Moreover, we follow Butterfill and Apperly (2013) in proposing that preverbal infants' expectations of the agent's behavior in implicit false belief tasks rely on a set of teleological reasoning principles. In addition, we propose that associations$_{pi}$ play a central role for encoding the agent's action goal. As spelled out in Section 3.1, infants as young as 6 months old generate associations$_{pi}$ when they observe an agent reaching repeatedly for a particular object. When successfully performing in implicit false belief tasks, 15-month-olds continue to expect an agent to reach for the particular object he or she reached for before. Crucially, in *all* the implicit false belief tasks, infants observe the agent reaching for and acting on the object placed at a particular location before infants expect the agent's subsequent behavior (and in some studies, the agent even placed the object at a particular location himself or herself, see e.g. Clements and Perner, 1994; Buttelmann *et al.*, 2009).

Once the infant has encoded the agent's action goal on the basis of association$_{pi}$, the infant expects the agent to pursue that goal in line with a number of teleological reasoning principles. The results of implicit false belief tasks clearly show that preverbal infants are capable of tracking an agent's persistent epistemic states (i.e. beliefs or belief-like states) and that infants expect an agent to pursue his or her goal in conformity with such states; one could call this the 'epistemic-congruency principle'. However, whether these results are evidence for preverbal infants' implicit *false* belief understanding remains an open question. In our account, independent of whether the agent's epistemic states are false or not, the infant expects, on the basis of nonlinguistic associations$_{pi}$, that the agent will pursue his or her goal in congruency with that agent's epistemic states (see Fiebich, 2015, for a discussion).

5. Summary and Discussion

In this article we have proposed a pluralistic approach to the explanation of social understanding according to which individuals make use of a variety of strategies when attempting to understand another person's behavior. This approach makes its contribution to the contemporary literature by integrating findings from the theory of mind literature and social psychology literature. Apart from theory and simulation, we have highlighted the roles that nonlinguistic and linguistic associations$_{pi}$ and associations$_{si}$ may play in social understanding. Methodologically, we have distinguished behavior expectations that rely on a variety of automatically activated associations and are indicated by spontaneous bodily gestures from behavior predictions and behavior explanations that rely on structured thoughts and may be verbally uttered. In addition, we have distinguished understanding another person as an intentional agent who acts on dispositions from understanding another person as a mental agent who acts on reasons. Whereas determining an agent's dispositions in behavior understanding typically relies on a variety of associations, attributing beliefs and desires to an agent relies on theorizing or simulating processes.

Our approach needs to be distinguished from other pluralistic approaches to the explanation of social understanding in the literature. Andrews (2012), for example, has provided a sophisticated pluralistic approach that resembles ours in various respects. Reviewing a number of empirical studies, Andrews highlights (as we have done) that on many occasions individuals understand the behavior of another person not by attributing beliefs and desires to the other person but instead by taking into account the person's character traits and history, social roles and rules, the situation the individual is engaged in, etc. Moreover, we agree with Andrews that preverbal infants' sensitivity towards other people's epistemic states in implicit false belief tasks is not evidence for an (implicit) representational understanding of the other person's epistemic states (including a grasp of the potential falseness of these states) but rather evidence for an understanding of these states as person-specific dispositions (see Andrews, 2012, pp. 32 ff. for a discussion). It is not possible at this point to highlight the many points in which we agree with Andrews. Most crucially, both

our approaches reject the assumption that is shared between TT and ST that there is a default procedure that individuals use whenever they are trying to understand another person's behavior. However, although pointing to a variety of procedures that may come into play in behavior understanding, Andrews provides no dedicated hypothesis concerning why an individual may use one procedure rather another one in a given context. Our approach is capable of filling this gap. By appealing to findings from social psychology, we stress that—as a rule of thumb—individuals make use of the cognitive procedure that is cognitively least effortful.

Following Kahneman (2011), we propose that people, when seeking behavior explanations, have a general tendency to use procedures that rely on fast cognitive processes that require little or no mental effort. Our pluralistic approach to explaining how social understanding is achieved predicts that when offering behavior explanations people tend to offer 'causal history explanations' by means of cognitively 'cheap' associations (if available) that help determining the agent's disposition to rather than cognitively more 'costly' theorizing or simulating activities. Moreover, we propose that, whatever kind of explanations people offer, when they make use of theory or simulation they do so in the least cognitively-effortful manner, i.e. they are prone to employ 'intuitive heuristics' (System 1). For example, when they make use of folk psychological rules in a situation that allows for psychological generalizations, people may use fewer rules and take into account fewer mental states of the agent than might have been required for adequate prediction or explanation of the agent's behavior. If people run into difficulties when offering causal history explanations or reason explanations, in contrast, they are likely to employ thoroughly deliberatively effortful and conscious processes (System 2) of theorizing or simulating.

Of course, narrative and communicative practices also play a crucial role in understanding other minds. Indeed, when we wonder why Laura is going to India, the most secure route for figuring out her reasons for doing so is to accept the reasons she gave us in a direct conversation, even though she may be lying, or self-deceived (Gallagher and Hutto, 2008). There is no need any more to predict or explain another person's behavior by offering 'reason explanations' if that very person simply tells us why she has acted in a certain way, or how she is planning to act in the future.

