

How the Cognitive Science of Belief Can Transform the Study of Mental Health

Eric Mandelbaum

The Graduate Center and Baruch College, CUNY

Nicolas Porot

Africa Institute for Research in Economics and Social Science, UM6P

NB: this is the penultimate version. We had to meet tight limits on word count and references for publication. They also didn't like a reference in here that we had to pull for, um, curious reasons. So please cite the published version which is a substantially condensed version of this paper. Published version is here: <https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2798509>

The cognitive science of belief is a burgeoning field, with insights ranging from detailing the fundamental structure of the mind,¹ to explaining the spread of fake news.² Here we highlight how new insights into belief acquisition, storage, and change can transform our understanding of psychiatric disorders. Although we focus on monothematic delusions, the morals apply more broadly.

Psychiatry & Belief

The Capgras syndrome patient thinks their loved one is an impostor yet doesn't call the cops. The Cotard's syndrome patient thinks they're dead but still eats dinner. These delusions seem like the workings of a belief system gone haywire. Why would one share the same bed as their spouse-replacing clone? How can one believe they're dead but still want seconds?

The problem is that belief is a concept in both “folk psychology” and science. Sometimes the former leads us astray, especially in clinical contexts. The new science of belief suggests that delusions function similarly to the beliefs of neurotypicals. By understanding how belief really works, we can better understand and serve psychiatric populations.

Belief Acquisition

Cognitive science uncovers an asymmetry between acquiring and rejecting beliefs.^{3,4,5} Acquisition is automatic, but rejection takes effort. These are the only options. There is no suspending belief; you can’t be neutral on a moving train.⁶ Information must be actively rejected, lest it’s passively accepted. Since the latter is easier, we do it more frequently. Because rejection requires attention, one cannot be distracted.

Before symptom onset, the Capgras patient regularly experienced the positive affect associated with seeing people they love; they saw their spouse and felt that low-level love buzz. But they now have a valenced-perception problem. Their spouse appears the same, but something is off: startlingly, seeing them doesn’t *feel* the same anymore.^{7,8,9,10,11} The patient then acts as anyone would, using their senses to guide belief, passively accepting their experience. Needing to explain why seeing their spouse no longer feels the same, they infer that their spouse is an impostor.

Patients are making understandable inferences from difficult-to-resolve premises. They acquired those premises not from a reasoning error but from a problem of affective perception combined with the formidable power of rationalization. To accept that their perceptual beliefs are misleading, they must continuously reject their perceptual experience. But the power of experience is strong;¹²

rejecting that information takes a level of cognitive effort that is hard to attain when one is under extraordinary cognitive load from unexpected hallucinations.

Belief Storage

A canonical picture of belief storage is that all beliefs are stored in a single “web of belief.”^{13,14,15} The evidential relations between beliefs constitute the web’s structure: changing one belief changes the topography of everything we believe. Knowing there’s no food in the house should make you believe the fridge is empty. Yet fridges still get fruitlessly opened.¹⁶ One explanation for this type of lapse is that there is no single web of belief. Instead, beliefs are cordoned off in smaller, discrete databases. This is the “Fragmentation of Belief.”⁴ Beliefs in one fragment are sequestered from beliefs in others, even when those beliefs are relevant.

Belief fragmentation helps explain how delusional beliefs become sequestered.¹⁷ Oftentimes the knowledge we need is stored but inaccessible. *Activated* beliefs are what dictates one’s behavior.¹⁸ One might think of Spain as both a beautiful vacation destination and as the location of one’s greatest heartbreak, yet never put both thoughts together until one lands in Barcelona. Similarly, a person might think: “Being alive doesn’t feel like it used to” and so infer that they’re dead. But at dinnertime, they still feel hungry. Knowledge that the dead don’t get hungry may appear in a belief fragment separated from their belief that they’re dead.

Belief Change

But even recognizing death's effects on metabolism won't necessarily dissolve the delusion, and the mechanisms that explain why are the same ones we use in our everyday life, those of a well-functioning *Psychological Immune System*.^{19,20} If the Cotard's patient realizes the dead don't get hungry, why don't they infer that they aren't dead? After all, their prior belief that they're alive is strong, and all the available evidence—e.g., their gurgling stomach—is confirmatory, assuming people reason like Bayesians.

