Arbitrating Between Theory-Theory and Simulation Theory: Evidence from a Think-Aloud Study of Counterfactual Reasoning

Meredith R. Wilkinson (m.r.wilkinson@lancaster.ac.uk)
Department of Psychology, Lancaster University
Lancaster, LA1 4YF UK

Linden J. Ball (l.ball@lancaster.ac.uk)
Department of Psychology, Lancaster University
Lancaster, LA1 4YF UK

Rachel Cooper (r.v.cooper@lancaster.ac.uk)
Department of Philosophy, Lancaster University
Lancaster, LA1 4YG UK

Abstract

How we engage in mental state reasoning remains a contentious issue, reflected in the debate between theory-theorists, who argue that we deploy theory-based rules, and simulation theorists, who argue that such reasoning is subserved by simulation processes. The present study examined whether theory-based or simulation-based reasoning is adopted in regret-oriented counterfactual scenarios involving mental state inferences. Participants thought aloud while reasoning through such scenarios from the perspective of themselves, themselves and another, or two other individuals. The scenarios also manipulated the controllability of event outcomes. Results revealed more theorizing in the uncontrollable than the controllable scenarios, and more simulation in the controllable than uncontrollable ones. More theorizing was also observed in the "other-and-other" than the "self-only" condition. These findings highlight the value of adopting a hybrid model of mental state reasoning, where theorizing and simulation are integrated within a common framework, with such processing being deployed in a context-sensitive manner.

Keywords: Simulation theory; theory-theory; counterfactual thinking; regret; think aloud protocols; mental models.

Introduction

The theory-theory (TT) versus simulation theory (ST) debate focuses on how we understand and reason about our own and others' *mental states* relating to beliefs, desires, intentions and the like. The TT approach (e.g., Carruthers, 1996) argues that we apply both tacit and non-tacit "theories" when understanding mental states. In contrast, ST posits that such understanding is attained either: (i) by "offline" simulation (e.g., Goldman, 2006), in which we take our own beliefs and desires offline, input those of another individual, and run a simulation process; or (ii) by imagining how we would feel in a given situation and by assuming that since other individuals are similar to ourselves then they would feel the same as we do (e.g., Gordon, 1986). Recently, a new variety of theorists have emerged who argue for a *hybrid* approach, in which both theory and simulation are adopted, dependent on the situation (e.g., Mitchell, Currie, & Ziegler, 2009). As we will demonstrate, this hybrid approach has considerable appeal, not least because it can capture the way in which mental state reasoning is sensitive to a multiplicity of factors associated with the prevailing situation, such as its familiarity. The factor of interest in the present paper relates to the protagonist's capacity to have control over the outcome associated with a situation.

Previous Attempts to Arbitrate Between TT and ST

Since the emergence of the TT/ST debate, psychologists and philosophers have been keen to find a test case to arbitrate between these accounts. This pursuit has largely focused on empirical findings that derive from comparisons between autistic individuals (who have various deficits in mental state understanding) and those without autism. The debate surrounding the correct theoretical interpretation of mental state reasoning deficits in autism appears to have reached an impasse (e.g., Carruthers, 1996, has argued that evidence from autism supports TT, while Goldman, 2006, argues that it corroborates ST), such that researchers have started to explore other ways to address the TT/ST debate.

Recent work in this latter vein reported by Kühberger, Kogler, Hug, and Mösl (2006), examined the TT/ST debate using the "position effect" (i.e., a bias to select the rightmost object in an array of identical objects when asked to make a preference judgment). In Experiment 1, participants observed a target person viewing a line of pantyhose and had to imagine viewing the items in the same manner. It was found that participants could predict the target's preference for the rightmost item, which seems to support ST since the availability of sufficient imaginative input (i.e., reasoning from the perspective of the target) enabled participants to predict the position bias. In Experiment 2, participants were given a verbal description about an actor rather than observing them, which failed to produce the results of Experiment 1. In Experiment 3, participants were informed about the position effect (but what it entailed was not explained), and were told to ignore it when making their selection. The fact that most people still demonstrated the

bias was interpreted as indicating deployment of an incorrect theory.