In our account, the experience of cognitive effort plays a (if not *the*) central role in an individual's choice to use one cognitive procedure over another. That is, people typically rely on cognitively 'cheap' procedures in behavior judgments and only revert to more effortful procedures if they experience cognitive strain (Kahneman, 2011). Notably, however, factors other than fluency may come into play. As pointed out by Smith *et al.* (2006), people may also engage in thoughtful reasoning processes in their judgments about people or objects, 'when people possess both high ability (e.g. adequate information, freedom from distraction) and high motivation (e.g. due to the object's self-relevance)' (p. 3). Moreover, good performance in the Cognitive Reflection Task has been found to correlate with people's preferring and generating explanatory deep judgments (Fernbach *et al.*, 2012). Finally, apart from information access 'impression management' (i.e. painting a flattering picture of another person as a rational agent) may also determine whether people offer reason explanations or

causal history explanations (Malle *et al.*, 2007). If several associations are available, the pragmatic context has been found to determine which association is activated; for example, perceiving an Asian woman applying cosmetics, or eating with chopsticks, is sufficient to activate either the gender, or the race stereotype (Macrae *et al.*, 1995).

To sum up, according to our pluralistic 2-System-approach to social cognition, people typically use that cognitive procedure that is experienced as cognitively least effortful in a given context. Apart from theory and simulation, we have emphasized the role that different kinds of associations play in social cognition. Our approach has the advantage over other approaches insofar as it accounts for the important roles of fluency and person-specific knowledge that have been neglected so far in the debate.

Institute of Philosophy II
Ruhr-University Bochum

Department of Cognitive Science
Macquarie University Sydney

References

Andrews, K. 2012: *Do Apes Read Minds?* Cambridge, MA: MIT Press.

Apperly, I., Riggs, K., Simpson, A., Chiavarino, C. and Samson, D. 2006: Is belief reasoning automatic? *Psychological Science*, 17(10), 841–44.

Apperly, I. and Butterfill, S. 2009: Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116(4), 953–70.

Baillargeon, R., Scott, R.M. and He, Z. 2010: False-belief understanding in infants. *Trends in Cognitive Science,* 14(3), 110–8.

Baron-Cohen, S. 1995: *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Bessenoff, G.R., and Sherman, J.W. 2000: Automatic and controlled components of prejudice toward fat people: evaluation versus stereotype activation. *Social Cognition*, 18, 329–53.

Borg, E. 2007: If mirror neurons are the answer, what was the question? *Journal of Consciousness Studies*, 14, 5–19.

Borkenau, P. and Liebler, A. 1992: Trait inferences: sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645–57.

Brochu, P.M. and Morrison, M.A. 2007: Implicit and explicit prejudice toward overweight and average-weight men and women: testing their correspondence and relation to behavioral intentions. *Journal of Social Psychology*, 147, 681–706.

Buttelman, D., Carpenter, M. and Tomasello, M. 2009: Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112, 337–42.

Butterfill, S. and Apperly, I. 2013: How to construct a minimal theory of mind. *Mind & Language*, 28, 606–37.

Carruthers, P. 2013: Mindreading in infancy. *Mind & Language*, 28, 141–72.

Clements, W.A. and Perner, J. 1994: Implicit understanding of belief. *Cognitive Development*, 9, 377–395.

Coltheart, M. 2010: The neuropsychology of delusions. *Annals of the New York Academy of Sciences*, 1191, 16–26.

Coltheart, M., Langdon, R. and McKay, R.T. 2011: Delusional belief. *Annual Review of Psychology*, 62, 271–98.

Csibra, G. and Gergely, G. 2007: 'Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124, 60–78.

Cunningham, W.A., Johnson, M.K., Raye, C.L., Gatenby, J.C., Gore, J.C. and Banaji, M.R. 2004: Separable neural components in the processing of black and white faces. *Psychological Science*, 15, 8068–13.

Eagly, A.H. and Steffen, V.J. 1984: Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46, 735–54.

Fabricius, W.V., Boyer, T., Weimer, A.A. and Carroll, K. 2010: True or false: do five-year-olds understand belief? *Developmental Psychology*, 46, 1402–16.

Fernbach, P.M., Sloman, S.A., Louis, R. and Shube, J.N. 2012: Explanation fiends and foes: how mechanistic detail determines understanding and preference. *Journal of Consumer Research*, 39(5), 1115–31

Fiebich, A. 2014: Mindreading with ease? Fluency and belief reasoning in 4- to 5-year-olds. *Synthese*, 191, 929–44.

Fiebich, A. 2015: *Varieties of Social Understanding*. Paderborn: mentis Verlag GmbH.

Frith, C.D. and Frith, U. 2008: Implicit and explicit processes in social cognition. *Neuron*, 60(3), 503–10.

Gallagher, S. 2001: The practice of mind: theory, simulation, or interaction? *Journal of Consciousness Studies*, 5–7, 83–108.

Gallagher, S. 2007: Simulation trouble. *Social Neuroscience*, 2(3/4), 353–65.

