**In press: Cognitive Unconscious and Human Rationality, Edition: 1, Chapter: 8, Publisher: The MIT Press, Editors: Laura Macchi, Maria Bagassi, Riccardo Viale, pp.167-180**

# The Basic Assumptions of Intuitive Belief: Laws, Determinism, and Free Will

Steven A. Sloman

Brown University

Aron K. Barbey

University of Illinois at Urbana-Champaign

Address correspondence to:

Steven Sloman

Department of Cognitive, Linguistic, and Psychological Sciences

Brown University, Box 1821

Providence, RI 02912

email: Steven_Sloman@brown.edu

Our goal in this chapter is to address the problems of determinism and free will. But don't give up on us quite yet. We do not purport to have anything to say about whether the world is deterministic or whether there are agents in it that have free will. Our goals are much more modest, to make some suggestions about what the human cognitive system assumes about determinism and free will. We want to know not whether determinism is true and free will exists, but whether people believe, at some level, in determinism and free will. Our task is much easier than that of a metaphysician because logical consistency is not axiomatic in psychology. Even if free will and determinism are in fact incompatible, people could still believe in both.

We also address a closely related question: Do people believe that the world is law governed? This is not the same as the question of determinism because laws are not necessarily deterministic, laws can be probabilistic (consider the laws of thermodynamics). It is also not the same as the question of free will because one might believe that free will is an exception to the laws that govern the world or one might believe that free will is itself governed by those laws. Alternatively, one might not believe in free will regardless of one's position on laws.

We think of these three questions as determining a set of basic assumptions made by the cognitive system. To make inferences, the cognitive system needs some direction, some driving force or principles of operation. We will call this a logic, and all logics make assumptions. What assumptions are made by whatever logic drives human inference? Our three questions are foundational. Any logic must take a position on them. Of course, the cognitive system could take different positions at different times. Indeed, it does. We have conversations in which we imagine the world is deterministic or it has free will, and we have conversations in which we imagine neither. But it seems plausible to us that there is a kind of basic, intuitive system for making inferences that is more consistent in the foundational assumptions it makes. It is that intuitive system – one whose operation we probably don't have conscious access to – that we focus on here.

**The space of possibilities**

We have 3 binary questions so there are 2 x 2 x 2 = 8 possible answers. People might believe the world is not law governed, not deterministic, and that there is no free will. Or they might believe that the world is law governed, not deterministic, and that there is no free will, or any other possible combination up to the possibility that they do believe in all three properties: that the world is law governed, that it is deterministic, and that there is free will.

These three questions are conceptually related in the sense that certain combinations of answers seem to make more sense than others. For instance, one could argue that a determinist should be more skeptical about free will than a non-determinist. But the questions are independent in the sense that all possibilities are up for grabs; one could take any position on any question regardless of the positions one takes on the others. Indeed, although we won't review it, we suspect that every combination of positions on the corresponding metaphysical questions is represented in philosophy (e.g., Hoefer, 2010; Salmon, 1998).

**Our hypothesis**

Before reviewing the evidence, we reveal our own position. On the first question, we believe that the cognitive system assumes that the world is law governed; in particular, the cognitive system assumes that the world is governed by causal laws, by mechanisms that take causes as inputs and deliver effects as outputs.

On the second question, we believe the cognitive system assumes that the world is deterministic. This is likely our most tendentious claim. Our claim, to be specific, is that the cognitive system believes that, whatever forces govern the world, they do so in such a way that a complete description of the world at time 1 entails a complete description of the world at time 2. There is no spontaneous noise that makes the world unpredictable if you have complete knowledge of it. I may accept quantum uncertainty, I may even accept as a matter of principle that there is randomness at the subatomic level. Our claim is that the intuitive cognitive system does not abide by this principle; it does not allow for quantum randomness (primarily because it does not understand it).