Although, Kühberger et al. claim their results support ST, we wonder whether they speak more to a hybrid account in which TT operates under some conditions and ST under others (e.g., Experiment 3 suggests an incorrect theory can be overridden by simulation). We also note a limitation of the study, which is that it focuses on epistemic mental states (i.e., beliefs), which are rather divorced from the richness of everyday mental state understanding. In order for TT and ST concepts to be useful for examining mental state reasoning (rather than just the currency of an esoteric philosophical debate; cf. Ratcliffe, 2007), such concepts must show general applicability to a wide variety of everyday mental state reasoning contexts.

In this paper we propose that counterfactual reasoning about mental states may provide a new test case for arbitrating between TT and ST. Counterfactual reasoning involves imagining how events associated with regret or disappointment could have turned out differently. Such reasoning is commonplace in everyday life and is vital for understanding how other people may be feeling in response to negative outcomes of real-world situations.

Mental Models and Counterfactual Thinking

The conceptual analysis of counterfactual reasoning is currently dominated by those adopting a mental models framework (e.g., Byrne, 2002, 2005), where mental models reflect representations of actual and counterfactual possibilities. Two particular phenomena in counterfactual reasoning that have been addressed by mental model theorists are: (i) the "action effect", which concerns the observation that greater regret intensity is elicited by acts of commission than acts of omission in the short term (e.g., Kahneman & Tversky, 1982), with the reverse being evident in the long term (e.g., Gilovich & Medvec, 1994); and (ii) the "temporal order effect", whereby we are more likely to reason about undoing the final event in a sequence of events that led to a negative outcome (e.g., Byrne, Segura, Culhane, Tasso, & Berrocal, 2000).

Atkinson, Bell, and Feeney (2009) explored both effects in a study using regret-oriented scenarios. To examine the action effect participants were asked to decide which of two protagonists would feel more regret, an actor or non-actor. To investigate the temporal order effect, participants were asked who would feel worse, the actor who was mentioned first or second. The study also manipulated the time that participants had available to register their response: either they had as long as they wished or they had to answer as quickly as possible. It was found that there was no effect of response time on the emergence of the temporal order effect. However, the action effect was disrupted in the speeded condition, with the actor being selected significantly less often than in the delayed condition. The finding that the action effect and temporal order effect are differentially influenced by the response-time manipulation is claimed by Atkinson et al. to be a consequence of reasoners needing to build complex representations when displaying the action effect (cf. Feeney & Handley, 2006). In particular, for the action effect to arise the reasoner has to compare events associated with both the actor and the non-actor.

We concur with Atkinson et al.'s interpretation, and also believe that their findings are relevant to understanding the role of theorizing and simulation in counterfactual reasoning. The observation that participants readily select the second actor in the temporal order scenario, regardless of time constraints, suggests that participants may be applying a straightforward, theory-based "rule" that the second actor would feel more regret. For the action effect, however, we propose that with sufficient time participants are likely to engage in mental simulation to pursue comparisons between the levels of regret felt by the actor and non-actor - in line with Atkinson et al.'s claim that participants flesh out mental models in these cases.

We illustrate this latter point with reference to a classic action/inaction scenario (Gilovich & Medvec, 1994):

> "Dave and Jim do not know each other but both are enrolled at the same elite East Coast University. Both are only moderately satisfied where they are and both are considering transferring to another prestigious school. Each agonizes over the decision, going back and forth between thinking he is going to stay and thinking he will leave. They ultimately make different decisions: Dave opts to stay where he is and Jim decides to transfer. Suppose their decisions turn out badly for both of them: Dave still doesn't like it where he is and wishes he had transferred, and Jim doesn't like his new environment and wishes he had stayed".

