Dual Processes in Mental State Understanding:
Is Theorising Synonymous with Intuitive Thinking and is Simulation Synonymous with Reflective Thinking?

Meredith R Wilkinson (mwilkinson@dmu.ac.uk)
Division of Psychology, School of Applied Social Sciences, Faculty of Health and Life Sciences, De Montfort University, The Gateway, Leicester, LE1 9BH, UK

Linden J. Ball (LBall@uclan.ac.uk)
School of Psychology, University of Central Lancashire, Darwin Building, Preston, Lancashire, PR1 2HE, UK

Abstract
In this paper we develop an idea first mooted by Wilkinson, Ball, and Cooper (2010), which is that the dichotomy between theory-based and simulation-based reasoning in the context of mental state understanding is synonymous with the distinction between intuitive and reflective thinking in dual-process accounts of human reasoning (e.g., Evans, 2010). To support this proposal we draw upon a range of concepts and findings deriving from both mainstream reasoning research and from studies of social cognition. We also consider the implications of our proposal for the formulation of an integrative approach to understanding reasoning in all of its many manifestations, whether undertaken for the attainment of socially-oriented goals or for the purposes of learning and discovery.

Keywords: Dual Processes, Intuitive thinking, Reflective thinking, Simulation Theory, Theory Theory

Introduction
The question of how we understand and reason about other people’s minds has resulted in considerable debate within psychology and philosophy (e.g., Bach, 2011; Wilkinson, Ball, & Cooper, 2010). Some researchers propose that such mental state reasoning is achieved through the adoption of tacit and non-tacit “theories” that are typically based around conditional inference rules (e.g., Carruthers, 1996). An example of such a theory might be that if someone fails to achieve something for which they have worked hard then they will feel upset. Other researchers, however, posit that mental state reasoning arises via a process of mental simulation. Such simulation might involve imagining how we would feel in a given situation and assuming that others are sufficiently like us that they will feel the same (e.g., Gordon, 1986). An alternative proposal (Goldman, 2006) is that we take our own beliefs and desires “offline”, input the beliefs and desires of the other person, and hence reason as if we had the beliefs and desires of the other person. It is important to note that when we refer to “simulation” in the present paper we adopt a restricted notion that relates solely to the simulation of mental states, whether our own or other people’s. We acknowledge that simulation can arise when reasoning in other domains such as design (e.g., see Ball & Christensen, 2009), but we do not extend our discussion to this issue. Second, we note that the term “simulation” has different meanings to different authors. As explained by Goldman (2006), there is high-level simulation, which refers to the type of simulation we are discussing in this paper, and also lower-level simulation, which refers to the functioning of mirror neurons when engaged in activities such as imitation. Mirror neurons are neurons that are activated both when we perform an action and when we observe the same action being performed (e.g., Gallese & Goldman, 1998). Although we acknowledge the evidence for the existence of mirror neurons, we, like others (e.g. Saxe, 2005) are unsure of the explanatory power of this form of simulation.

Recently, theorists have started to move away from polarised views as to whether theorising or simulation is adopted in mental state reasoning and have instead acknowledged that both processes may be at play. This has resulted in a flurry of hybrid approaches appearing in the literature (e.g., Bach, 2011; Mitchell, Currie, & Ziegler, 2009), which not only propose that both theorising and simulation can occur in mental state reasoning, but which also claim that there are content-based effects that govern the mechanism that is triggered. For example, Mitchell et al. (2009) have argued that we deploy simulation as a “default” process, using theorising in familiar situations. We have recently provided empirical support for a hybrid view in a study that required people to think aloud when reasoning about counterfactual scenarios pertaining to mental states (Wilkinson et al., 2010). Participants adopted both theorising and simulation for these scenarios, with content effects being evident in that more simulation and less theorising arose with scenarios involving “controllable” compared to “uncontrollable” events.

In the present paper we extend an argument first presented by Wilkinson et al. (2010) to the effect that the theorising versus simulation distinction is synonymous with the “intuitive” versus “reflective” distinction as described in contemporary dual-process theories of thinking and reasoning (see Evans, 2010, for an overview). According to the dual-process framework, intuitive thinking is classed as fast, automatic, high capacity, low effort and independent of working memory resources, whereas reflective thinking is classed as slow, controlled, low capacity, high effort and dependent on working memory. We argue here that these characteristics of intuitive and reflective thinking align well with key features of theorising and simulation in contexts associated with mental state understanding. In subsequent sections we support this proposal by drawing on concepts
and findings from contemporary reasoning research and from studies of social cognition. We suggest that conceptualising theorising and simulation in a dual-process manner has the potential to enable researchers to move towards a more compelling account of mental state reasoning that can be subjected to rigorous empirical examination. We conclude the paper by addressing issues that researchers might wish to consider further if they find merit in our proposed dual-process conceptualisation of the processes underpinning mental state understanding.

