

SIMULATION THEORY

Mental simulation is the simulation, replication or re-enactment, usually in imagination, of the thinking, decision-making, emotional responses, or other aspects of the mental life of another person. According to simulation theory, mental simulation in imagination plays a key role in our everyday psychological understanding of other people. The same mental resources that are used in our own thinking, decision-making or emotional responses are redeployed in imagination to provide an understanding of the thoughts, decisions or emotions of another.

Simulation theory stands opposed to the ‘theory theory’ of folk psychology. According to the theory theory, everyday psychological understanding depends on deployment of an empirical theory or body of information about psychological matters, such as how people normally think, make decisions or respond emotionally. Simulation theory does not altogether deny that third-personal psychological knowledge is implicated in our folk psychological practice of prediction, interpretation and explanation. But it maintains that, over a range of cases, the first-personal methodology of mental simulation allows us to avoid the need for detailed antecedent knowledge about how psychological processes typically operate.

- 1 Simulation in reality and in imagination**
- 2 Mental simulation and prediction**
- 3 Mental simulation, interpretation and explanation**
- 4 Varieties of simulation theory**

1 Simulation in reality and in imagination

The behaviour of an object under particular actual or hypothetical circumstances may be simulated by the behaviour of another object of the same kind under the same or similar circumstances. Thus, for example, we can simulate the increase of pressure in a cylinder of gas that would result if its temperature were raised. The simulation exercise involves heating another cylinder of the same gas, starting from the same initial temperature and pressure. We can make a prediction about what would happen to the first cylinder by observing what actually happens to the second. Similarly, we can simulate the effect of a drug on one person by having another person ingest the same substance and, once again, the simulation may generate a prediction. In fact, I can simulate the effect of a drug on another person by taking the drug myself. In these cases of the cylinder and the drug, the simulation takes place in reality: the cylinder is really heated, and I really take the drug. The same processes occur in the simulation as would be operative in generating the behaviour of the object being simulated. This is vital if simulation is to provide an alternative to the deployment of theoretical knowledge about gases or drugs. If we merely imagine heating the cylinder or imagine taking the drug then, in order to move the simulation forward, we shall need to draw on a body of information about temperature and pressure in gas cylinders or about the effects of the drug on human beings. So any prediction that we arrive at will depend on this antecedent knowledge and the role of the simulation will be insignificant.

In these examples, simulation in reality is ‘process-driven’ while simulation in imagination is ‘theory-driven’ (Goldman 1989) and it is only simulation in reality that constitutes a genuine alternative to the use of empirical theory in making a prediction. When it comes to simulating the mental life of another person, I may sometimes be able to place myself, in reality, in the other person’s mental and physical situation. But mental

simulation in reality is often not a practical option. I may, for example, want to predict the decision of someone whose experiences, attitudes and emotions are very different from my own and I cannot simply change my mental life to bring it into line with theirs. This may suggest that mental simulation will, for the most part, be theory-driven simulation (Dennett 1981). But a distinctive claim of simulation theory can be expressed by saying that mental simulation in imagination can be process-driven rather than theory-driven. This is because at least some mental processes operate in just the same way when we imagine being in a particular situation as they would if we were really in that situation.

This way of putting the claim would not, however, be congenial to all simulation theorists. It may suggest a picture of mental simulation in which the simulator observes the operation of a piece of their own mental machinery and then draws an inference about the operation of the mental machinery in the other person. According to this picture, the gaze of the simulator would be first inward and then outward to the person being simulated and it would be the inward gaze that would distinguish simulation from the use of a third-personal empirical theory about psychological processes. But at least some simulation theorists would stress a different point about direction of gaze. Since the simulator replicates or re-enacts in imagination the thinking of the person being simulated, the simulator's gaze is neither inward nor upon the person being simulated but, primarily, upon the (imagined) circumstances about which the person being simulated is thinking (Heal 1986; Gordon 1995).