When considering the mental models constructed when reasoning about Dave, Feeney and Handley (2006) argue that participants construct the actual state of affairs in which he stayed and was unhappy and the counterfactual state in which he moved and was happy. However, we propose that a third possibility may be constructed for the short term of Dave moving and being unhappy:

Actual: Stays – Unhappy [Regret]
Counterfactual: Moves – Happy [No regret]
Counterfactual: Moves – Unhappy [Regret]

When fleshing out this model set it is apparent that three possibilities have to be constructed. These would not be easy to derive from theory-based processing alone, which underpins our proposal that simulation may be necessary for participants to imagine options that the protagonist may be considering - along with their related emotional impact. We further propose that when individuals are presented with a counterfactual scenario they will typically evoke a two-stage reasoning process. The first stage involves bringing to mind an initial model based on theory-driven processing (e.g., "If failing to take an action turns out badly then one will feel regret"). Assuming that a response can be generated at Stage 1 without any perceived need for further processing then this will be done on the basis of the initial model. However, if more processing seems to be needed then individuals may

engage in simulation. This Stage 2 process would be more cognitively effortful than theorizing as well as more sequential and controlled in nature.

Predictions of the Study

These aforementioned ideas allow us to develop predictions in relation to the experiment we report below. One focus of the research concerned the distinction between "controllable" and "uncontrollable" outcomes within regret-oriented counterfactual situations. Scenarios involving uncontrollable outcomes limit the consideration of how the outcome could have turned out better. We therefore predict that such scenarios will be susceptible to reasoning based on the application of theory-based inferences. For controllable outcomes, however, although theory-based reasoning is available as a starting point, there is also the potential for the reasoner to flesh out possible ways in which events could have turned out differently. We therefore predict increased simulation-based reasoning when participants engage in mental state understanding in the controllable case relative to the uncontrollable case.

Most philosophers have focused on arbitrating between TT and ST with reference to situations involving reasoning about another person. However, it is also interesting to examine people's reasoning about their own mental states. Evidence that there are differences in how we reason about ourselves compared to others comes from a study by Girotto, Ferrante, Pighin, and Gonzalaz (2007), who presented participants with a scenario where they could win a prize by solving a problem. One key experiment involved two conditions. In the "actor" condition participants were presented with two sealed envelopes; one they were told contained an easy problem, one a difficult one (in fact, both contained an insoluble problem). In the "reader" condition, participants read about a protagonist who had to make the same choice as in the actor condition, with an identical outcome (i.e., failure to solve the problem). Participants were either assigned to the actor or reader condition and were afterwards asked to name one way in which the outcome could have turned out better. Responses were coded as either modifying choices (e.g., selecting the other envelope), or as modifying problem features (e.g., having more time). Girotto et al. found the actors were more likely to alter problem features, while readers were more likely to alter a choice, such as the selection of the envelope.

The finding that participants reason about different aspects of a scenario in the self/actor versus other/reader condition suggests that different processes may be occurring. The increased likelihood of undoing a problem feature in the actor condition indicates the consideration of more possibilities than in the reader condition, which tended to involve just choosing the other envelope. This points to the idea that more possibilities are considered when reasoning from the perspective of oneself than that that of another, with the implication being that more simulation may arise in the former than the latter situation.

We also note that when people engage in counterfactual reasoning about another individual then theory-based reasoning may take precedence as people tend to possess a wealth of generalized rules concerning how people will feel in regret-oriented situations (e.g., "A person will be upset if they miss out on something they desire greatly"). However, when reasoning about ourselves we may be more likely to progress onto the simulation stage since we possess more specialized knowledge concerning ourselves and the nuances of our own reactions to events. In this way, people who are engaged in self-oriented reasoning may move away from the application of generalized folk psychological theories toward the simulation of multiple eventualities.

In sum, by using regret-oriented counterfactual scenarios involving mental state reasoning we assumed that we would gain useful insights to address two issues surrounding the TT/ST debate. First, such scenarios should usefully inform whether theory-based or simulation-based reasoning dominate in mental state understanding, or whether both forms of reasoning are deployed. Second, the manipulation of factors such as outcome controllability and the self/other distinction should clarify whether contextual and instructional aspects of the presented scenarios determine whether individuals are more likely to theorize or simulate.