**Proposed Parallels: Theorising as Intuition; Simulation as Reflection**

Evans (2010) describes multiple distinctions between intuitive and reflective reasoning and we take these distinctions as a foundation for demonstrating how one can arbitrate between theorising and simulation within a dual-process framework. The first distinction that Evans notes is that intuitive reasoning is fast whereas reflective reasoning is slow. We similarly propose that theorising is a fast process whereas simulation is slower. This in turn is linked to the cognitive effort required for theorising and simulation. Like intuitive reasoning, which Evans proposes involves low cognitive effort, we view theorising as being low effort compared to simulation, which is high effort, much like reflective reasoning. Evans further argues that intuitive reasoning is high capacity and reflective reasoning is low capacity. We contend that the same holds for theorising and simulation, respectively. In addition, for people to engage in theorising they need to have a store of pre-existing theories pertaining to others' mental states, with these theories being drawn upon in an automatic manner when primed by particular contexts. Simulation, however, takes the form of a concurrent and incremental reasoning process (Goldman, 2006), which will require more controlled than automatic processing. Again, this distinction parallels the notion that intuitive reasoning is automatic whereas reflective reasoning is controlled. Evans has made further claims concerning the links between intuitive versus reflective reasoning and working memory. He argues that reflective reasoning is dependent upon working memory resources whereas intuitive reasoning is independent of such resources. We propose that this distinction holds for theorising and simulation too, with simulation being highly dependent upon working memory and executive functioning (e.g., Currie, 1996; Goldman, 2006).

**Empirical Evidence for the Proposed Parallels**

A robust finding in the reasoning literature concerns the phenomenon of “belief bias”, which is typically studied in relation to people’s abilities at syllogistic inference. Within a standard conclusion-evaluation paradigm participants are presented with two premises that they should assume are true and an associated conclusion. They are then required to determine whether the conclusion follows logically from the premises. Many studies have shown that participants are biased by the conclusion’s believability when making evaluations, rather than reasoning on the basis of the conclusion’s validity (e.g., see Stupple, Ball, Evans, & Kamal-Smith, 2011).

Numerous dual-process accounts have been forwarded as to why belief-bias occurs (see Ball, 2011, for a review). For the purposes of our argument, however, we draw on the “selective processing model” of Evans (e.g., 2000), itself an example of a more general class of dual-process models referred to as “default-interventionist” theories (Evans, 2007). According to the selective processing model of belief bias, intuitive reasoning cues a response that may or may not be overridden by a reflective process. The default, intuitive response is to accept or reject conclusions based solely on their believability. If, however, reflective reasoning is applied then this reasoning is influenced by the conclusion’s belief status such that participants will search for confirming models when a conclusion is believable and for disconfirming models when it is unbelievable. Whether a logically correct evaluation ensues for a problem is, therefore, dependent on the interplay between the intuitive and reflective processes, with certain problems (e.g., those with invalid but believable conclusions) being especially difficult because the belief status of the conclusion biases both the default response and the confirmation-oriented reflective response (Stupple et al., 2011).

We propose that in tasks of mental state reasoning people can be similarly biased by their personal beliefs. This is demonstrated by the so-called “curse of knowledge”, whereby participants are unable to pass false belief tasks because they cannot inhibit viewing a situation from their own perspective (e.g., Birch & Bloom, 2007). In such tasks (e.g., Baron-Cohen, Leslie, & Frith, 1985) participants (typically young children) are introduced to two protagonists in a room, both of whom are aware of a particular state of affairs, such as a marble in a basket. Then one protagonist leaves the room and the remaining protagonist moves the marble to a box. The participant is asked, “Where will the protagonist who left the room look for the marble upon returning?” If the individual is able to reason about another person’s beliefs then they should state that the protagonist will look for the marble in the basket. Individuals who fall foul of the curse of knowledge will respond by saying that the protagonist will look in the box (where they themselves know the marble is currently located), demonstrating a form of belief-biased reasoning. We suggest that overcoming this bias, especially when encountering such a situation for the first time, requires the deployment of a controlled process of mental simulation in which the reasoner takes their own beliefs off-line and reasons from the beliefs of the protagonist (e.g., Mitchell et al., 2009). This is equivalent, we propose, to the way that people can engage in reflective reasoning in an effort to overcome belief bias in syllogistic reasoning (Stupple et al., 2011), although, as noted above, even reflective reasoning does not guarantee success since it may itself be biased.