2 Mental simulation and prediction

The claim that mental simulation in imagination can be process-driven is particularly plausible in the case of the mental activities of theoretical and practical reasoning. Thinking and decision-making proceed in the same way from hypothetical or imagined premises about my situation as from premises that I really believe and this is vital to my ability to consider what I would think or decide in various actual or hypothetical circumstances. Thus, for example, by engaging in hypothetical practical reasoning I can consider what I would decide to do if I were to learn that I had inherited one million pounds or that I was wanted by the police. But then I can, in the same way, reach a prediction about what someone else relevantly similar to me will decide to do if they find themselves in such circumstances. I can make a prediction about the other person's decisions or thoughts by placing myself in imagination in their situation and then making use of my own ability to engage in practical or theoretical reasoning.

The use of mental simulation in imagination to make a prediction about another person seems to rely on two assumptions: first, that the relevant mental processes operate in the same way in imagination as in reality, and second, that the other person is relevantly similar to me. To see why the second assumption seems to be required, suppose that the person being simulated differs from me in their desires for consumer goods or their beliefs about the police. In that case, what I would think or decide if I were to learn that I had inherited one million pounds or that I was wanted by the police will be an imperfect guide to what that person will think or decide. But, in fact, the predictive use of mental simulation in imagination does not require the full strength of the second assumption. Given the first assumption, I can compensate for differences by taking on in imagination, not only the other person's belief that they have inherited one million pounds or are wanted by the police, but also a host of their other attitudes. Rather than projecting myself, with my attitudes, into the other person's situation, I need to identify in imagination with the other person, to imagine being them.

Mental simulation in imagination can allow for differences in attitudes. But a version of the second assumption is still needed. The other person's theoretical or practical reasoning may be different from mine, not because of a difference in attitudes, but because of a difference in mental processes. It might be, for example, that the other person's reasoning processes are operating abnormally because they have just taken a drug. I cannot compensate for this difference between the other person and myself just by imagining taking the same drug. In order to adjust my simulation in imagination, I need to make use of empirical information about how the drug influences reasoning.

So the use of mental simulation in imagination relies on an assumption of similarity in the basic operation of mental processes and where that assumption is false, we need to invoke elements of empirical theory. But even where correct prediction requires an intrusion of psychological theory it need not be the case that the prediction owes everything to theory and nothing to simulation. The empirical information that is drawn on might simply specify the ways in which the reasoning of someone who has taken the drug will differ from what is normal, without also providing an account of normal reasoning. In that case, mental simulation could be used to generate an initial prediction that would then be modified in the light of the empirical information about the effects of the drug.

An important objection to simulation theory begins from examples of prediction failure (Stich and Nichols 1992, 1995; Nichols et al. 1996). If folk psychological predictions draw on a body of empirical theory about psychological processes, as the theory theory of folk psychology says, then use of an incorrect theory may well lead to incorrect predictions. So prediction failures have a ready explanation in terms of the theory theory. According to simulation theory, however, mental simulation should yield correct predictions about another person's decision-making provided that two conditions are met. The simulator must correctly take on the other person's attitudes and there must be no difference in the basic operation of mental processes between the simulator and the person whose practical reasoning is being simulated.

Opponents of simulation theory use examples of prediction failure in which both these conditions appear to be met in order to support their argument that, even if mental simulation is a possible method for making folk psychological predictions, it is not the method that people actually use. In these examples, it is indeed plausible that there are no gross differences in mental processing between the group of simulators and the group of people being simulated, for members of both groups are selected randomly from the same population of subjects. But the simulation theory of folk psychological prediction may still be defended. The crucial claims for a defence are that the people who are being simulated are influenced by non-rational factors (subtle analogues of drugs or tiredness) and that, even if the simulators correctly take on these people's attitudes, they do not take these non-rational factors into account (Harris 1992; Heal 1996a).

3 Mental simulation, interpretation and explanation

According to simulation theory, our ability to engage in mental simulation in imagination is an important component in the basis of our folk psychological practice of prediction, interpretation and explanation. We have just seen that prediction of another person's thoughts and decisions that are arrived at by theoretical or practical reasoning is a favourable case for simulation theory while mental simulation in imagination is not a good way to predict the mental effects of being drugged, or tired, or drunk. Mental simulation can be used to predict thoughts, decisions, and ultimately behaviour, on the basis of information about a subject's attitudes. It can also play a part in the attribution of

attitudes given information about a subject's behaviour or about the thoughts and decisions to which those attitudes lead. For the attributed attitudes should plausibly give rise to the thoughts, decisions and behaviour that the subject actually produces and this can be tested by mental simulation.