There is however a huge caveat to our claim: Belief in determinism is conditioned on having a complete description of the world. Yet we never have such a description, and we know it. So we allow for probability despite our belief that there are deterministic laws, and we justify our use of probability in terms of uncertainty. We know we are ignorant about some things, and that ignorance means that we can't be perfect predictors. For example, we may believe that smoking causes cancer deterministically without believing that every person who smokes will get cancer. After all, a smoker could get hit by a car at a young age and never have the opportunity to develop cancer. Such an event would be unknowable, not because it is inherently unpredictable, but because the world is too complicated. Given our limited cognitive resources and limited exposure to events, we cannot have enough knowledge to allow us to predict all the complex interactions that occur in the real world. So even though we believe that all events are determined by their causes, we also know that all events are impossible to predict because we are necessarily ignorant. Probability arises from ignorance of initial conditions or relevant variables or governing laws.

On the third question, we believe that people believe in free will. The cognitive system assumes that people make choices that they are responsible for, and we take that assumption to heart, even when making inferences about ourselves.

## The evidence

### People believe the world is law governed

In principle, the mind need not assume the world is law governed. It could assume for instance that all it need do is represent the structure in experience, whatever that structure is. Indeed, this is a common assumption in psychology and even in cognitive science. Historically, psychologists have tended to think the mind imposes as little structure as possible and so behaviorists such as Pavlov, Watson, and Skinner believed that animals of all kinds would learn whatever associations they were taught. Watson was famously bold about it:

> Give me a dozen healthy infants, well-formed, and my own specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select – doctor, lawyer, artist, merchant-chief and, yes, even beggar-man and thief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors. (Watson, 1930) (p. 82)

Learning was just a matter of internalizing experience according to one or another principle of learning and reward. It was not until the seminal work of Garcia (Hankins et al., 1976), Premack (Premack and Putney, 1962), and many others that the vacuity of this idea was fully realized. All animals, including people, are predisposed to learn certain kinds of relations and not others.

Garcia's work suggests that the predisposition is to learn relations that are consistent with the mechanisms that operate in their environment. So a rat will be predisposed to learn that illness is associated with a food eaten and not with a light but that a shock is associated with a light but not with food. In other words, the rat is more likely to learn the relations that make causal sense.

The tendency to rely on experience as the basis for mental representation did not die along with behaviorism during the cognitive revolution of the '70s. Connectionism allowed for more sophisticated learning, representational, and inferential processes than behaviorism but it retained the view that what is learned is what is experienced. In connectionist models, experience is represented in terms of correlations and higher-order correlations among elements of experience. Modern neural net models that are based largely on connectionist principles retain this property. They usually have structure, but the structure does not reflect any laws that might or might not govern the world; the structure reflects neural structure, the structure of the brain.

Another type of model that is current is the probability model. Probabilistic models come in more than one variety. Some (e.g., Anderson, 1993) are intended to be direct representations of the statistics of the world. Such models explicitly reject the notion of representing laws; they only represent outcomes and claim that the cognitive system is tuned to the relative frequency of states in the world.

Other models (e.g., Tenenbaum et al., 2006) try to have it both ways. They represent statistics, but in a Bayesian way. That is, their representations are biased in favor of prior probabilities. Whether or not such models assume that the world is law governed depends on where their priors come from. If their priors are "flat" (uniform distributions sometimes called ignorance or noninformative priors) or if they are chosen because they are mathematically tractable (e.g., conjugate priors; see Griffiths et al., 2010), then they are not assuming a law-governed world. At least, they are not assuming the world is governed by any kind of natural law but rather by a law intended to ease calculation.