We now return to Evans’ (2010) description of the characteristics of intuitive versus reflective reasoning in
order to assess the evidence for the proposed parallels between dual-process views and the theorising/simulation distinction in mental state reasoning. Evans argues that intuitive reasoning is fast whereas reflective reasoning is slow, a view that we propose aligns well with the theorising/simulation distinction. Evidence for our claim comes from Atkinson, Bell, and Feeney (2009), who examined the influence of a speeded-response requirement on how participants reasoned about counterfactual scenarios that were constructed to tap into two robust effects: (1) the “action effect”, which is a tendency to regret action more than inaction in the short term (e.g., Kahneman & Tversky, 1982), with the reverse being the case in the long term (e.g., Gilovich & Medvec, 1994); and (2) the “temporal order effect” (e.g., Byrne, Segura, Culhane, Tasso, & Berrocal, 2000), which is the tendency to attribute more negative affect to the person committing the final act in a sequence of actions when a negative outcome occurs. Atkinson et al. (2009) asked participants to reason about the presented scenario, which entailed reading a vignette describing two agents in a negative situation, either in a speeded condition, in which they had to answer as quickly as possible, or in a non-speeded condition, in which they were able to take as long as they wished. Whereas the temporal order effect was unaffected by the speed manipulation, the action effect was disrupted, with the actor selected significantly less in the speeded compared to the non-speeded condition. In relation to the temporal order manipulation we propose that participants uniformly access a “theory” that if a person acts last in a sequence of events leading to a negative outcome then they will feel worse. However, in the case of the action effect, we propose that participants need to run simulations of the mental states of both protagonists to evaluate the interplay between action/inaction and the time that events arose. Such an evaluation process is time-consuming relative to accessing a pre-stored theory, which would explain why the speeded-response requirement only affects the action effect and not the temporal order effect. These findings further demonstrate how a simulation may overturn an initial theory. The fact that the actor was chosen less often in the speeded compared to the non-speeded condition suggests that people held an initial theory that the non-actor would feel more regret, but when afforded the time to run simulations they could overturn this initial response, much as reflective reasoning can overturn an initial belief-biased response in syllogistic inference (Stupple et al., 2011).

The fact that simulation seems to be slower than theorising also speaks to likely discrepancies in the cognitive effort required for these reasoning types. We propose that theorising is low effort, like intuitive reasoning, whereas simulation is high effort, like reflective reasoning. As such, simulation will be dependent on general cognitive resources, including executive functioning and cognitive inhibition (e.g., Currie, 1996; Goldman, 2006) whereas theorising will work independent of such mechanisms. In addition, simulation will be dependent upon working memory whereas theorising will not. Evidence in support of this claim comes from a study using the “director task” (Lin, Keysar, & Epley, 2010), where participants are presented with a grid that contains slots that can be seen by both themselves and a director, who is actually a confederate. Some items, however, are only visible to the participant, since they are occluded from the director’s view. Participants are instructed what object to move. On critical trials the perspective between director and participant differs so the director may say “move the small mouse” when there are three mice and only the smallest one can be seen by the participant. This requires the participant to engage in simulation by shifting their perspective to that of the director’s in order to fulfill the instruction correctly.

Importantly, Lin et al. (2010) found that participants with higher working memory capacities performed better on the director task than those with lower working memory capacities. This provides a link between simulation and working memory that bears strong similarities to the link between reflective reasoning and working memory. For example, De Neys (2006) has shown that individuals with greater working memory resources perform better on belief-oriented syllogistic reasoning tasks, and Stanovich, West, and Toplak (2011) have argued that reflective reasoning is dependent upon executive functioning resources. We propose that just as the intuitive/reflective distinction is associated with differential involvement of working memory, so too is the theorising/simulation distinction.

Theorists proposing dual-process accounts of reasoning have also recently begun to draw upon neuroscientific evidence to support their claims. For example, Goel (2003) has presented evidence for “dual pathways” in syllogistic reasoning, with intuitive processes associated with the frontal-temporal pathway and reflective processes associated with the parietal pathway. In a review article, Goel (2007) acknowledges that the question of which neural regions are responsible for particular types of processing is one that has demonstrated differing findings, but that the evidence nevertheless points towards a fractionated system for deductive reasoning rather than a unitary one.