Method

Participants

Participants were 90 individuals at Lancaster University who received either course credit or payment. None had prior knowledge of research on reasoning or theory of mind.

Design

A 2 x 3 mixed between-within participants design was adopted. The between participants factor was the perspective that participants had to reason from, which had three levels: self; self-and-other; other-and-other. The within participants factor was outcome controllability, with two levels: controllable versus uncontrollable.

Materials and Pre-Test

Participants received two controllable and two uncontrollable scenarios. Controllable scenarios concerned: (i) two individuals performing poorly on a University assignment (assignment scenario); and (ii) individuals changing or not changing a minor subject to a major at University and then not enjoying the course and receiving a poor course grade (course scenario). The uncontrollable scenarios concerned: (i) losing a game of table football by scoring an own goal, followed by one's opponent scoring a winning goal (football scenario); and (ii) an individual missing their flight by 5 mins, with the plane having been delayed, and another individual missing their flight by 30 mins, with the plane leaving on time (plane scenario).

In the self-only and self-and-other conditions the participant had to take on the role of one of the individuals within the scenario. In those conditions that involved other individuals, we presented "personas" (i.e., brief bio-sketches

of the named individuals) in an effort to increase the realism of the scenarios. For the self-and-other and other-and-other conditions participants were required to state who they thought would feel more regret, upset or frustration (dependent on scenario). For the self-only condition participants had to state how much regret/upset/frustration they would feel with essentially equivalent scenarios.

To validate our controllability manipulation we gave 13 participants the scenarios from the self-only condition. After reading each scenario they had to use a 10-point scale to rate it for familiarity in their everyday life, and for the controllability of the outcome. We also included a mutability question in which participants were asked simply to list all the ways in which the situation could have turned out for the better. Using paired samples t-tests we found that controllable scenarios were rated as significantly more controllable (M = 6.54) than the uncontrollable scenarios (mean = 5.31), $t(12) = 2.66$, $p = .02$. There was no difference, however, in ratings of familiarity (means of 5.58 versus 4.88 for controllable vs. uncontrollable), $t(12) = 1.17$, $p = .26$. For the mutability measure there was a mean of 2.38 mutations for controllable scenarios and 2.69 mutations for uncontrollable scenarios, which was unreliable, $t(12) = 1.67$, $p = .12$. Overall, these pre-test data reveal a solid effect of controllability in the predicted direction, but no confounding effects of mutability or familiarity.

Procedure

Participants were randomly assigned to one of the three perspective conditions and were given associated instructions. They were then presented with a booklet containing the scenarios and were asked to think-aloud whilst reasoning about each one. Scenario order was independently randomized for each participant.

Results

Data Coding

To code the data we adopted Ball and Christensen's (2009) scheme in which each line was coded as reflecting theory-based or simulation-based reasoning. An "ambiguous" code was used when: (i) lines were evenly split across categories (there were 13 instances of these); or (ii) it was difficult to be certain whether theory-based or simulation-based reasoning was being adopted. Theory-based reasoning concerned instances in which the participant adopted tacit or non-tacit theories to make inferences about their own or others' mental states. Such reasoning tended to involve the participant stating general rules regarding mental states, typically involving a grammatical construction such as "The person will feel x because of y". The following excerpt illustrates theorizing taking place, with a participant adopting a tacit rule that captures the notion that action will elicit greater regret in the short term than inaction:

> "I think Mike's gonna feel the more regret in the short term coz he's actually chan- he actually made a bad decision whereas Timmy's decided - Timmy's chosen not to make the decision, so he doesn't know whether or not he'd prefer the other - you can assume he can".

