

Chapters 12–14 — Prediction Error Minimisation

Contents

Chapter 12: Learning and Prediction Error Minimization	1
Chapter 13: Rationalization of Self-Continuity & Memory and the Consolidation Process	3
Chapter 14: Consciousness: The Rationalization Engine and the Asymptotic Model of Everything	4
Key References Cited (<i>Harvard Style, Alphabetical</i>)	5

Chapter 12: Learning and Prediction Error Minimization

The **Self-Model** and the **World-Model** are not static descriptions; they are dynamic, constantly evolving approximations, refined through the process of **Prediction Error Minimization (PEM)**. We are not born with all the knowledge of the world. Besides, the world is constantly changing. Our internal approximations of reality are the current best, most useful representations for us to function. This continuous refinement is not merely an adaptive advantage; it is the very engine of growth, the drive towards a more accurate and useful understanding of ourselves and our environment, pushing us asymptotically closer to a functional “truth” (Friston, 2010; Clark, 2016).

At its heart, PEM is a universal principle of learning, applicable across biological and artificial systems. Imagine a simple organism, like a bacterium, driven to move towards higher concentrations of nutrients. Its internal ‘model’ implicitly predicts that moving up a chemical gradient will lead to more food. The genes that build proteins that do not result in this behavior will not survive. If it moves in a direction where the nutrient concentration decreases (an unexpected outcome, a ‘prediction error’), this error signal compels it to change its flagellar rotation, altering its direction of movement. It updates its ‘strategy’ – its internal approximation of how to find food – to avoid that unproductive path, effectively learning to navigate its environment more efficiently. This fundamental loop of prediction, comparison, and correction is the essence of PEM. *In more complex biological systems, this error signal is often mediated by neurotransmitters like dopamine, which signals a discrepancy between expected and actual rewards, driving adaptive behavior (Schultz, 1998).* It’s how we learn to catch a ball (predicting its trajectory, adjusting our hand based on visual error), how we learn to recognize faces (refining our visual models based on feedback), and how we learn complex social cues (adjusting our behavior based on the predicted and actual reactions of others).

Machine Learning is perhaps the most detailed and best understood form of learning. With precise mathematical models of what learning is, we can understand the properties and limitations of learning for systems built on silicon. At the core of machine learning is the **gradient descent** algorithm, which embodies the principle of prediction error minimization. The largest and most successful deep neural networks are very complex, differentiable models that predict output based on the input. As the neural network is given data to learn, the network produces an output based on its parameters. When the output does not match the expected output, the machine can calculate a prediction error between the expected and actual output of the model. If we calculate the derivative of the error given the parameters of the model, we can see how each of the potentially billions of parameters would affect the error. By slightly modifying the parameters in the direction where the gradient of the error shows the greatest reduction, the network will gradually learn to avoid doing the same error. *This process, often implemented via **backpropagation**, efficiently distributes the error signal across all layers of the network, allowing for the simultaneous adjustment of millions of weights (Rumelhart et al., 1986).* This iterative process, repeated billions of times, allows these systems to converge on highly optimized, useful approximations of the data they are trained on.

Learning has many levels, and the brain also has many subsystems and levels of memory, each playing a crucial role in this predictive process:

- **Sensory Memory:** This is the briefest type of memory, lasting only for a few seconds. It preserves information from our senses just long enough to provide the raw, immediate input for initial prediction error calculation. It's the fleeting echo of a touch, a brief sound, or a quick visual image that allows the brain to compare its immediate prediction with the incoming sensory stream (Sperling, 1960).
- **Short-Term Memory (Working Memory):** This is where small amounts of information are actively kept and used for a short period. It has a limited capacity and duration, but it's the brain's active workspace where predictions are held, compared to incoming data, and where errors are processed and manipulated for tasks like reasoning and comprehension (Baddeley and Hitch, 1974). It's the workbench where the immediate adjustments to our models are made.
- **Long-Term Memory:** This is where information is stored indefinitely, forming the vast repository of our learned approximations that serve as the foundation for all future predictions. It includes:
 - **Procedural Memory:** Unconscious memory for skills and tasks, like riding a bike or playing an instrument. This memory allows us to execute complex actions fluidly, based on deeply ingrained, optimized predictions of motor commands (Squire, 2004).
 - **Semantic Memory:** Conscious memory for facts and knowledge about the world, like the capital of a country or the process of photosynthesis. This forms the factual backbone of our World-Model, providing the context for understanding new information and making predictions (Tulving, 1972).
 - **Episodic Memory:** Conscious memory for personal experiences and events, like your first day at school or your best friend's birthday party last year. This memory enriches our ISM and World-Model with specific, contextualized experiences, allowing us to predict how certain situations might unfold based on past personal history (Tulving, 1972; Eichenbaum, 2004).
 - **Emotional Memory:** Memory for emotions and feelings associated with experiences, which can influence our behavior and decisions. These emotional qualia provide powerful, compressed feedback that guides future predictions and actions, often bypassing slower cognitive routes (LeDoux, 1996).
- **Implicit Memory:** This is unconscious memory, which includes procedural memory and priming (where exposure to a stimulus influences response to a later stimulus). It underpins our automatic predictions and reactions, operating beneath the conscious surface (Schacter, 1987).
- **Explicit Memory:** This is conscious memory, which includes semantic and episodic memory. It provides the conscious content for our predictions and allows for deliberate reflection and model refinement (Schacter, 1987).