On the theory theory approach to our folk psychological practice, all three strands – prediction, interpretation and explanation – draw on knowledge of the generalisations of a psychological theory. Folk psychological explanation is a kind of covering law (deductive nomological or subsumptive) explanation, prediction and explanation work in essentially the same way, and interpretation is inference to the best explanation. We have just seen that simulation theory provides linked accounts of the prediction and interpretation strands. But, for at least two reasons, it is not so clear how simulation theory is to account for folk psychological explanation. First, subsumptive explanations require generalisations; but mental simulation is supposed not to depend on antecedent knowledge of psychological generalisations. Second, it is a striking feature of the earlier examples of the gas cylinder and the drug that the use of simulation to generate predictions does not provide any explanation of why the simulated objects behave in the way that they do.

One possible response to the first apparent problem for simulation theory is to note that, although mental simulation does not depend on antecedent theoretical knowledge about psychology, mental simulation may yield, not just predictions about individuals, but also knowledge of generalisations. These generalisations could then figure in subsumptive explanations. However, simulation theorists more often maintain that folk psychological explanation is different in kind from deductive-nomological explanation so that it does not require knowledge of empirical generalisations about how psychological processes typically operate (Heal 1986). It is also said that folk psychological explanation yields a distinctive kind of understanding of the other person 'from the inside' (Heal 1998a) and this highlights a difference between mental simulation, as these theorists conceive it, and simulation as it figured in the examples of the cylinder and the drug. In those examples, the person conducting the simulation had no access to the processes taking place; it was 'black box simulation.' But in mental simulation, my mental activity conducted within the scope of imaginative identification with another person is not a process that is hidden from me. If this account of folk psychological explanation is correct then even theory-driven simulation, although it does not offer a distinctive methodology for prediction, may contribute to explanations of a kind different from that which the theory theory itself recognises.

There is an important similarity between what simulation theory says about folk psychological understanding and R.G. Collingwood's (1946) claim that historical understanding is to be achieved by the re-enactment of the historical character's thought. Indeed, simulation theory has marked affinities with the hermeneutic tradition of Vico, Herder, Dilthey, Weber and Croce, as well as Collingwood.

4 Varieties of simulation theory

Simulation theorists (and opponents) differ over the level of description at which the theory is primarily pitched and over its epistemological status. Some regard the theory as being about the architecture of an information-processing system while others focus on conscious, thinking persons. Some take the theory as being straightforwardly empirical while others regard it as having a more a priori character.

On one construal of what is at issue, the simulation theory and the theory theory are competing subpersonal-level accounts of the information-processing machinery that

suberves decision-making (Goldman 1989, 1992; Stich and Nichols 1992). This machinery takes representations corresponding to the subject's beliefs and desires as inputs and produces a representation of a decision as output. The production of this output representation typically leads to a piece of behaviour on the part of the subject. The simulation theory says that this information-processing component can be taken 'off line' from its usual inputs and outputs and can be fed input representations that correspond to 'pretend beliefs' and 'pretend desires' rather than real beliefs and desires. Operating in this off-line mode, the system then produces output representations that correspond to 'pretend decisions' and that do not lead to real behaviour. From the point of view of this construal, considerable interest attaches to discoveries that components of the visual system are also active during visual imagery and that components of the system that is implicated in intentional action are also involved in motor imagery (Currie 1995a; Currie and Ravenscroft 1997). For these discoveries make it more empirically plausible that imagined decision-making should make use of some of the same information-processing components as real decision-making.