Simulation occurred when participants took their own beliefs and desires offline and inputted those of other individuals (e.g., Goldman, 2006). This simulation process typically involved the participant running through their own or another individual's mental states in relation to the possibilities arising within the scenario so as to determine how they or others would feel (cf. Gordon, 1986). The following excerpt demonstrates such simulation, with the participant imagining themselves in a given situation and stating how they would feel and also how the protagonist might feel, rather than simply stating a rule such as "People feel upset when they receive a poor grade":

> "I myself am not particularly competitive erm, so I might be kind of disappointed and think, 'Oh well, that's kind of surprising, I'll erm, I'll have to find out why I went wrong'. But perhaps Jim might be slightly more likely to think, 'Oh I should have worked harder I should have'".

For each scenario the application of this coding scheme by the first author resulted in a percentage of theorizing and simulation for each participant as a function of all coded lines, including ambiguous ones. An independent coder checked a 10% sample of transcripts after first being trained in the application of the coding scheme. Inter-rate reliability was good, with 74% agreement. All areas of disagreement were resolved through discussion between the coders.

Theory-Based Reasoning

Table 1 presents the percentage of theory-based reasoning as a function of controllability and perspective. A 2 x 3 mixed design ANOVA revealed a main effect of controllability, $F(1, 87) = 15.81$, $MSE = 589.91$, $p < .001$, $\eta_p^2 = 0.15$, with theory-based reasoning being more prevalent in uncontrollable than controllable scenarios. There was also a main effect of perspective, $F(2, 87) = 11.41$, $MSE = 1238.82$, $p < .001$, $\eta_p^2 = 0.21$, with the other-and-other condition evoking the greatest level of theorizing and the self-only condition the least. The controllability by perspective interaction was not reliable, $F(2, 87) = 0.75$, $MSE = 589.91$, $p = .48$, $\eta_p^2 = 0.02$. Post hoc comparisons showed significant differences between the self-only condition and the other-and-other and self-and-other conditions ($ps < .01$). No difference was found between the self-and-other and other-and-other conditions ($p = .23$).

Table 1: Mean percentage of theorizing as a function of outcome controllability and perspective (SDs in brackets).

Perspective	Outcome Controllability		
	Controllable	Uncontrollable	*M*
Self-only	33 (33)	46 (40)	40
Self-&-other	49 (29)	69 (23)	59
Other-&-other	65 (25)	75 (28)	70
M	49	63	

Table 2: Mean percentage of simulation as a function of outcome controllability and perspective (SDs in brackets).

Perspective	Outcome Controllability		
	Controllable	Uncontrollable	*M*
Self-only	52 (38)	40 (37)	46
Self-&-other	24 (27)	20 (21)	22
Other-&-other	20 (21)	15 (24)	18
M	32	25	

Simulation-Based Reasoning

Table 2 presents the percentage of simulation as a function of controllability and perspective. A 2 x 3 mixed design ANOVA revealed a main effect of controllability, $F(1, 87) = 3.97$, $MSE = 555.95$, $p = .049$, $\eta_p^2 = 0.04$, with greater simulation in controllable than uncontrollable scenarios. There was also a main effect of perspective, $F(2, 87) = 12.63$, $MSE = 1122.86$ $p < .001$, $\eta_p^2 = 0.23$, with more simulation in the self-only condition relative to the other-and-other and self-and-other conditions. No interaction was observed between perspective and controllability, $F(2, 87) = 0.57$, $MSE = 555.95$, $p = .57$, $\eta_p^2 = 0.01$. Post-hoc comparisons revealed a significant difference between the self-only condition and the self-and-other other-and-other conditions ($ps < .001$), but no difference between the self-and-other and other-and-other conditions ($p = .90$).

Discussion

To our knowledge this is the first study to examine the TT/ST debate through the prism of mental state reasoning with counterfactual scenarios. The results from our protocol analysis indicate that although theorizing dominated overall, there were nevertheless differences across conditions when theorizing and simulation data were analyzed separately.

