

# Can reliabilism explain how conscious reflection justifies beliefs?

Dr. S. Kate Devitt

Queensland University of Technology

## Abstract

This poster addresses the justificatory role of conscious reflection within a naturalized, reliabilist epistemology. Reliabilism is the view that implicit, mechanistic (System 1) processes can justify beliefs, e.g. perceptual beliefs formed after a history of consistent exposure to normal lighting conditions are justified in a given context with normal lighting. A popular variant of reliabilism is virtue epistemology where the cognitive circumstances and abilities of an agent play a justificatory role, e.g. the cooperation of the prefrontal cortex and primary visual cortex of the individual perceiving the Müller-Lyer illusion partly justify the belief that the lines are equi-length. While virtue epistemology is a well-endorsed reliabilism for implicit beliefs, its application to explicit, consciously reflective (System 2) processes is more controversial. Critics ask: How can iterations of dumb reliabilist processes produce higher order justification? To respond to this concern, I draw on another agent-centred, normative and reliabilist epistemology—Bayesian epistemology. A Bayesian virtue epistemology argues that reflective hypothesis-testing generated by (largely) implicit Bayesian mechanisms offers higher order reliabilist justification for beliefs. Iterative Bayesian mechanisms (e.g. hierarchically nested probabilistic models) explain the development of higher order beliefs about abstract concepts such as causation, natural laws and theoretical entities traditionally explained by recourse to vague concepts such as ‘the a priori’, ‘intuition’ or ‘the intellect’. A hybrid Bayesian virtue epistemology offers an iterative reliabilist framework to explain how conscious reflection justifies beliefs. However, I acknowledge limitations on Bayesian accounts of justification such as confirmational holism, commutativity, and the frame problem.

## Glossary

**Conscious reflection:** Top-down reasoning that requires attention and leads to understanding and explicit confidence in the justification of a belief.

**Naturalized:** A philosophical method whereby philosophers consider empirical experiments and rigorous models by cognitive scientists rather than depend exclusively on their naive or intuitive conceptions of human functioning.

**Reliabilism:** “S’s belief in p is justified iff it is caused (or causally sustained) by a reliable cognitive process, or a history of reliable processes” (Goldman, .

**Virtue epistemology:** A variant of reliabilism in which the cognitive circumstances and abilities of an agent play a justificatory role

**Bayesian epistemology:** typical beliefs exist (are performed and evaluated) in degrees, rather than absolutes, represented as credence functions.

## Virtue Epistemology

ANIMAL KNOWLEDGE:

*System 1.* Operates automatically and quickly, with little or no effort and no sense of voluntary control (Kahneman, 2011).

