

# NEH Application Cover Sheet (FA-232860)

## Fellowships for University Teachers

### PROJECT DIRECTOR

---

Michael Rescorla

(b) (6)

(b) (6)

(b) (6)

**E-mail:** rescorla@philosophy.ucsb.edu

**Phone:** (b) (6)

**Fax:**

**Field of expertise:** Philosophy of Science

### INSTITUTION

---

University of California, Santa Barbara

Santa Barbara, CA 93106-2050

### APPLICATION INFORMATION

---

**Title:** *Bayesian Modeling of the Mind: Conceptual and Explanatory Foundations*

**Grant period:** From 2016-09-01 to 2017-08-31

**Project field(s):** Philosophy of Science

**Description of project:** Illuminating how the mind works has been a central concern of humanistic research stretching back to Plato. I seek to advance this enterprise by analyzing Bayesian cognitive science, a scientific research program that models the mind using probabilities. My analysis hinges upon the mind's capacity to represent the world. I will argue that Bayesian cognitive science assigns a central explanatory role to mental representation. Bayesian modeling reveals that core mental activities such as perception, action, and decision-making all crucially involve representational aspects of mentality. My analysis should advance our understanding of the mind by establishing that mental representation is an indispensable theoretical notion. As an illustrative case study, I will discuss Bayesian modeling of autism. My discussion of this case study should clarify some important points of similarity and difference between typically developing individuals and individuals with autism.

### REFERENCE LETTERS

---

Tyler Burge

Distinguished Professor

Philosophy

University of California, Los Angeles

burge@ucla.edu

Mohan Matthen

Professor and Canada Research Chair

Philosophy

University of Toronto

mohan.matthen@utoronto.ca

# NEH Supplemental Information for Individuals

This form should be used by applicants to the NEH Fellowships, Fellowships for Advanced Social Science Research on Japan, Awards for Faculty, and Summer Stipends Programs.

Field of Project:

Field of Project #2:

Field of Project #3:

Project Director Field of Study:

The mailing address provided on the SF 424-Individual is for your  work  home

## Institutional Affiliation

Are you affiliated with an institution? (If yes, provide information below.)  Yes  No

Institution Name:

Street 1:

Street 2:

City:

County:

State:

Province:

Country:

Zip / Postal Code:

DUNS Number:

Employer/Taxpayer Identification Number (EIN/TIN):

Status:  Senior Scholar  Junior Scholar

# Reference Letters

## Reference 1

First Name:

Last Name:

Email:

Title:

Department Name:

Institution:

## Reference 2

First Name:

Last Name:

Email:

Title:

Department Name:

Institution:

## Nominating Official (Summer Stipends Applicants Only)

Are you exempt from nomination? If not, provide information below.  Yes  No

First Name:

Last Name:

Email:

Title:

Institution:

## **Bayesian Modeling of the Mind: Conceptual and Explanatory Foundations**

### ***Theoretical background***

Rev. Thomas Bayes was an 18<sup>th</sup> century English mathematician and theologian who passed his adult life in relative obscurity as a minister in a provincial town. After Bayes's death in 1761, his friend Richard Price found among his papers an unpublished essay on probability. Recognizing the essay's immense significance, Price saw to its posthumous publication. Bayes's insights gave birth to *Bayesian decision theory*, which has become the standard mathematical model of reasoning and decision-making under uncertainty.

The core notion of Bayesian decision theory is *subjective probability* --- a quantitative measure of the degree to which an agent believes some statement. I may assign a very low subjective probability to the statement that a meteor shower will occur today, a higher subjective probability to the statement that it will rain today, a still higher subjective probability to the statement that a Republican will win the 2016 presidential election, and an even higher subjective probability to the statement that a presidential election will occur in 2016. As I acquire new evidence, I typically revise my subjective probabilities. For example, if I look outside and see that the sky is overcast, then I may assign a higher probability to the statement that it will rain today. Bayesian decision theory codifies these intuitive ideas through precise mathematical rules that govern how rational agents should update their subjective probabilities. The resulting framework has proven remarkably fruitful within statistics, medical science, robotics, economics, cognitive science, and philosophy.

Bayesianism originated as a theory of how people *should* reason, not a theory of how they *actually* reason. (Analogy: ethics studies how people should behave, without pretending to describe how they actually behave.) Nevertheless, cognitive scientists increasingly use Bayesian models to describe the workings of the human brain. On a Bayesian approach, the brain updates subjective probabilities in a rational or near-rational fashion. Bayesian models offered within cognitive science have clarified numerous mental phenomena, such as *perception, action, decision-making, and social cognition*.