Although many contributions to the simulation versus theory debate are cast in terms of information-processing mechanisms, some influential advocates of the mental simulation approach to folk psychology cast their theory in personal-level terms. They regard speculations about subpersonal-level machinery as interesting and even plausible, but strictly speaking inessential. These personal-level simulation theorists differ amongst themselves over the epistemological status of the theory. On the one hand, Heal (1998b, p. 478) argues that 'it is an a priori truth that simulation, in some sense, must be given a substantial role in our personal-level account of psychological understanding'. Heal (1996a) also restricts the domain of mental simulation to rational (or otherwise intelligible) transitions between mental states that have propositional contents. One result of this restriction is that there is a close connection (though not identity) between Heal's version of simulation theory and what might be called the 'normative theory theory' of folk psychology. According to this latter theory, our folk psychological practice draws on knowledge of an a priori theory about right reasoning rather than an empirical theory about how mental processes typically operate. Heal (1986) makes the point that, to the extent that mental simulation requires knowledge of principles, these are normative or semantic rather than causal.

Gordon (1986, 1992, 1996), on the other hand, does not regard simulation theory as having close connections with issues about normativity and assigns mental simulation a correspondingly broader domain including, importantly, the simulation of emotion. It is a striking fact that our emotional responses to imagining being in a terrifyingly dangerous situation are similar to the responses that we would have if we were really in such a situation (Walton 1997). So, on Gordon's account, prediction of emotional responses is a favourable case for simulation theory, like prediction of decisions that are arrived at by practical reasoning and unlike prediction of the mental effects of being drugged or tired. Because the facts about emotional responses to imagined danger are not plausibly a priori, Gordon does not share Heal's assessment of the epistemological status of simulation theory.

Both simulation theory and the theory theory are primarily about normal adult folk psychological practice. But simulation theorists and theory theorists also offer competing accounts of the developmental transitions that lead up to the mature adult state and of the failure of folk psychological understanding that is characteristic of autism (see Harris 1992; Gopnik and Wellman 1992; Leslie and German 1995; Currie 1996; MIND, CHILD'S THEORY OF).

See also: FOLK PSYCHOLOGY.

References and further reading

- Carruthers, P. and Smith, P.K. (eds) (1996) *Theories of Theories of Mind*, Cambridge: Cambridge University Press. (An excellent collection of philosophical and empirical papers on the theory versus simulation debate, ranging over mature adult folk psychological understanding and its development, competing explanations of autism, and 'theory of mind' in non-human primates.)
- *Collingwood, R.G. (1946) *The Idea of History*, Oxford: Oxford University Press; Revised Edition, 1993. (The papers in Part V ('Epilogomena') explore the idea that historical understanding requires re-enacting the thoughts of the protagonists.)
- *Currie, G. (1995a) 'Visual Imagery and the Simulation of Vision'. *Mind and Language* 10: 25–44. (Suggests that mental imagery results from running the visual system in an off-line mode.)
- Currie, G. (1995b) 'Imagination and Simulation: Aesthetics Meets Cognitive Science', in M. Davies and T. Stone (eds) *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell Publishers, 151–69. (Proposes that a mental simulation mechanism stands to imagination as H₂O stands to water.)
- Currie, G. (1995c) *Image and Mind: Film, Philosophy and Cognitive Science*, Cambridge: Cambridge University Press. (An innovative book that *inter alia* deploys simulation theory as part of a theory of how we engage with fiction.)
- *Currie, G. (1996) 'Simulation-Theory, Theory-Theory, and the Evidence from Autism', in P. Carruthers and P. K. Smith (eds) *Theories of Theories of Mind*, Cambridge: Cambridge University Press, 242–56. (Compares theory theory and simulation theory explanations of autism and concludes that simulation theory explanations are slightly better placed than theory theory explanations.)
- *Currie, G. and Ravenscroft, I. (1997) 'Mental Simulation and Motor Imagery', *Philosophy of Science* 64: 161–80. (Reviews the evidence that components of the motor system such as pre-motor cortex are active but off line in motor imagery.)
- Davies, M. and Stone, T. (eds) (1995a) *Folk Psychology: The Theory of Mind Debate*, Oxford: Blackwell. (Contains many of the seminal papers in the simulation versus theory debate along with a substantial editorial introduction.)
- Davies, M. and Stone, T. (eds) (1995b) *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell Publishers. (Continues the debate from Davies and Stone (eds) (1995a) and includes papers that apply the idea of mental simulation in ethics, psychopathology and artificial intelligence.)
- Davies, M. and Stone, T. (1998) 'Folk Psychology and Mental Simulation', in A. O'Hear (ed.) *Contemporary Issues in the Philosophy of Mind*, Cambridge: Cambridge University Press, 53–82. (Expansion of the material in this entry; includes discussion of normativity, prediction failure, the circumscribed domain of simulation, and Collingwood on re-enactment.)
- *Dennett, D.C. (1981) 'Making Sense of Ourselves', *Philosophical Topics* 12: 63–81; repr. in *The Intentional Stance*, Cambridge, MA: MIT Press, 1987, 83–101. (A discussion of the role in folk psychological interpretation of an assumption of rationality. It contains an influential argument that simulation must, in the end, be a kind of theorising.)