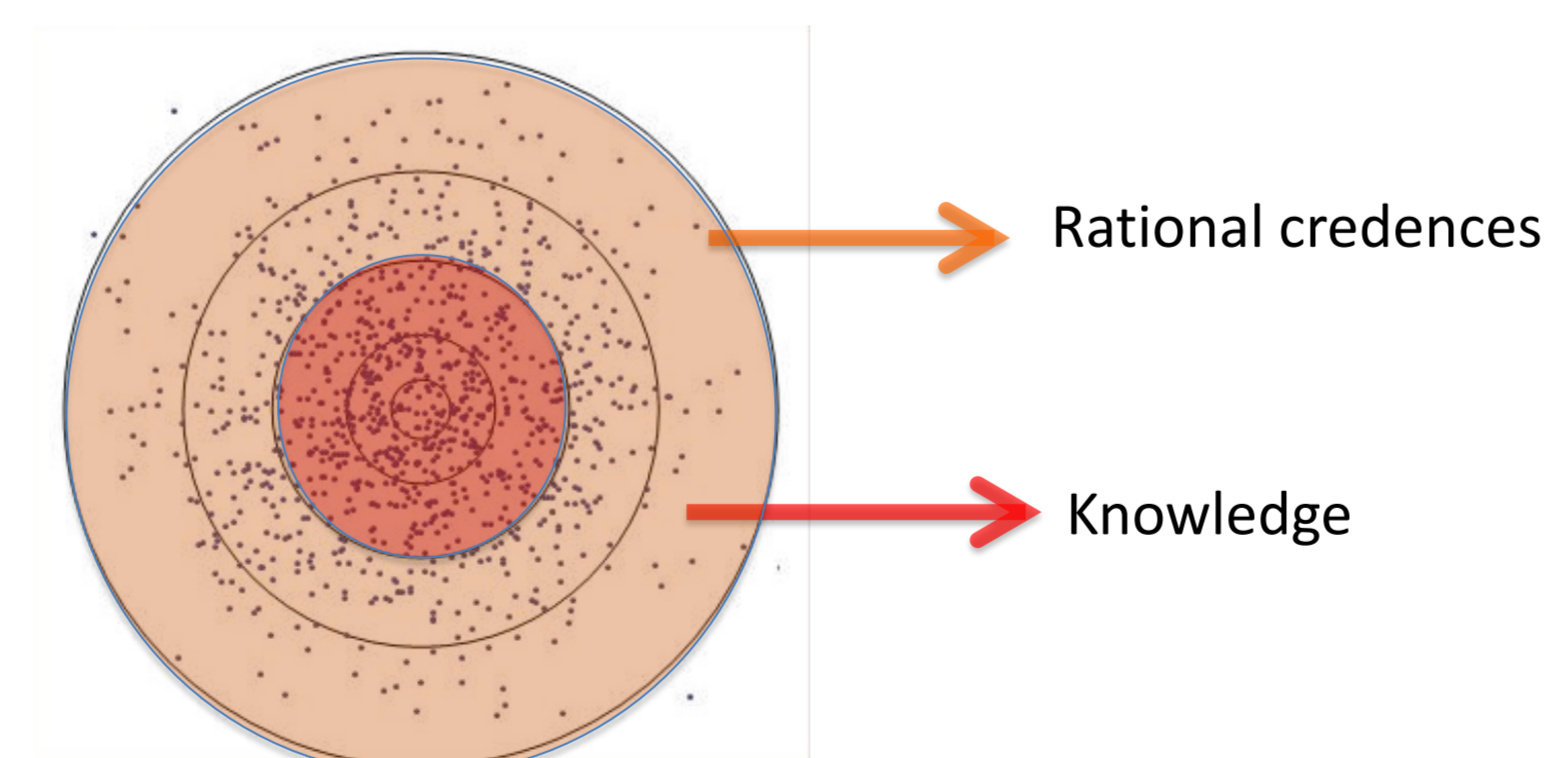
REFLECTIVE KNOWLEDGE:

*System 2:* Allocates attention to the effortful mental activities that demand it, including complex computations. The operations of system 2 are often associated with the subjective experience of agency, choice, and concentration (Kahneman, 2011).



**Figure #1** An agent has animal knowledge if belief p is accurate, the agent is adroit, and p is accurate due to adroit processes. An agent has reflective knowledge p if they have animal knowledge p, p coheres with other beliefs and the agent has a meta-apt understanding of why it is true and “how in which it is sustained as reliably truth-conducive” (Sosa, 2009, p.138).

## Bayesian Virtue Epistemology



**Figure #2** Bayesian virtue epistemology (Devitt, 2013) values justification brought by reliable competencies producing both rational credences and knowledge.

**Reflective knowledge:** If one is to know that h, then h must be reliably produced, be true and the most likely amongst a set of plausible hypotheses  $h_m$  to  $h_n$  that cohere with the rest of one’s beliefs at lower likelihoods (Devitt, 2013).

## Reliabilist Reflective Knowledge

A Bayesian virtue account of reflective knowledge aligns with work in cognitive psychology on how people develop theories of how the mind works—the so-called ‘theory theory.’ The theory theory attempts to explain higher order belief revision by proposing that all humans, from babies to adults, come to understand the world abstractly by forming hypotheses and then rigorously testing them and updating them in light of evidence. Hypotheses explain and justify beliefs with higher order beliefs about causality, unobservable theoretical entities, relations and so on. Recently (Gopnik & Wellman, 2012), the ‘theory theory’ research program has merged with new work being done on *hierarchically nested probabilistic models* (HNPM) to show how complex thoughts can be achieved through iterations of the same justificatory processes that underlie basic probabilistic processes.