But we are not strictly Bayesian updaters.²⁰ Consider the Cotard's patient's mindset. They've told people that they're dead, and in return, received blank stares and accusations of irrationality. Yet they still cling to the belief *I'm not irrational*. This is understandable. Giving up their core belief in their rationality would have existentially crushing consequences,²¹ so instead they infer that sometimes the dead's stomachs growl.

It's easy to accept one's wrong when one doesn't self-identify with a belief. An erroneous belief about Vermont's capital can be updated without much friction. But evidence that assails one's sanity cannot merely be taken up without a drastic reimagining of the self. The same principle holds for millennial cultists who predict end times, conspiracy theorists who are certain Q exists deep in the CIA, and the rest of us, who generally reason about misfortune using an unconscious, obstinate 'belief in a just world.'²²

This is the Psychological Immune System at work: we tend to think of ourselves as good, smart, sane people. Giving up core beliefs isn't like giving up a belief about Vermont's capital. Evidence that one is not good, smart, or sane won't be integrated in line with probabilistic norms, but will be twisted to preserve psychological stability.

This is the belief updating system properly functioning. If one's spouse leaves, one might infer that they are coming back, it'll just be a little longer. After a while, that belief may become unsustainable, so one rationalizes that *it's good that my spouse left me*; the idea that it's truly awful that one's spouse left is too devastating to handle.²³ Similar morals hold for those who lose jobs¹⁹ or limbs. Major blows to the psyche cannot be blithely accepted, except perhaps by the clinically depressed. They don't tend to believe they are good, smart, and rational. Lacking these core beliefs may be among the causes of both their depression and depressive realism.²⁴

The New Science of Belief and Delusions

Three key aspects of beliefs--that they're acquired automatically and rejected only with effort, stored in disconnected fragments, and updated to protect our core values--illuminate how delusions are shaped by the same belief mechanisms we all use.^{25,26} When we focus on patterns of reasoning, those suffering from monothematic delusions no longer look distinctively irrational. Instead, they may be adequately reasoning given their unfortunate and misleading evidence.^{27,28,29} Parallel morals apply to other pathological patterns of belief, including polythematic delusions in schizophrenia and self-image delusions in personality and eating disorders.

These lessons can be used in treatment to reduce delusional belief. For example, interventions that focus on what one's proud of can strengthen the Psychological Immune System, thus increasing receptivity to counterattitudinal information.³⁰ Before drawing attention to delusion-inconsistent evidence, clinicians can highlight information that bolsters the patient's sense of self, such as reflecting on their values. In this way cognitive science holds new, powerful tools for not only

reducing the negative feeling induced by counterattitudinal information, but also for addressing delusions more broadly.

References

1. Porot N, Mandelbaum E. The Science of Belief. In: Hemmer P., Musolino, J. eds. *The Cognitive Science of Beliefs*. Cambridge University Press; in press.
2. Pennycook G, Rand DG. The Psychology of Fake News. *Trends Cogn Sci*. 2021;25(5):388-402.
doi:10.1016/j.tics.2021.02.007
3. Gilbert DT. How mental systems believe. *American Psychologist*. 1991;46(2):107-119.
doi:[10.1037/0003-066X.46.2.107](https://doi.org/10.1037/0003-066X.46.2.107)
4. Mandelbaum E. Thinking is Believing. *Inquiry*. 2014;57(1):55-96.
doi:[10.1080/0020174X.2014.858417](https://doi.org/10.1080/0020174X.2014.858417)
5. Asp E, Tranel D. False Tagging Theory: Toward a Unitary Account of Prefrontal Cortex Function. Oxford University Press Accessed June 23, 2022.
<https://oxfordmedicine.com/view/10.1093/med/9780199837755.001.0001/med-9780199837755-chapter-29>
6. Zinn, H. (2018). *You can't be neutral on a moving train: A personal history of our times*. Beacon Press.
7. Anderson DN. The Delusion of Inanimate Doubles: Implications for Understanding the Capgras Phenomenon. *The British Journal of Psychiatry*. 1988;153(5):694-699.
doi:10.1192/bjp.153.5.694
8. Ellis, H. D., & Young, A. W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*, 157, 239–248.