But if the prior is chosen to reflect causal structure (Pearl, 2000; Spirtes, Scheines, and Glymour, 1993/2000), then it is making an assumption about natural law, namely that the world is governed by mechanisms that obey causal logic. Causal Bayes nets are Bayesian probabilistic models that represent the world using graphs intended to reflect causal structure. The structure is specifically causal in that it supports intervention: Intervening to impose a value on a variable in the model changes the structure of the model so that new inferences are only made about effects of the intervened-on variable, not its normal causes (Sloman, 2005, offers a painless introduction). Causal Bayes nets actually entail a causal logic. For instance, they imply that if A causes B and B causes C then A causes C (we do have to make additional assumptions, like the absence of a second causal path in which A inhibits C; Halpern and Hitchcock, 2013).

It is not our intention to argue that human reasoning relies on Causal Bayes nets, in fact we do not believe that it does (see Sloman & Lagnado, in press). But we do believe that some form of causal logic governs how people reason; that people are natural causal reasoners (Waldmann & Holyoak, 1992).

We reason very comfortably and naturally about causal structure and not about other kinds of logical structure. When engaging in everyday reasoning, we are bad at propositional reasoning, syllogistic reasoning, and probabilistic reasoning but good at qualitative causal reasoning (for supporting arguments, see Holyoak and Cheng, 2011; Chater and Oaksford, 2013).

Some of the evidence in favor of our position comes from demonstrations that causal knowledge trumps probabilistic knowledge in reasoning. Bes et al. (2012) report an example of such trumping. They tested the hypothesis that conditional probability judgments can be influenced by causal structure even when the statistical relations among variables are held constant. They informed participants in a series of experiments that a specific set of correlations held among three variables A, B, and C (for instance, a person's quality of sleep, their level of magnesium, and their muscle tone). They also offered an explanation for these correlations, a different explanation offered to each of three groups. Each explanation took the form of a causal model. One group was told the variables were correlated by virtue of a common cause ("An increase in the level of magnesium leads to an increase in the quality of sleep. An increase in the level of magnesium also leads to an increase in muscle tone"). Another group was told the correlations were due to a causal chain ("An increase in the level of magnesium leads to an increase in the quality of sleep which in turn leads to an increase in muscle tone") and a third to a diagnostic chain ("An increase in muscle tone leads to an increase in the quality of sleep which in turn leads to an increase in the level of magnesium"). They were all asked the same question: They were told that Mary, 35 years old, has good quality of sleep and were asked "According to you, what is the probability that Mary has good muscle tone?" In other words, they were asked to judge a conditional probability, P(C|A), whose value had not been given, but could easily be calculated from the data that they were given.

The results showed that conditional probability judgments were highest when the explanation was a causal chain, next when the explanation was a diagnostic chain, and lowest after the common cause explanation. Bes et al. also showed that the data could not be accounted for by a large class of Bayesian learning models, the most natural class to use. Specifically, they considered Bayesian models that assume that conditional probability judgments are informed not just by data but also by prior beliefs about causal structure. They also constrained the parameters of the models to a single value in the causal direction and a single value in the diagnostic direction (after all, any model with enough free parameters could fit the data perfectly after the fact). Such models predict that conditional probability predictions on the common cause model should be between those of the chain and diagnostic models. The reason they make this prediction is that the causal chain involves two causal inferences, the diagnostic chain entails two diagnostic inferences, whereas the common cause involves one causal and one diagnostic inference.

The fact that the common cause model engendered the lowest judgments suggests that people were not combining priors with new data in the way that Bayesianism prescribes. Instead, people were likely trying to understand the relation between the variables they were judging in qualitative terms. That relation was the most elaborate in the common cause case because reasoning involved both a diagnostic link from C to B and a causal link from B to A. The other causal models required tracing either two causal links or two diagnostic links. In other words, the conditional probability judgments were influenced by the difficulty of constructing an explanation from causal structure, not by Bayesian reasoning.