Gallagher, S. and Hutto, D. 2008: Understanding others through primary interaction and narrative practice. In J. Zlatev, T. Racine, C. Sinha and E. Itkonen (eds), *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam: John Benjamins, 17–38.

Goldman, A. 2006: *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.

Gopnik, A. 1998: The scientist as child. In A. Gopnik and A. Meltzoff (eds), *Words, Thoughts, and Theories*. Cambridge. MA: MIT Press, 13–47.

Gordon, R. 1996: 'Radical' simulationism. In P. Carruthers and P.K. Smith (eds), *Theories of Theories of Mind*. Cambridge: Cambridge University Press, 11–21.

He, Z., Bolz, M. and Baillargeon, R. 2011: False-belief understanding in 2.5-year-olds: evidence from violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*, 14(2), 292–305.

Kahneman, D. 2011: *Thinking, Fast and Slow*. Harmondsworth: Penguin Books.

Kelly, D., Quinn, P., Slater, A.M., Lee, K., Gibson, A., Smith, M., Ge, L. and Pascalis, O. 2005: Three-month-olds, but not newborns, prefer own-race faces. *Developmental Science*, 8, F31–F36.

Lin, M.H., Kwan, V.S.Y., Cheung, A. and Fiske, S.T. 2005: Stereotype content model explains prejudice for an envied outgroup: scale of Anti-Asian American stereotypes. *Personality and Social Bulletin*, 31, 34–47.

Lu, H.P. and Hsiao, K.L. 2010: The influence of extro/introversion on the intention to pay for social networking sites. *Information & Management*, 47(3), 150–7.

Luo, Y. and Johnson, S.C. 2009: Recognizing the role of perception in action in at 6 months. *Developmental Science*, 12, 142–9.

Macrae, C.N. and Bodenhausen, G.V. 2000: Social cognition: thinking categorically about others. *Annual Review of Psychology*, 51, 93–120.

Macrae, C.N., Bodenhausen, G.V. and Milne, A.B. 1995: The dissection of selection in person perception: inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology*, 69, 397–407.

Macrae, C.N. and Quadflieg, S. 2010: Perceiving people. In S. Fiske, D.T. Gilbert and G. Lindzey (eds), *The Handbook of Social Psychology*, 5th edn. New York: McGraw-Hill, 428–63.

Malle, B.F., Knobe, J.M. and Nelson, S.E. 2007: Actor-observer asymmetries in explanations of behavior: new answers to an old question. *Journal of Personality and Social Psychology*, 93(4), 491–514.

Mason, M.F. and Macrae, C.N. 2004: Categorizing and individuating others: the neural substrates of person perception. *Journal of Cognitive Neuroscience*, 16(10), 1785–95.

Nyström, P. 2008: The infant mirror neuron system studied with high density EEG, *Social Neuroscience*, 3(3–4), 334–347.

Onishi, K.H. and Baillargeon, R. 2005: Do 15-month-old infants understand false beliefs? *Science*, 308, 255–8.

Oppenheimer, D.M. 2008: The secret life of fluency. *Trends in Cognitive Sciences*, 12(6), 2372–41.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C. and Fabre-Grenet, M. 1995: Mother's face recognition by neonates: a replication and an extension. *Infant Behavior and Development*, 1, 79–85.

Perner, J. 1999: Metakognition und Introspektion in entwicklungspsychologischer Sicht: Studien zur 'Theory of Mind' und 'Simulation'. In W. Janke and W. Schneider (eds), *Hundert Jahre Institut für Psychologie und Würzburger Schule der Denkpsychologie*. Göttingen: Hogrefe, 411–431.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., and Banaji, M. R. 2000: Performance on indirect measures of race evaluation predicts amygdale activation, *Journal of Cognitive Neuroscience*, 12(5), 729–738.

Puhl, R. M., Moss-Racusin, C. A., Schwartz, M. B. and Brownell, K. D. 2008: Weight stigmatization and bias reduction: perspectives of over-weight and obese adults. *Health Education Research*, 23, 347–58.

Scott, R. M. and Baillargeon, R. 2009: Which penguin is this? Attributing false beliefs about identity at 18 months. *Child Development*, 80, 1172–96.

Shanton, K. and Goldman, A.I. 2010: Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Sciences*, 1(4), 527–38.

Southgate, V., Senju, A. and Csibra, G. 2007: Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), 587–92.

Smith, E.R., Miller, D.A., Angela, T.M., Crump, S.A., Garcia-Marques, T. and Mackie, D.M. 2006: Familiarity can increase stereotyping. *Journal of Experimental Social Psychology*, 43(4), 471–8.

Tomasello, M. 1999: *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

Wellman, H.M., Cross, D. and Watson, J. 2001: Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, 72, 655–84.

White, G.L. and Shapiro, D. 1987: Don't I know you? Antecedents and social consequences of perceived familiarity, *Journal of Experimental Social Psychology*, 23, 75–92.

Woodward, A.L. 1998: Infants selectively encode the goal object of an actor's reach. *Cognition*, 69, 1–34.

Woodward, A. L. 1999: Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, 22(2), 145–60.