Neuroscientific evidence for theorising and simulation in mental state reasoning is equally complex and has been criticised for failing to provide a clear differentiation between brain regions specialised for such processing (e.g., Apperly, 2008; Wilkinson & Ball, 2012). Nevertheless, findings are suggestive. For example, Mitchell, Banaji, and Macrae (2005), found that the ventral medial prefrontal cortex was activated when participants made judgements concerning facial expressions. This provides evidence of a partial locus for theory-based reasoning, since an intuitive judgement is all that would be required for this task, with no simulation being necessitated. We believe that our claims for the intuitive/theorising parallels here are strengthened by the observation that this brain region is also known to be activated in syllogistic reasoning tasks when participants provide belief-biased responses (e.g., Goel & Dolan, 2003).

As for simulation in mental state reasoning, it is admittedly not easy to locate a specific brain region.
responsible for such processing. Part of the difficulty originates from the differing conceptualisations of simulation within the literature, with some theorists (e.g., Gallese & Goldman, 1998) arguing for lower-level simulation (Goldman, 2006). Studies of higher-level simulation in paradigms such as the director task suggest the involvement of a number of brain regions, including the superior dorsal medial prefrontal cortex and the superior/middle temporal sulci extending to the extrastriate body area and the posterior superior temporal sulcus (Dumontheil, Küster, Apperly, & Blakemore, 2010). Admittedly, the evidence for the localisation of simulation is not clear-cut, but again is suggestive of distinct brain regions being associated with theorising and simulation.

![Diagram of hybrid dual-process model](image)

**Figure 1:** A schematic representation of a hybrid dual-process model of theorising and simulating.

**Issues Arising from the Proposed Parallelism**

In the previous section we outlined how Evans’ (2010) distinction between intuitive and reflective thinking can map onto the concepts of theorising and simulation in mental state reasoning. However, there is still much for theorists to consider if they see benefits in examining such apparent parallels more extensively. In this respect we note that reasoning researchers have now begun to move away from normative accounts of human reasoning towards a descriptivist agenda (e.g., Elqayam & Evans, 2011). We generally support this approach and suggest that using theorising and simulation within a descriptivist dual-process framework affords an opportunity to develop a rich and innovative programme of empirical and theoretical research. However, several questions need to be borne in mind when pursuing such a project, which we consider below.

A first, critical question is this: exactly how do theorising and simulation function within a dual-process framework? The reasoning literature contains both sequential and parallel dual-process models of phenomena such as belief bias. Sequential models of the default-interventionist variety propose that intuitive reasoning generates a default response and that reflective reasoning serves either to confirm or override this initial judgement (e.g., Evans, 2006). Parallel models (e.g., Sloman, 2002; Stupple & Ball, 2008) propose that intuitive and reflective reasoning compete in generating a response. Evans (2009) has also proposed a “hybrid” dual-process model, which combines sequential and parallel processes. We think it likely that theorising and simulation can operate both sequentially and in parallel such that a hybrid model may capture key subtleties most effectively (see Figure 1). We propose an initial stage of pre-attentive theorising, whereby representational structures such as scripts and schemas are attended to before a decision is made to apply further theorising or simulation. If a theory is insufficient for generating an inferential response then people can switch to first-person or third-person simulation, with the possibility of returning to theorising. A simulation might also override an initial theory-based response. Furthermore, we follow Mitchell et al. (2009) in suggesting that there will be occasions when an appropriate theory is unavailable, such that people will have to engage in simulation to make some kind of inference. Our model therefore operates in a highly content-dependent manner.

Undoubtedly, empirical evidence needs to be provided for the model presented in this paper. In fact, this model grew out of a series of experiments that we conducted, including one reported by Wilkinson et al. (2010), which showed that people simulate more when reasoning about scenarios involving controllable rather than uncontrollable events, with the reverse pattern for theorising. We propose that our model can readily accommodate such evidence for content-dependency in mental-state reasoning. We suggest that more simulation is evoked for controllable events because participants are more readily able to engage in hypothetical thinking and planning in relation to such scenarios, whereas in the case of uncontrollable events participants are likely to engage in the extraction of a theory since there is little more that they are able to do. Wilkinson et al. (2010) also found that participants often switched between theorising and simulation within the same response. This finding aligns well with Figure 1, which can accommodate this process and the inter-dependence of theorising and simulation in that participants may start out theorising and then adopt simulation to develop their answer further. Furthermore, Wilkinson et al. (2010) noted that participants’ theory-based responses tended to be much quicker than their simulation-based responses. This is explicable given that the extraction of theories is assumed to arise in a high-capacity but low-effort manner, whereas simulation is assumed to be more involved, requiring a longer and more controlled reasoning process (e.g., Goldman, 2006). We acknowledge, however, that such claims would benefit from corroborative evidence via the deployment of chronometric measures.