Another demonstration suggesting that causal knowledge is basic to human judgment was reported by Park and Sloman (2013). They investigated what kind of information people use to make predictions. Causal Bayes nets theory implies that people should follow a particular structural constraint called the Markov property or “the screening-off rule” when reasoning. We will describe the property informally. It states that if the causes of an effect are held constant, then the effect should be treated as independent of other variables related to the effect only via the cause. For example, consider a common cause structure in which B causes A and B also causes C. If the value of B is known, then A tells us nothing we don’t already know about C. So A and C should be independent conditional on B. But a variety of previous work shows that people frequently violate this principle (Chaigneau et al., 2004; Lagnado and Sloman, 2004; Rehder and Burnett, 2005; Waldmann and Hagmayer, 2005; Walsh & Sloman, 2008).

Park and Sloman (2014) showed that this violation only occurs when B causes both A and C in the same way, when the same causal mechanism is responsible (Park and Sloman, 2014). For instance, consider the following pair of causal relations:

> *Smoking more than three packs of cigarettes a day often causes impairment of lung function.*

> *Smoking more than three packs of cigarettes a day often causes damage to blood vessels.*

In both cases, the causal relation is supported by similar mechanisms, namely the tar and nicotine in cigarettes does damage to the body. But now consider a different pair of causal relations:

> *Smoking more than three packs of cigarettes a day often causes impairment of lung function.*
> *Smoking more than three packs of cigarettes a day often causes a financial burden on the family budget.*

In this case, the causal relations are supported by quite different mechanisms. The financial impact has to do with spending money, not damage to the body. Park and Sloman's main finding is that violations of screening-off only occurred for the first type of example, in which mechanisms were the same, not when they were different. Specifically, when people were asked for the probability that an individual would have damage to blood vessels given that they smoked 3 packs a day and their lung function was normal, they gave a higher estimate than when asked for just the probability that an individual would have damage to blood vessels given that they smoked 3 packs a day. In contrast, when asked for the probability that an individual would have a financial burden given that they smoked 3 packs a day and their lung function was normal, they gave the same estimate that they gave when asked for the probability of a financial burden given that they smoked 3 packs a day.

Our explanation for this finding is that people use knowledge about underlying mechanisms to infer latent structure to make conditional probability judgments, especially when their expectations are violated. When told that someone who smokes a lot has normal lung function, there is something to be explained. It can be explained by introducing a disabler (e.g., the person must have smoked highly filtered cigarettes) or by introducing a mediating mechanism (e.g., the person did not inhale). When the two mechanisms are the same, the disabler or mediating mechanism is likely to apply to the other effect too (if the person didn't inhale, then their blood vessels are also likely not damaged). But if the mechanisms are different, then the disabler or mediating mechanism has no implications for the other effect (not inhaling doesn't relieve the financial burden).

What these experiments show is that people construct a causal understanding based on the evidence they are given when they are given enough information to do so. Then they make probability judgments in a way that is compatible with their causal understanding. Both the Bes et al. (2012) and Park and Sloman (2014) studies imply that the causal understanding comes first, that the cognitive system is designed to generate a causal explanation and derive probability judgments from that causal explanations. Because causal explanations require the assumption that the world is governed by causal mechanisms and such mechanisms are manifestations of natural laws, we conclude that the cognitive system assumes the world is law governed.

Further evidence that causal understanding comes before probabilistic knowledge comes from the literature on causal learning. Only rarely and with some difficulty are people able to induce causal structure from probabilistic data. Most of the time people rely on single-event cues to causal relations, cues that are grounded in perception (White, 2014), temporal relations, instruction, or intervention (for a review, see Lagnado, Waldmann, Hagmayer, & Sloman, 2007).

Causal knowledge represents our understanding of how the world works, capturing systematic (i.e., law governed) relationships. Rather than being reducible to associative links, causal knowledge is based on mental representations of cause-effect relations that reflect the workings of the external world. So if we are right that the logic of intuition is causal, then we can conclude that, at the most fundamental level, the human reasoning system assumes that the world is governed by natural laws, namely causal laws.