### ***Research and contribution***

I contend that Bayesian cognitive science has striking implications regarding the human mind, implications hitherto unrecognized by scientists and philosophers. More specifically, Bayesian modeling decisively advances a debate that has raged in the philosophical community for the past few decades over *mental representation*.

Just as the heart serves to pump blood and the stomach serves to digest food, one of the mind's principal functions is to represent the world. For instance, I have various beliefs about Barack Obama: that he is president of the United States, that he is married to Michelle Obama, and so on. These beliefs represent Barack Obama as being a certain way. Thus, the mind somehow reaches beyond itself to external reality, depicting the world as having certain features. In that sense, the mind is a *representational organ*. John Locke, David Hume, Immanuel Kant, and many other philosophical giants have emphasized the mind's representational capacity as one of its most important properties. Building on this tradition, contemporary *representationalists* such as Tyler Burge and Jerry Fodor hold that mental representation should play a vital role within the explanation of various core mental phenomena. *Anti-representationalists* such as Paul Churchland and W. V. Quine hold that representational notions are too obscure or otherwise problematic to figure in good explanations. Anti-representationalists often claim that a proper science of the mind should exclusively discuss networks of neurons, without mentioning what those networks represent.

I will defend representationalism. My main thesis is that Bayesian cognitive science assigns a central role to mental representation. To attach a probability to some state of affairs, one must be able to represent that state of affairs. For example, one can attach a probability to a Republican winning the 2016 presidential election only if one can represent the possibility that a Republican wins the 2016 presidential election. Thus, I claim that Bayesian modeling presupposes the picture of the mind as a representational organ. Bayesian cognitive science invests that picture with unprecedented scientific substance, elaborating it into mathematically rigorous models. Admittedly, Bayesian cognitive scientists do not usually *explicitly* mention mental representation. Nevertheless, I will argue that their research *implicitly presupposes* mental representation. Bayesian models reveal that core mental activities such as *perception, action, decision-making, navigation, speech comprehension, and social cognition* all crucially involve representational mental states. We cannot even begin to understand these activities if we overemphasize neurons to the exclusion of representational relations between mind and world. While neurons are certainly very important, they do not tell nearly the whole story about the human mind.

Illuminating how the mind works has been a central concern of humanistic research stretching back to Plato. I seek to advance this enterprise by establishing *mental representation* as an indispensable theoretical notion. I will show how the traditional picture of the mind as a representational organ can be integrated into an empirically well-confirmed scientific framework. My analysis should provide fresh support for representationalism by demonstrating that diverse quantitatively precise scientific explanations become available only once we consider representational aspects of mentality. My project should also benefit cognitive science itself, by unveiling theoretical presuppositions that underlie Bayesian modeling yet that do not receive explicit mention within the scientific community.

To elaborate my representationalist position, I will analyze how exactly Bayesian models explain mental and behavioral outcomes. A key claim I will defend is that Bayesian models are explanatory because they specify how an agent's subjective probabilities *causally influence* mental and behavioral outcomes. To illustrate, suppose I infer from your facial expression that you are angry. Intuitively, my inference depends upon a prior conviction that certain emotions are likelier in light of certain facial expressions. Bayesian models codify this intuition by specifying in precise mathematical terms how my subjective probabilities influence my interpretation of your facial expression. More generally, Bayesian models explain mental and behavioral outcomes by depicting in precise mathematical terms how those outcomes causally depend upon the agent's subjective probabilities. Thus, Bayesian explanation is a kind of causal explanation. On my analysis, representational relations between mind and world crucially inform how mental states causally interact during probabilistic inference.

As a detailed case study that illustrates my approach, I will discuss Bayesian modeling of autism. Autism is a mental disorder marked by social cognition deficits and restricted, repetitive interests and behaviors. Recent work by researchers such as Jakob Hohwy and Elizabeth Pellicano indicates that Bayesian models can explain many characteristic marks of autism, including social cognition deficits. For example, individuals with autism are relatively poor at recognizing the emotions of others. A Bayesian framework can explain this deficit by positing that individuals with autism have atypical subjective probabilities involving the link between facial expressions and underlying emotions. My discussion of the Bayesian explanatory strategy should clarify some important points of similarity and difference between individuals with autism and typically developing individuals. A key similarity is that all these individuals execute the same basic mental activity: probabilistic inference. A key difference is that individuals with autism deploy subjective probabilities that diverge markedly from those employed by typically developing individuals. By analyzing how Bayesian models accommodate autism, I hope to showcase the great explanatory power of representational discourse. We can extend the representationalist paradigm beyond the case usually emphasized by philosophers --- typically developing individuals --- to illuminate the minds of atypically developing individuals. This shows that the representationalist paradigm is more flexible and robust than even many of its staunchest proponents have previously suspected.