9. Ellis HD, Young AW, Quayle AH, De Pauw KW. Reduced autonomic responses to faces in Capgras delusion. *Proc R Soc Lond B*. 1997;264(1384):1085-1092. doi:[10.1098/rspb.1997.0150](https://doi.org/10.1098/rspb.1997.0150)
10. Pacherie E. Perception, Emotions and Delusions: The Case of the Capgras Delusion. In: Bayne T, Fernández J, eds. *Delusion and Self-Deception: Affective and Motivational Influences on Belief Formation*. Psychology Press; 2008.
11. Coltheart M, Davies M. What is Capgras delusion?. *Cogn Neuropsychiatry*. 2022;27(1):69-82. doi:10.1080/13546805.2021.2011185
12. Bongiorno F. Spinozan Doxasticism About Delusions. *Pacific Philosophical Quarterly*. n/a(n/a). In press. doi:[10.1111/papq.12373](https://doi.org/10.1111/papq.12373)
13. Quine WVO. Two Dogmas of Empiricism. *Philosophical Review*. 1951;60(1):20-43.
14. Duhem PMM. *The Aim and Structure of Physical Theory*. Princeton University Press; 1954:85-87.
15. Fodor J. *Modularity of Mind*. MIT; 1983: 101-118.
16. Bendana J, Mandelbaum E. The Fragmentation of Belief. In: Borgoni C, Kindermann D, Onofri A, eds. *The Fragmented Mind*. forthcoming.
17. Davies M, Egan A. Delusion: Cognitive approaches—Bayesian inference and compartmentalization. *The Oxford handbook of philosophy and psychiatry*. Published online 2013:689-727. doi:10.1093/oxfordhb/9780199579563.001.0001
18. Quilty-Dunn J, Mandelbaum E. Against dispositionalism: belief in cognitive science. *Philosophical Studies*. 2018;175(9).
19. Gilbert D, Pinel E, Wilson T, Blumberg S, Wheatley T. Immune Neglect: A Source of Durability Bias in Affective Forecasting. *Journal of personality and social psychology*. 1998;75:617-638. doi:[10.1037/0022-3514.75.3.617](https://doi.org/10.1037/0022-3514.75.3.617)

20. Mandelbaum E. Troubles with Bayesianism: An introduction to the psychological immune system. *Mind and Language*. 2019;34(2).
21. Festinger L, Riecken H, Schachter S. *When prophecy fails: A social and psychological study of a modern group that predicted the destruction of the world*. Lulu Press, Inc; 2017.
22. Lerner M. *The Belief in a Just World: A Fundamental Delusion*. Springer US; 1980. doi:10.1007/978-1-4899-0448-5
23. McKay R, Langdon R, Coltheart M. "Sleights of mind": delusions, defences, and self-deception. *Cogn Neuropsychiatry*. 2005;10(4):305-326. doi:10.1080/13546800444000074
24. Moore MT, Fresco DM. Depressive realism: a meta-analytic review. *Clin Psychol Rev*. 2012;32(6):496-509. doi:10.1016/j.cpr.2012.05.004
25. Ritunnano R, Bortolotti L. Do delusions have and give meaning? *Phenom Cogn Sci*. Published online August 23, 2021. doi:[10.1007/s11097-021-09764-9](https://doi.org/10.1007/s11097-021-09764-9)
26. Flores C. Delusional evidence-responsiveness. *Synthese*. 2021;199(3-4):6299-6330.
27. Maher BA. Delusional thinking and perceptual disorder. *J. Individ. Psychol*. 1974 May;30(1):98-113
28. Maher BA. Anomalous experience in everyday life: Its significance for psychopathology. *The Monist*. 1999 October;82(4):547-570
29. Compare to similar conclusions Rosenhan draws about "the insane" Rosenhan DL. On being sane in insane places. *Science*. 1973;179(4070):250-258. doi:10.1126/science.179.4070.250
30. Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When Beliefs Yield to Evidence: Reducing Biased Evaluation by Affirming the Self. *Personality and Social Psychology Bulletin*, 26(9), 1151–1164. <https://doi.org/10.1177/01461672002611011>

ACKNOWLEDGMENTS: Molly Balikov, Adam Bradley, Chiara Brozzo, Michael Coleman, Carolina Flores, Stephen Gadsby, Bert Mandelbaum, Todd Moore, Rachel Osias, Marc Porot, the SVP reading group, Joshua Shepherd, Jennifer Ware.