**People are deterministic. Probability arises from ignorance of initial conditions or governing laws**

We propose that people are determinists for whom probability originates from ignorance about causal laws rather than representing a central feature of causal knowledge. This is an old and established position articulated by Laplace (1814/1902) among others including Pearl (2000). We address the evidence in favor of the determinist assumption and the origins of probability separately.

**Evidence in favor of determinism.** Accumulating evidence indicates that people have a deterministic representation of causal laws. Causal determinism is the idea that all events have causes (Friedman, 1982; Hirschfeld and Gelman, 1994). It predicts that people will infer unobserved causes whenever events appear to occur spontaneously. There is considerable evidence that both adults and children do this (Chandler and Lalonde, 1994;

Friedman, 1982). For example, Schultz and Sommerville (2006) report a series of experiments demonstrating that children infer unobservable causes whenever events appear to occur spontaneously. In these experiments, a toy is placed on a device that illuminates only when the switch on the device is turned ON (deterministic condition) or when the switch is turned ON or OFF (stochastic condition) and children are asked to perform an action to enable or disable the device. Rather than inferring probabilistic causation, their actions indicate that children assume that causes produce their effects deterministically. They believe in the stochastic condition that an inhibitory cause is at work (e.g., there is a device in the experimenter's hand that blocks the effect of the switch). Thus, the reported findings indicate that children resist believing direct causes can act stochastically, instead preferring a deterministic representation of causal laws.

Other evidence in favor of determinism is less direct. English speakers recognize there is a difference in meaning between causes and enabling conditions. Yet the two are identical in their effect on probability; both increase the probability of their associated effect. The difference between them cannot be described in terms of probability (Cheng and Novick, 1991; Wolff et al., 2010) and so must have some other basis. A likely possibility is that they differ in the role they play in causal mechanisms (Sloman et al., 2009). Such differences in role are consistent with a deterministic model.

Another argument comes from learning. Reasoners often infer a causal relation from a single observation (e.g., Schlottmann and Shanks, 1992; Ahn et al., 1995; White, 2014). But, if causal assertions are probabilistic, single observations should rarely suffice to establish cause and effect because probabilistic representations tolerate exceptions. Single observations may be sufficient in the context of extremely strong prior beliefs that require very little additional information to compel a conclusion. But learning from a single instance occurs more often than such a view can justify.

Furthermore, people are happy to conceive of interventions that initiate a causal chain in a deterministic way (Pearl, 2000; Tenenbaum and Griffiths, 2001; Gopnik et al., 2004; Hagmayer and Sloman, 2009). Interventions are actions that, on most views, cannot even be assigned a probability (Spohn, 1977) and are generally modeled using deterministic operators.

**Evidence that probability is associated with ignorance**. There is no direct evidence that people believe that probability originates in ignorance; perhaps such evidence is impossible. Nevertheless, the claim is at least consistent with data showing that changing people's feelings of ignorance influence their judged probability. The Ellsberg Paradox (Ellsberg, 1961) is that people prefer to bet on an urn of known distribution than one of unknown distribution. In the simplest case, you are offered two urns, one with 50 balls of one color (say red) and 50 balls of a different color (say blue). The other urn also has 100 balls and all of them are either red or blue, but you are not told how many there are of each color. You are told you can pick an urn, one ball will be selected from that urn, and if it's red, you will win $50. Which urn do you choose? Given the reasonable assumption of symmetry – that anything relevant to the likelihood of picking a red ball applies equally to picking a blue ball – the effective probability of winning from either urn is .5. The probability of winning from either urn is identical, and yet most people choose the first urn, the urn of known composition. The greater knowledge (of the composition of balls) associated with the first urn makes people feel more confident or certain about the first urn and therefore more willing to bet on it.

Heath and Tversky (1991) showed that willingness to bet on an event is proportional to one's sense of competence regarding the event, where competence refers to what one knows relative to what could be known. So less ignorance is associated with a greater willingness to bet, and willingness to bet is an operational definition of subjective probability (Savage, 1954).