### ***Methods and work plan***

I plan to write four papers during the fellowship period. The first paper, “Mental Representation in Bayesian Cognitive Science,” will defend the key thesis that Bayesian cognitive science makes essential appeal to representational mental states. (My previous publications have defended that thesis for two special cases: *perception* and *action*. The present paper will extend the thesis to Bayesian cognitive science more generally.) The second paper, “Representationalism versus Anti-Representationalism,” will relate the analysis from the first paper to the longstanding philosophical debate between representationalists and anti-representationalists. The third paper, “Intentional Explanation: An Interventionist Treatment,” will analyze how Bayesian cognitive science deploys subjective probabilities to provide causal explanations of mental and behavioral outcomes. This paper will draw upon recent philosophical discussions of causal explanation, especially the writings of James Woodward and Christopher Hitchcock. The final paper, “Bayesian Modeling of Autism,” will discuss autism as a case study that illustrates ideas from the first three papers. To write these four papers, I will study the relevant scientific research and subject that research to careful philosophical analysis. An NEH fellowship would afford me the time necessary to accomplish those tasks. I will eventually integrate the various elements of my project into a book, which Oxford University Press has already expressed interest in publishing.

### ***Competencies, skills, and access***

I have published numerous papers on these matters. A previous NEH fellowship (Grant FA-56081-11) yielded four papers on representational aspects of mental computation. Those four papers, along with my other work on computation and representation, led to my receiving the 2015 Herbert A. Simon Award in Computing and Philosophy from the International Association for Computing and Philosophy. I have also published multiple papers on Bayesian cognitive science, including “Bayesian Perceptual Psychology” (*The Oxford Handbook of the Philosophy of Perception*, ed. Mohan Matthen, Oxford University Press) and “Bayesian Sensorimotor Psychology” (forthcoming in *Mind and Language*). Finally, I have published a long paper on the philosophical foundations of probability theory, entitled “Some Epistemological Ramifications of the Borel-Kolmogorov Paradox” (*Synthese*). These papers demonstrate my extensive, up-to-date knowledge of cognitive psychology and Bayesian modeling.

During the fellowship period, I will build upon my previous publications to elucidate how Bayesian cognitive science illuminates representational aspects of mentality. I will remain based in Santa Barbara, where all the articles and books I need are available through the university library.

### ***Final product and dissemination***

I am trained as a philosopher, and I publish mainly in philosophy journals. However, I anticipate that my project will engage researchers outside philosophy. Since my project addresses the underpinnings of cognitive science, I hope that cognitive scientists will find my project stimulating. I am communicating with several leading scientific researchers on these topics, so my project has the potential to impact how actual science proceeds. I particularly hope that my work will interest both scientists and humanists who study autism. A major goal of the humanities is to provide greater insight into people who differ from oneself. My project promotes this goal by highlighting how a single framework --- Bayesian modeling --- encompasses both typically developing individuals *and* individuals with autism. My project should promote greater understanding of individuals with autism by clarifying how they resemble typically developing individuals in some respects but not in other respects. The results should interest researchers concerned with the minds of atypically developing individuals, including researchers in disability studies. Overall, I hope to demonstrate how we can gain insight into the minds of both typically and atypically developing individuals through an interdisciplinary methodology that connects the humanities with current cognitive science.