A further important issue in relation to theorising and simulation processes within a hybrid model is whether these processes operate independently or whether there are dependencies, with the output of one process determining the likelihood of deploying the other process (see Elqayam,
2009, for a discussion of this issue in the context of dual-process theories of reasoning). It is too soon to speculate on this matter, although what has been established in our own research using think-aloud techniques is that people are readily inclined to switch between theorising and simulating within the same reasoning task (Wilkinson et al., 2010). This evidence is at least suggestive of a degree of interdependence between the two processes.

We finally turn to the question of what happens in our proposed model when conflict arises between theorising and simulation. For example, imagine a scenario in which a student fails an important assignment and where our inferential goal is to understand what they might be feeling. Using theorising we might infer that the person will be upset, since if someone fails to pass an assessment they are likely to be distraught. However, if we are also presented with the information that the individual spent every evening drinking in the pub during the week prior to the assignment deadline, we may run a simulation of the person’s mental state to draw the conclusion that the assignment was not of much importance to them. We propose that when such conflict arises a “type 3” conflict-resolution process comes into play (see Evans, 2009). This process would arise subsequent to processes of theorising and/or simulation, but before the generation of a final response. In this way it would be possible for simulation to override a theory-based decision when conflict occurs between the two processes, much as reflective reasoning can override a belief-oriented response in syllogistic inference tasks (e.g., see Ball, 2011).

Conclusions
We have presented arguments for why theory-based reasoning can be viewed as synonymous with intuitive reasoning and simulation-based reasoning can be viewed as synonymous with reflective reasoning within a dual-process framework. This argument was originally advanced by Wilkinson et al. (2010), but we have extended it here so as to provide a more complete and compelling explanation of the parallels between these two hitherto separate conceptual dichotomies. We have additionally considered some of the key questions that need to be addressed by researchers who see value in exploring these suggested parallels further.

Our proposals also resonate with recent calls for greater integration between theorising and simulation accounts of social cognition (Bach, 2011). Bohl and van den Bos argue that the general notion of “theory of mind” is primarily focused on type 2 processing (reflective thinking) rather than type 1 processing (intuitive thinking). We contest this point and instead propose that the traditional distinction between theory-based and simulation-based inferences is best viewed as aligning with the intuitive (type 1) versus reflective (type 2) distinction.

In terms of the development of dual-processes, Evans (2011) has stated that he does not wish to propose that reflective reasoning replaces intuitive reasoning, but rather that the two co-occur in adulthood. This proposal is similar to Mitchell et al.’s (2009) claim that theorising and simulation operate side-by-side in adulthood. Of course, it is difficult at this point to map out a developmental progression of intuitive and reflective reasoning. Mitchell et al. propose that we start out by simulating and Bach (2011) has suggested that theories may grow out of repeated simulations. These ideas initially seem counterintuitive, since simulation is so dependent upon general cognitive resources (Currie, 1996; Goldman, 2006), which tend not to be well developed in young children. However, the claim is not that children necessarily make correct inferences; indeed false belief studies show that they start out by giving incorrect responses by answering from their own perspective (Mitchell et al., 2009). When it comes to the development of intuitive and reflective reasoning the pattern should, of course, align fully with that for theorising and simulation, with reflective reasoning developing first and intuitive reasoning later, which does seem to capture aspects of the development of expertise. For example, when one first learns to drive a car the process is deliberate and controlled, but after time things become automated. We propose that this represents a shift from deliberate reflection to automatic intuitive reasoning.

We trust that by advancing an account of how theorising and simulation align with intuitive and reflective thinking we have provided inspiration for future empirical work and theoretical development and will enliven future discussion concerning the processes involved in mental state reasoning.

Acknowledgements
We thank Rachel Cooper for previous discussions relating to the ideas presented in this paper. We also thank the ESRC for funding the first author’s interdisciplinary PhD in Psychology and Philosophy at Lancaster University, UK.

References