Looking first at the controllability factor, our results indicated that theorizing was more prevalent in uncontrollable than controllable scenarios, with the reverse being the case for simulation. The observation that people theorize more and simulate less in uncontrollable scenarios relative to controllable ones is consistent with the view that uncontrollable scenarios evoke less consideration and modeling of alternative possibilities. In essence, participants appear to be minimizing cognitive effort in these cases. For the controllable condition, simulation may have been facilitated because it was possible to consider more alternatives to reality, thereby provoking a more detailed examination of how an individual might feel in a situation. Participants also appeared to be more likely to engage in reasoning about *how* they might feel in such scenarios, using this to infer how the protagonist might feel.

How do these findings concerning the effect of outcome controllability on mental state reasoning fit in with other theories? Mitchell et al.'s (2009) hybrid account argues that simulation is used by default, but in cases where a situation is familiar they suggest that people might use rule-based theorizing as a shortcut strategy. However, our controllable and uncontrollable scenarios were equated for familiarity in a pre-test. As such, since controllability was not confounded with familiarity it is not immediately apparent how Mitchell et al.'s account might address the observed influence of controllability on rates of theorizing and simulation. It may be the case, however, that both theorizing and simulation reflect different strategies for engaging in mental state understanding, with one or other strategy being elicited by different factors in the prevailing context, including familiarity and event controllability - and potentially other cues (e.g., the emotionality of the situation).

Our study also set out to examine whether differences arise in how people reason about themselves versus others. Our analysis showed that the self-only condition elicited more simulation and less theorizing than the other-and-other condition, with the self-and-other condition occupying a middle position on both the theorizing and simulation measures. One reason for relatively more simulation arising in the self-only condition may be that it is triggered by direct emotional engagement with presented scenarios arising from specific memories of personal experiences.

Overall, our findings indicate that both theorizing and simulation occur in mental state reasoning about regret-oriented counterfactual scenarios. This supports a hybrid view of mental state understanding along the general lines espoused by Mitchell et al. (2009), and suggests the traditional TT/ST debate may be misconceived in its attempt to emphasize the deployment of a unitary reasoning approach based purely around *either* theorizing *or* simulation. Our results also have implications for mental models accounts of counterfactual reasoning. So far these accounts have been dominated by studies of the action effect (e.g., Feeney & Handley, 2006), and the temporal order effect (e.g., Atkinson et al., 2009), with less work examining issues relating to the controllability of regret outcomes. Our research suggests that an initial model may be formed by theory-based reasoning, with subsequent models being fleshed out through a mental simulation process involving the identification of multiple alternative possibilities. Although speculative, these ideas resonate with previous findings relating to the action effect, and represent a useful area for future research.

Reflecting on our results more generally, we wonder whether they also speak to dual-process accounts (e.g., Evans, 2003, 2006), which contend that human reasoning involves the interplay between two distinct reasoning processes. On the one hand Type 1 or heuristic processes are fast, automatic, high capacity and involve low cognitive effort. On the other hand, Type 2 or analytic processes are slow, controlled, low capacity and involve high cognitive effort. Under some dual-process accounts, Type 1 processes act by default to provide an initial response that can be overturned through the application of Type 2 processes (e.g., Evans, 2006). We suggest that theory-based reasoning may map onto Type 1 processing, and simulation-based reasoning may map onto Type 2 processing. Our findings suggest that there was little simulation that was not also driven by an initial phase of theorizing, which implies that

theorizing may be primary, and that if further processing is required this arises through simulation and may serve either to confirm or override a theory-based decision.

Evidence for this dual-process view of mental state reasoning comes from Atkinson et al.'s (2009) study, where the absence of an influence of speeded responding on the temporal order effect suggests the rapid and automatic deployment of a rule-based process, in line with TT assumptions that we possess a set of folk psychological theories. Furthermore, Atkinson et al.'s observation that speeded responding modulated the emergence of the action effect is indicative of slower, controlled, Type 2 processing linked to simulation. These dual-process arguments also resonate with Apperly and Butterfill's (2009) claims for two processing systems in mental state reasoning, with the proposal being that infants possess a cognitively efficient but inflexible method for tracking belief states that runs parallel to a later-developing adult system which is more flexible but cognitively demanding.