More evidence that feelings of ignorance are associated with judged probability was reported by Fox and Tversky (1995). They showed that the Ellsberg paradox only occurred in within-participants designs, not when people were stating their willingness to bet on each urn separately. They call this "comparative ignorance"; it is the comparison to another event that induces the sense of ignorance. Similar demonstrations of reduced probability judgment come from experiments varying whether more knowledgeable individuals happen to be in the room. Simply stating that an expert is present makes people feel less confident, presumably by making them feel less knowledgeable.

**There is free will**

Evidence from psychology indicates that the belief in free will originates from the subjective experience of choosing and acting. Monroe and Malle (2010) investigated the psychological foundations of free will, providing evidence that people do believe in free will and conceptualize it as a choice that satisfies their desires, free from constraints. Even when faced with evidence to suggest that free will is an illusion (i.e., that behavior is caused by neural impulses that precede our impression of agency), people nonetheless defend the notion that free will is central to human thought and behavior. Indeed, Nichols (2004) provides evidence to suggest that free will originates in childhood from the attribution of choice and agency to human actors. For example, children attribute agency to a human actor but not to an object, believing that the human actor could have acted differently (based on free will), whereas the object must behave deterministically (Nichols, 2004).

Recent psychological evidence further indicates that the belief in free will is central to moral judgment, providing a basis for holding people responsible for their actions (Mele, 2009). In a series of experiments, Clark et al. (2014) provide evidence that the belief in free will is modulated by the desire to hold others morally responsible for their wrongful behaviors. For example, participants report greater belief in free will after considering an immoral action than a morally neutral one, consistent with the desire to punish morally wrongful behavior. The belief in free will may therefore be motivated by the human desire to blame and punish others for wrongful behavior.

More indirect evidence concerning people's assumptions about free will comes from work on intervention that was alluded to above. Interventions afford different inferences than observations (Spirtes et al., 1993; Meek and Glymour, 1994; Pearl, 2000). In particular, they afford different diagnostic inferences about the causes of an event. For instance, observing someone drink 10 shots of whiskey in a row does make it more likely that the person is an alcoholic; intervening to make the person drink 10 shots of whiskey provides no such evidence. People are exquisitely sensitive to this difference in the diagnostic inferences that are afforded by observation versus intervention; they draw different conclusions in the two cases when other things are equal (Sloman and Lagnado, 2005; Waldmann and Hagmayer, 2005; Hagmayer and Sloman, 2009).

Intervention is not easy to define (Woodward, 2003) and what constitutes an intervention may well be different in different situations. We would all agree that an experiment is an example of an intervention but we might disagree about whether there are “natural” experiments (does a hurricane provide an unconfounded test of a tree’s resilience?). We can state one sufficient condition for an intervention: an action taken by virtue of an agent’s own free will. The locus of the action is then an intervention. The data showing that people make appropriate inferences from intervention suggest that people can and do represent interventions in a way distinct from observations. To the extent that such interventions arise from free will, representing them presupposes the existence of free will. In that sense, people do act as if they and others have free will.

## Conclusion

This chapter is speculative in that all of our conclusions are defeasible. None of the evidence is overly compelling. But it does paint a sensible portrait of human cognition. We propose that human intuition is premised on 3 assumptions: that people believe the world is law-governed, deterministic, and that free will exists. These are all commonsensical assumptions that we suspect would fly unnoticed at most dinner table conversations. We are well aware that uncertainty abounds and that people realize that. But the cognitive system could easily attribute that uncertainty to ignorance. It’s hard to make predictions, especially about the future, and it’s hard to know things, especially about the world. There’s just too much to know.

What is more surprising to us is how much cognitive theory does not respect these simple, basic assumptions. Probabilistic models are fine either as approximations or as models of knowledge and inference in the face of uncertainty. But they should be grounded in a representation that, with sufficient knowledge, would be law-governed and deterministic. We would go so far as to suggest that the basic problem for cognitive psychology is to understand how causal laws are represented and processed.