- *Goldman, A.I. (1989) 'Interpretation Psychologized', *Mind and Language* 4: 161–85; repr. in Davies and Stone (eds) (1995a). (Defends simulation theory as a more plausible account of the attribution of mental states to others than the rationality approach or the folk-theory approaches; also introduces the distinction between process-driven and theory-driven simulation.)
- *Goldman, A.I. (1992) 'In Defense of the Simulation Theory'. *Mind and Language* 7: 104–19; repr. in Davies and Stone (eds) (1995a). (A response to the arguments against simulation theory advanced by Stich and Nichols (1992); stresses the relative parsimony of simulation theory and discusses the connection between simulation and empathy.)
- *Gopnik, A. and Wellman, H.M. (1992) 'Why the Child's Theory of Mind Really *Is* a Theory', *Mind and Language* 7: 145–71; repr. in Davies and Stone (eds) (1995a). (Sets out in detail the analogy between the child's development of a theory of mind and the development of natural scientific theories.)
- *Gordon, R.M. (1986) 'Folk Psychology as Simulation', *Mind and Language* 1: 158–71; repr. in Davies and Stone (eds) (1995a). (Along with Heal 1986 and Goldman 1989, this paper inaugurated the contemporary debate. It introduces the idea of simulated practical reasoning as a method for predicting one's own behaviour and that of others.)
- *Gordon, R.M. (1992) 'The Simulation Theory: Objections and Misconceptions', *Mind and Language* 7: 11–34; repr. in Davies and Stone (eds) (1995a). (Argues that the simulator is not using one individual, himself, as a model of another and illustrates the point that the simulator's direction of gaze is outwards to the external world.)
- *Gordon, R.M. (1995) 'Simulation Without Introspection of Inference from Me to You', in M. Davies and T. Stone (eds) *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell Publishers, 53–67. (Rejects the idea that mental simulation involves an analogical inference from oneself to another person and develops an account of self-attribution of beliefs in terms of an 'ascent routine'.)
- *Gordon, R.M. (1996) "'Radical" Simulationism', in P. Carruthers and P. K. Smith (eds) *Theories of Theories of Mind*, Cambridge: Cambridge University Press, 11–21. (Further develops the version of simulation theory that involves a transformation in imagination of oneself into the other.)
- *Harris, P.L. (1992) 'From Simulation to Folk Psychology: The Case for Development', *Mind and Language* 7: 120–44; repr. in Davies and Stone (eds) (1995a). (An important empirical paper providing a detailed simulation-theoretic account of the major stages in the child's development of mature folk psychological understanding.)
- *Heal, J. (1986) 'Replication and Functionalism', in J. Butterfield (ed.) *Language, Mind and Logic*, Cambridge: Cambridge University Press, 135–50; repr. in Davies and Stone (eds) (1995a). (An important paper that introduces the idea of mental simulation (replication) and contrasts it with functionalist approaches. It explicitly addresses the objection that simulation requires the deployment of a theory.)
- Heal, J. (1995) 'How to Think About Thinking', in M. Davies and T. Stone (eds) *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell Publishers, 33–52. (Argues that simulation must have an important role in the prediction of a person's beliefs; also discusses the role that theory may play in prediction and in other aspects of folk psychological practice.)