We conclude by returning to the two issues mentioned in our introduction that we hoped our research might address, that is: (i) whether mental state understanding is based on *either* theory-based reasoning *or* simulation-based reasoning - or whether both types of processing are deployed; and (ii) whether the manipulation of factors such as outcome controllability and the self/other distinction might determine the propensity for individuals to theorize or simulate. In relation to the first issue, we have demonstrated by means of think-aloud protocols and the adoption of counterfactual thinking scenarios that both theory-driven and simulation-driven reasoning play out in mental state understanding, with all participants deploying theorizing *and* simulation to greater or lesser degrees for many of the scenarios. In relation to the second issue, we have shown that people are more likely to engage in simulation when thinking about themselves rather than when thinking about other individuals. Furthermore, they are more likely to engage in simulation when reasoning about controllable than uncontrollable regret outcomes. Moreover, these two factors (i.e., perspective and controllability) appear to combine additively to determine the relative levels of simulation and theorizing that arise in mental state reasoning.

Standard, unitary TT and ST accounts do not seem to be able to accommodate our observations that the processes underpinning mental state understanding are influenced by content, context and perspective effects. Although these accounts may be able to develop ways to explain the present evidence, it remains for the proponents of these theories to take up this challenge. In contrast, hybrid accounts that embrace both TT and ST seem better able to deal with our findings. We suggest that hybrid theories represent an important new direction in research examining the processes associated with mental state reasoning.

Acknowledgements

We acknowledge the ESRC for an interdisciplinary studentship awarded to the first author.

References

Apperly, I.A., & Butterfill, S.A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116,* 953-970.

Atkinson, L., Bell, D., & Feeney, A. (2009). The relationship between counterfactual thinking and emotional reactions to event outcomes: Does one account fit all? *Psychonomic Bulletin & Review, 16,* 724-728.

Ball, L.J., & Christensen, B.T. (2009). Analogical reasoning and mental simulation in design: Two strategies linked to uncertainty resolution. *Design Studies, 20,* 169-186.

Byrne, R.M.J. (2002). Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences, 6,* 426-431.

Byrne, R.M.J. (2005). *The Rational Imagination: How People Create Alternatives to Reality.* Cambridge, MA: MIT Press.

Byrne, R.M.J., Segura, S., Culhane, R., Tasso, A., & Berrocal, P. (2000). The temporality effect in counterfactual thinking about what might have been. *Memory & Cognition, 28,* 264-281.

Carruthers, P. (1996). Autism as mind-blindness: An elaboration and partial defence. In P. Carruthers & P.K. Smith (Eds.), *Theories of Theories of Mind* (pp. 257-273), Cambridge: CUP

Evans, J.St.B.T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454-459.

Evans, J.St.B.T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13,* 378-395.

Feeney, A., & Handley, S.J. (2006). Comparisons, mental models, and the action effect in judgments of regret. *Memory & Cognition, 34,* 1422-1430.

Gilovich, T., & Medvec, V.H. (1994). The temporal pattern to the experience of regret. *Journal of Personality and Social Psychology, 67,* 357-365.

Girotto, V., Ferrante, D., Pighin, S., & Gonzalez, M. (2007). Postdecisional counterfactual thinking by actors and readers. *Psychological Science, 18,* 510-515.

Goldman, A.I. (2006). *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading.* New York: Oxford University Press.

Gordon, R.M. (1986). Folk psychology as simulation. *Mind & Language, 1,* 158-171.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American, 246,* 160-173.

Kühberger, A., Kogler, C., Hug, A., & Mösl. E. (2006). The role of the position effect in theory and simulation. *Mind & Language, 21,* 610-625.

Mitchell, P., Currie, G., & Ziegler, F. (2009). Two routes to perspective: Simulation and rule-use as approaches to mentalizing. *British Journal of Developmental Psychology, 27*, 513-543.

Ratcliffe, M. (2007).*Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation.* Basingstoke: Palgrave Macmillan.