The implications of belief in free will are less obvious, and less manageable using the mechanistic tools of cognitive science. One view is that free will arises from the quantum nature of the brain but we fail to see how this solves any problems. At minimum, it suggests that we need to take seriously representations that distinguish intervention from observation (Pearl, 2000).

**References**


Ahn, W.K., Kalish, C.W., Medin, D.L., and Gelman, S.A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition* 54**,** 299-352.

Anderson, J.R. (1993). *Rules of the mind.* Hillsdale, N.J.: L. Erlbaum Associates.

Bes, B., Sloman, S., Lucas, C.G., and Raufaste, E. (2012). Non-Bayesian inference: causal structure trumps correlation. *Cogn Sci* 36**,** 1178-1203.

Chaigneau, S.E., Barsalou, L.W., and Sloman, S.A. (2004). Assessing the causal structure of function. *Journal of Experimental Psychology-General* 133**,** 601-625.

Chandler, M.J. and Lalonde, C.E. (1994). Surprising, miraculous, and magical turns of events. *British Journal of Developmental Psychology* 12**,** 83-95.

Chater, N., and Oaksford, M. (2013). Programs as causal models: speculations on mental programs and mental representation. *Cogn Sci* 37**,** 1171-1191.

Cheng, P.W., and Novick, L.R. (1991). Causes versus enabling conditions. *Cognition* 40**,** 83-120.

Clark, C.J., Luguri, J.B., Ditto, P.H., Knobe, J., Shariff, A.F., and Baumeister, R.F. (2014). Free to punish: A motivated account of free will belief. *J Pers Soc Psychol* 106**,** 501-513.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75**,** 643-669.

Fox, C. R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *The Quarterly Journal of Economics*, *110*(3), 585-603.

Friedman, W.J. (1982). *The developmental psychology of time.* New York: Academic Press.

Glymour, C., Scheines, R., Spirtes, P.. (1993/2001). Causation, Prediction, and Search - 2nd Edition", in: *Adaptive Computation and Machine Learning.* (Cambridge, Mass.: MIT Press,).

Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T., and Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review* 111**,** 3-32.

Griffiths, T.L., Chater, N., Kemp, C., Perfors, A., and Tenenbaum, J.B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends Cogn Sci* 14**,** 357-364.

Hagmayer, Y., and Sloman, S.A. (2009). Decision makers conceive of their choices as interventions. *Journal of Experimental Psychology-General* 138**,** 22-38.

Halpern, J.Y., and Hitchcock, C. (2013). Compact representations of extended causal models. *Cogn Sci* 37**,** 986-1010.

Hankins, W.G., Rusiniak, K.W., and Garcia, J. (1976). Dissociation of odor and taste in shock-avoidance learning. *Behav Biol* 18**,** 345-358.

Heath, C., Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty* 4**,** 5-28.

Hirschfeld, L.A., and Gelman, S.A. (1994). *Mapping the mind : domain specificity in cognition and culture.* Cambridge ; New York: Cambridge University Press.

Hoefer, C. (2010). Causal determinism. *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2010/entries/determinism-causal/>.

Holyoak, K.J., and Cheng, P.W. (2011). Causal learning and inference as a rational process: the new synthesis. *Annu Rev Psychol* 62**,** 135-163.

Lagnado, D.A., and Sloman, S. (2004). The advantage of timely intervention. *J Exp Psychol Learn Mem Cogn* 30**,** 856-876.

Lagnado, D. A., Waldmann, M. R., Hagmayer Y., & Sloman, S.A., (2007). Beyond covariation: Cues to causal structure. In Gopnik, A. & Schulz, L. (Eds.), Causal learning: Psychology, philosophy, and computation. Oxford: Oxford University Press, pp. 154-172.

Laplace, P.S. (1902). *A philosophical essay on probabilities.* New York: J. Wiley.