## Bibliography

- Baker, C., and Tenenbaum, J. 2014. "Modeling Human Plan Recognition Using Bayesian Theory of Mind." In G. Sukthankar, R. P. Goldman, C. Geib, D. Pynadath, D., and H. Bui, eds., *Plan, Activity, and Intent Recognition: Theory and Practice*. Waltham: Morgan Kaufmann.
- Bayes, T., and Price, R. 1763. "An Essay Towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F.R.S. Communicated by Mr. Price, in a Letter to John Canton, A.M.F.R.S." *Philosophical Transactions of the Royal Society of London* 53: 470-418.
- Burge, T. 2007. *Foundations of Mind*. Oxford: Oxford University Press.
- Burge, T. 2010. *Origins of Objectivity*. Oxford: Oxford University Press.
- Campbell, J. 2007. "An Interventionist Approach to Causation in Psychology." In A. Gopnik and L. Schulz, eds., *Causal Learning: Psychology, Philosophy and Computation*. Oxford: Oxford University Press.
- Chalmers, D. 2011. "Frege's Puzzle and the Objects of Credence." *Mind* 120: 587-365.
- Churchland, P. 1981. "Eliminative Materialism and the Propositional Attitudes." *Journal of Philosophy* 78: 67-90.
- Clark, A. 2013. "Expecting the World: Perception, Prediction, and the Origins of Human Knowledge." *Journal of Philosophy* 110: 469-496.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford: Clarendon Press.
- Dennett, D. 1987. *The Intentional Stance*. Cambridge: MIT Press.
- Fodor, J. 1975. *The Language of Thought*. New York: Thomas Y. Crowell.
- Fodor, J. 2008. *LOT2*. Oxford: Clarendon Press.
- Grush, R. 2004. "The Emulation Theory of Representation: Motor Control, Imagery, and Perception." *Behavioral and Brain Sciences* 27: 377-442.
- Helmholtz, H. von. 1867. *Handbuch der Physiologischen Optik*. Leipzig: Voss.
- Hohwy, J. 2014. *The Predictive Mind*. Oxford: Oxford University Press.
- Knill, D., and Richards, W., eds. 1996. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.
- Lawson, R., Rees, G., and Friston, K. 2014. "An Aberrant Precision Account of Autism." *Frontiers in Human Neuroscience* 8 (article 302): 1-10.
- Quine, W. V. 1960. *Word and Object*. Cambridge: MIT Press.
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge: MIT Press.
- Peacocke, C. 1993. "Externalist Explanation." *Proceedings of the Aristotelian Society* 67: 203-230.
- Pellicano, E., and Burr, D. 2012. "When the World Becomes Too Real." *Trends in Cognitive Science* 16: 504-510.
- Pouget, A., Beck, J., Ma, W. J., and Latham, P. 2013. "Probabilistic Brains: Knowns and Unknowns." *Nature Neuroscience* 16: 1170-1178.
- Putnam, H. 1975. *Mind, Language, and Reality: Philosophical Papers, vol. 2*. Cambridge: Cambridge University Press.
- Searle, J. 1983. *Intentionality*. Cambridge: Cambridge University Press.
- Stalnaker, R. 1999. *Context and Content*. Oxford: Oxford University Press.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science*. Cambridge: MIT Press.
- Strevens, M. 2008. *Depth*. Cambridge: Harvard University Press.
- Trommershäuser, J., Körding, K., and Landy, M., eds. 2011. *Sensory Cue Integration*. Oxford: Oxford University Press.
- Wolpert, D. 2007. "Probabilistic Models in Human Sensorimotor Control." *Human Movement Science* 26: 511-524.
- Woodward, J. 2003. *Making Things Happen*. Oxford: Oxford University Press.
- Woodward, J., and Hitchcock, C. 2003. "Explanatory Generalizations II: Plumbing Explanatory Depth." *Nous* 37: 181-199.
- Yablo, S. 2003. "Causal Relevance." *Philosophical Issues* 13: 316-327.

## Michael Rescorla - Résumé

### Employment

Associate Professor, Summer 2009 to present  
Department of Philosophy, University of California, Santa Barbara  
Assistant Professor, Fall 2003 to Spring 2009  
Department of Philosophy, University of California, Santa Barbara

### Education

Harvard University, Ph.D., Philosophy, June 2003  
Dissertation: *Is Thought Explanatorily Prior to Language?*  
Harvard University, B.A., *Summa Cum Laude*, Philosophy and Mathematics, June 1997  
Senior Thesis: *Forcing, Atoms, and Choice*

### Fellowships, Honors, and Awards

- Herbert A. Simon Award for Outstanding Research in Computing and Philosophy, awarded annually by the International Association for Computing and Philosophy (2015).
- National Endowment for the Humanities Fellowship for University Teachers, Grant FA-56081-11 (Winter 2011-Fall 2012).
- R. M. Martin Fellowship in Philosophy, Harvard University (2002-3).
- John Parker Scholarship, Harvard University (2001-2).
- Andrew W. Mellon Fellowship in Humanistic Studies (1997-8, declined).
- National Science Foundation Graduate Research Fellowship in Mathematical Sciences (Fall 1997 through Summer 1999, Fall 2000 through Summer 2001).
- Thomas T. Hoopes Prize for Senior Thesis, Harvard University (Spring 1997).
- Junior Phi Beta Kappa, Harvard University (Spring 1996).

### Publications Resulting from NEH Grant FA-56081-11

- “From Ockham to Turing --- and Back Again,” *Turing 100: Philosophical Explorations of the Legacy of Alan Turing*, in *Boston Studies in the Philosophy and History*, eds. Alisa Bokulich and Juliet Floyd. Springer (in press).
- “The Causal Relevance of Content to Computation,” *Philosophy and Phenomenological Research* 88 (2014): pp. 173-208.
- “A Theory of Computational Implementation,” *Synthese* 191 (2014): pp. 1277-1307.
- “How to Integrate Representation into Computational Modeling, and Why We Should,” *The Journal of Cognitive Science* 13 (2012): pp. 1-38.