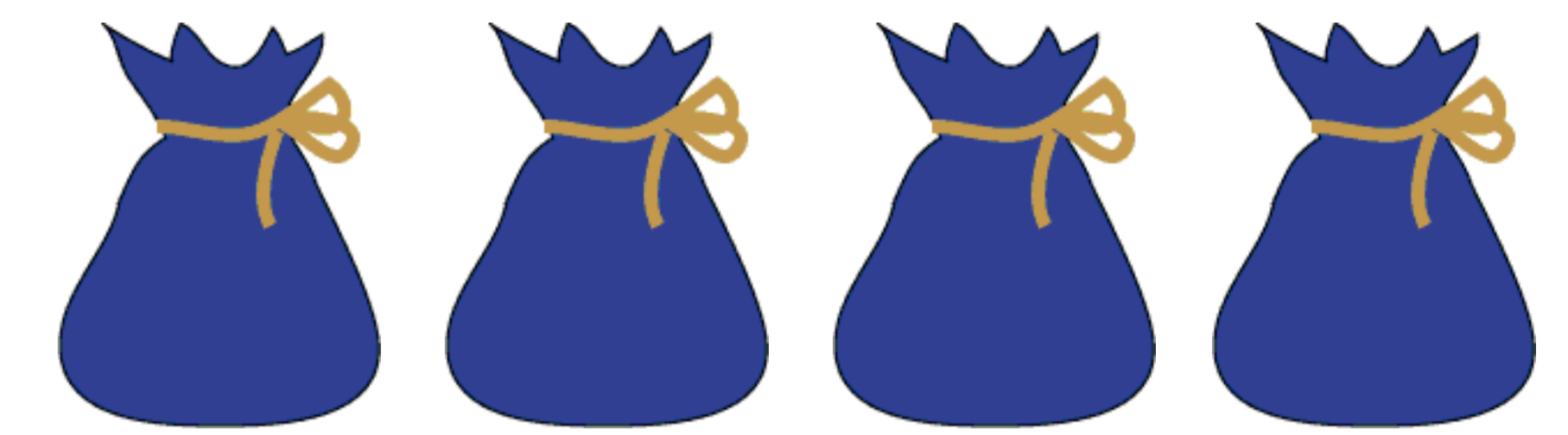


**Figure #3** Children learn abstract generalizations relating to higher level principles, specific theories and details about particular experiences during the same exposure to stimuli (Gopnik & Wellman, 2012).

With an HNPM (as opposed to typical probability models) the *same* inputs impact *multiple* hypotheses across many levels of the network using the *same* Bayesian calculus to achieve layered outputs. HNPM can account for multiple levels of knowledge, including:

1. Abstract generalizations relating to higher level principles
2. Specific theories about a set of instances
3. Particular experiences.

HNPM explains how abstract generalizations arise from specific theories that are, in turn, learnt from particular experiences. What is strikingly different about these models (compared with classic empiricist, foundational accounts of knowledge acquisition) is that abstract generalizations can *precede* specific ones.



**Figure #4** Suppose many bags of marbles were placed in front of you and your job was to identify the color of the marbles in each bag (Gopnik & Wellman, 2012). An experimenter takes a red marble out of the first bag and asks you what color you think the next marble will be. She then goes on removing marbles from that bag, pausing between each one to ask the same question, and each time the marble is red. The experimenter then repeats the procedure with the second bag of marbles. This time a succession of blue marbles appears. You will quickly assume that all the bags of marbles contain only one color. Predicting that the contents of any bag of marbles will match the color of the first one is an abstract generalization—an *overhypothesis*. Learning this *overhypothesis* precedes learning about the contents of bag 3 (for example, that all marbles are purple).

## Conclusion

In sum, reflective knowledge is not iterations of dumb reliabilist processes. New work in HNPM suggests that many aspects of higher-level knowledge can be created iteratively from low-level justificatory processes. Newer probabilistic models can restrict the scope of hypotheses considered and explain how multiple levels of knowledge can be learnt at once. Brute, reliabilist processes can generate reflective knowledge if they yield cohering beliefs across a broad explanatory domain.

S. Kate Devitt | @skdevitt | <http://skdevitt.com>

### References

- Devitt, S. K. (2013). *Homeostatic epistemology: Reliability, coherence and coordination in a Bayesian virtue epistemology*. Ph.D., Rutgers The State University of New Jersey - New Brunswick. Retrieved from <http://eprints.qut.edu.au/62553/> QUT ePrints database.
- Goldman, A. I. (1994). Naturalistic epistemology and reliabilism. In P. French, T. Uehling & T. Wettstein (Eds.), *Midwest Studies in Philosophy XIX: Philosophical Naturalism*. Minneapolis: University of Minnesota Press.
- Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, 138(6), 1085-1108. doi: 10.1037/a0028044
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux
- Sosa, E. (2009). *Reflective Knowledge: Apt Belief and Reflective Knowledge* (Vol. 2). Oxford, UK: Oxford University Press.