- *Heal, J. (1996a) 'Simulation and Cognitive Penetrability', *Mind and Language* 11, 44–67. (Responds to the prediction failure objection to simulation theory and restricts the domain of mental simulation to rational or otherwise intelligible transitions between mental states with semantic content.)
- Heal, J. (1996b) 'Simulation, Theory, and Content', in P. Carruthers and P. K. Smith (eds) *Theories of Theories of Mind*, Cambridge: Cambridge University Press, 75–89. (Discusses the scope of the simulation approach and argues that it is central in dealing with the attribution of states with content.)
- *Heal, J. (1998a) 'Understanding Other Minds from the Inside', in A. O'Hear (ed.) *Contemporary Issues in Philosophy of Mind*, Cambridge: Cambridge University Press, 83–99. (Defends a version of simulation theory that gives central importance to rationality and stresses the role of folk psychological understanding in co-operative enterprises.)
- *Heal, J. (1998b) 'Co-Cognition and Off-Line Simulation: Two Ways of Understanding the Simulation Approach', *Mind and Language* 14: 477–98. (Distinguishes between simulation theory as an a priori claim about the relations between various personal-level cognitive abilities and empirical hypotheses about the workings of subpersonal-level cognitive machinery.)
- Heal, J. (2000) 'Other Minds, Rationality and Analogy', *Proceedings of the Aristotelian Society, Supplementary Volume* 74: 1–19. (Stresses the connection between simulation and rationality and opposes versions of simulation theory that rely on the idea of an argument by analogy from self to other.)
- *Leslie, A.M. and German, T.P. (1995) 'Knowledge and Ability in "Theory of Mind": One-eyed Overview of a Debate', in M. Davies and T. Stone (eds) *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell Publishers, 123–50. (Argues that psychological knowledge, and not just psychological ability, is required to account for the child's acquisition of mature folk psychological understanding.)
- *Nichols, S., Stich, S., Leslie, A. and Klein, D. (1996) 'Varieties of Off-Line Simulation', in P. Carruthers and P. K. Smith (eds) *Theories of Theories of Mind*, Cambridge: Cambridge University Press, 39–74. (Reviews simulation-based accounts of a number of cognitive abilities including counterfactual reasoning, empathy and mental imagery and presents experimental results in support of the prediction failure objection against the simulation theory.)
- *Stich, S. and Nichols, S. (1992) 'Folk Psychology: Simulation or Tacit Theory?', *Mind and Language* 7: 35–71; repr. in Davies and Stone (eds) (1995a). (An important and detailed critical assessment of simulation theory as presented by Gordon and Goldman and defence of the theory theory; includes an early presentation of the prediction failure objection.)
- *Stich, S. and Nichols, S. (1995) 'Second Thoughts on Simulation', in M. Davies and T. Stone (eds) *Mental Simulation: Evaluations and Applications*, Oxford: Blackwell Publishers, 87–108. (Reviews different versions of the competing theories and allows that simulation theory can give a plausible account of the prediction of beliefs; develops the prediction failure objection against the simulation theory of behaviour prediction.)
- Stone, T. and Davies, M. (1996) 'The Mental Simulation Debate: A Progress Report', in P. Carruthers and P. K. Smith (eds) *Theories of Theories of Mind*, Cambridge: Cambridge University Press, 119–37. (Suggests that there are important differences amongst the questions that are being addressed by participants in the

debate, offers a reply to the prediction failure objection, and draws attention to the normative theory theory.)

*Walton, K.L. (1997) 'Spelunking, Simulation and Slime: On Being Moved by Fiction', in M. Hjort and S. Laver (eds) *Emotion and the Arts*, Oxford: Oxford University Press, 37–49. (Argues, with vivid examples, that emotional responses to works of fiction result from engaging in mental simulation and that these responses reveal something about the simulator's actual, and not just imagined, personality.)

MARTIN DAVIES
TONY STONE