### Partial List of Other Publications

- “Bayesian Sensorimotor Psychology,” *Mind and Language* (in press).
- “The Computational Theory of Mind,” *The Stanford Encyclopedia of Philosophy* (in press).
- “The Representational Foundations of Computation,” *Philosophia Mathematica* (published online by Oxford Journals in 2015; print version in press).
- “Some Epistemological Ramifications of the Borel-Kolmogorov Paradox,” *Synthese* 192 (2015): pp. 735-767.



- “Computational Modeling of the Mind: What Role for Mental Representation?”, *Wiley Interdisciplinary Reviews: Cognitive Science* 6 (2014): pp. 65-73.
- “The Causal Relevance of Content to Computation,” *Philosophy and Phenomenological Research* 88 (2014): pp. 173-208.
- “Perceptual Constancies and Perceptual Modes of Presentation,” *Philosophy and Phenomenological Research* 88 (2014): pp. 468-476.
- “Bayesian Perceptual Psychology,” *The Oxford Handbook of the Philosophy of Perception*, ed. Mohan Matthen. Oxford University Press (published by Oxford Handbooks Online in 2013; print version in press).
- “Against Structuralist Theories of Computational Implementation,” *The British Journal for the Philosophy of Science* 64 (2013): pp. 681-707.
- “Rationality as a Constitutive Ideal,” *A Companion to Davidson*, eds. Ernie Lepore and Kirk Ludwig. Wiley-Blackwell (2013): pp. 472-488.
- “Millikan on Honeybee Navigation and Communication,” *Millikan and Her Critics*, eds. Dan Ryder, Justine Kingsbury, and Kenneth Williford. Wiley-Blackwell (2013): pp. 87-102.
- “Are Computational Transitions Sensitive to Semantics?”, *Australasian Journal of Philosophy* 90 (2012): pp. 703-721.
- “Predication and Cartographic Representation,” *Synthese* 169 (2009): pp. 175-200.
- “Shifting the Burden of Proof?”, *The Philosophical Quarterly* 59 (2009): pp. 86-109.
- “Epistemic and Dialectical Regress,” *Australasian Journal of Philosophy* 87 (2009): pp. 43-60.
- “Assertion and its Constitutive Norms,” *Philosophy and Phenomenological Research* 79 (2009): pp. 98-130.
- “Chrysippus’s Dog as a Case Study in Non-Linguistic Cognition,” *The Philosophy of Animal Minds*, ed. Robert Lurz. Cambridge University Press (2009): pp. 52-71.
- “Cognitive Maps and the Language of Thought,” *The British Journal for the Philosophy of Science* 60 (2009): pp. 377-407.
- “Church’s Thesis and the Conceptual Analysis of Computability,” *Notre Dame Journal of Formal Logic* 48 (2007): pp. 253-280.
- “Convention,” *Stanford Encyclopedia of Philosophy*, ed. Edward Zalta (published in Fall 2007; substantive revision in Fall 2011).


### Partial List of Invited Talks

- “Levels of Computational Explanation,” *International Association for Computing and Philosophy Annual Conference*, Keynote Speech, scheduled for June 2015.
- “Is Computation Formal?”, *University of Groningen*, June 2014.
- “Bayesian Modeling of the Mind,” *Conference on Non-Propositional and Imagistic Representations*, *University of Antwerp*, June 2014.
- “Is Computation Formal?”, *Columbia University*, October 2013.
- “Modest Foundationalist Solutions to the Regress Problem,” *Workshop on Infinite Regress*, *Vanderbilt University*, October 2013.
- “The Representational Foundations of Computation,” *SoCal PhilMath + PhilLogic + FoM Workshop*, *University of California, Irvine*, February 2013.
- “Essential Meanings,” *University of Cincinnati*, *46th Annual Philosophy Conference*, May 2010.
- “Mental Syntax,” *Southern Society for Philosophy and Psychology*, *Annual Meeting*, April 2010.
- “Mental Syntax,” *University of California, Los Angeles*, February 2010.
- “Logical Form and Cartographic Representation,” *Depiction and Description Conference*, *Singapore National University*, January 2010.

May 24, 2015

To Whom It Concerns:

(b) (6)



Tyler Burge  
Distinguished Professor of Philosophy  
UCLA

25 May 2015

(b) (6)



(b) (6)



Mohan Matthen,  
Professor of Philosophy and Canada Research Chair,  
University of Toronto