Meek, C., Glymour, C. (1994). Conditioning and intervening. *The British Journal for the Philosophy of Science* 45**,** 1001-1021.

Mele, A. R. (2009). *Effective intentions: The power of conscious will*. New York: Oxford University Press.

Monroe, A. E. & Malle, B. (2010). From uncaused will to conscious choice: The need to study, not speculate about people's folk concept of free will. *Rev Phil Psych* 1, 211-224.

Nichols, S. (2004). Folk concepts and intuitions: from philosophy to cognitive science. *Trends Cogn Sci* 8**,** 514-518.

Park, J. & Sloman, S. A. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology, 67*, 186-216.

Park, J. & Sloman, S. A. (2014). Causal explanation in the face of contradiction. *Memory and Cognition, 1*, 1-15.

Pearl, J. (2000). *Causality : models, reasoning, and inference.* Cambridge, U.K. ; New York: Cambridge University Press.

Premack, D., & Putney, R.T. (1962). Relation between intersession interval frequency of competing responses and rate of learning. *J Exp Psychol* 63**,** 269-274.

Rehder, B., & Burnett, R.C. (2005). Feature inference and the causal structure of categories. *Cogn Psychol* 50**,** 264-314.

Salmon, W.C. (1998). "Causality and explanation". (New York ; Oxford: Oxford University Press,).

Savage, L.J. (1954). *The foundations of statistics.* New York,: Wiley.

Schlottmann, A., and Shanks, D.R. (1992). Evidence for a distinction between judged and perceived causality. *Q J Exp Psychol A* 44**,** 321-342.

Schulz, L.E., and Sommerville, J. (2006). God does not play dice: causal determinism and preschoolers' causal inferences. *Child Dev* 77**,** 427-442.

Sloman, S., Barbey, A.K., and Hotaling, J.M. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cogn Sci* 33**,** 21-50.

Sloman, S.A. (2005). "Causal models how people think about the world and its alternatives". (Oxford: Oxford University Press,).

Sloman, S. A. & Lagnado, D. (in press). Causality in thought. *Annual Review of Psychology*.

Sloman, S.A., and Lagnado, D.A. (2005). Do We "do"? *Cogn Sci* 29**,** 5-39.

Spirtes, P., Glymour, C.N., and Scheines, R. (1993/2000). *Causation, prediction, and search.* New York: Springer-Verlag.

Spohn, W. (1977), Where Luce and Krantz do really generalize Savage's decision model, *Erkenntnis,* 11, pp. 113-134.

Tenenbaum, J.B., and Griffiths, T.L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences* 24**,** 629-640; discussion 652-791.

Tenenbaum, J.B., Griffiths, T.L., and Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn Sci* 10**,** 309-318.

Waldmann, M.R., and Hagmayer, Y. (2005). Seeing versus doing: two modes of accessing causal knowledge. *J Exp Psychol Learn Mem Cogn* 31**,** 216-227.

Waldmann, M.R., and Holyoak, K.J. (1992). Predictive and diagnostic learning within causal models: asymmetries in cue competition. *Journal of Experimental Psychology-General* 121**,** 222-236.

Walsh, C. & Sloman, S. A. (2008). Updating beliefs with causal models: Violations of screening off. Gluck, M. A., Anderson, J. R , & Kosslyn, S. M. (Eds.). *Memory and Mind: A Festschrift for Gordon H. Bower*. New Jersey: Lawrence Erlbaum Associates.

Watson, J.B. (1930). *Behaviorism.* New York,: W.W. Norton & company, inc.

White, P.A. (2014). Singular clues to causality and their use in human causal judgment. *Cogn Sci* 38**,** 38-75.

Wolff, P., Barbey, A.K., and Hausknecht, M. (2010). For want of a nail: How absences cause events. *Journal of Experimental Psychology-General* 139**,** 191-221.

Woodward, J. (2003). *Making things happen : a theory of causal explanation.* Oxford ; New York: Oxford University Press.