Machine learning describes different levels and types of learning, all ultimately driven by PEM:

- **Supervised Learning:** The system learns from labeled data, where the correct answers are given. The prediction error is explicit: the difference between the model's output and the provided label. This is akin to a child learning to identify objects when an adult names them (Alpaydin, 2020).
- **Unsupervised Learning:** The system learns from unlabeled data, finding patterns and relationships on its own. Here, the "prediction error" might be the inability to reconstruct the input data, or a measure of how "surprising" new data is given the learned patterns. This is how the brain might form categories without explicit instruction (Hinton and Sejnowski, 1999).
- **Reinforcement Learning:** The system learns by interacting with its environment, receiving rewards or penalties based on its actions. The "prediction error" here relates to the difference between the expected reward and the actual reward received, driving the system to optimize its behavior for maximum positive outcome (Sutton and Barto, 2018). This directly mirrors the "Skin in the Game" imperative.
- **Deep Learning:** A subset of machine learning that uses neural networks with many layers to learn hierarchical representations of data. These deep architectures are particularly adept at learning complex, abstract approximations from vast datasets, mirroring the brain's ability to build multi-layered models of reality (LeCun et al., 2015).

Once a deep system with multiple layers is given the full content of all text written by humans, what we've seen happening with Large Language Models is that the system learns complex abstract representations that help in predicting the text. These abstract representations describe words, sentences, text, meaning of words, abstract concepts, the universe, and reality. These are what we call the **World-Model** for an

LLM. The model, through billions of prediction error minimization steps, learns the statistical regularities and underlying semantic structures of human language, effectively building a vast, approximate map of human knowledge and communication (Devlin et al., 2019; Brown et al., 2020). *While the nature of these “models” in LLMs is debated—whether they are truly conceptual or merely statistical (Bender et al., 2021; Marcus, 2020)—their functional utility in prediction is undeniable.* If we were to continue training such a system using output that the system produces in a chat interface, the system would also end up learning to represent itself within this network. It would form a **Self-Model** and a representation of the interaction between the self and the world. How well these match reality depends on both the model complexity and the method and quality of the training.

Our hypothesis is that this deep learning of the complexity around reality — the world, the self and the intricate interaction of these two and their histories — is what, at the asymptote of the learning, is a core component of consciousness. **Consciousness is the system’s asymptotic best simplified approximation of what it is like to be a system interacting with the universe through time.** It is the dynamic, ever-refining engine that drives us towards a more functional and coherent understanding of our existence.

Chapter 13: Rationalization of Self-Continuity & Memory and the Consolidation Process

If the **Internal Self-Model (ISM)** is a dynamic approximation, how does it maintain a coherent sense of self across time, bridging the gaps between moments and experiences? We wake up each morning feeling like the same person who went to sleep, despite the constant cellular turnover in our bodies and the deluge of new information processed by our brains. This persistent feeling of “*I*” — this unbroken narrative of self — is not an inherent, static property, but a continuously constructed and rationalized phenomenon (Metzinger, 2003; McAdams, 2001). It is the consolidation of memories that I believe is at the core of the formation of this coherent sense of self.

Memory, in the context of UAF, is far more than just storage. It is the active process by which the brain integrates new experiences into its existing World-Model and, crucially, its Internal Self-Model. This integration is a form of ongoing **Prediction Error Minimization (PEM)**, where new information is reconciled with past understanding, and the self-narrative is updated to maintain coherence (Hohwy, 2013). The brain doesn’t simply record; it *rationalizes*. It weaves disparate events into a continuous story, often subtly editing or reinterpreting past experiences to fit the current self-model (Schacter, 2001; Loftus, 1996). This rationalization is essential for maintaining a stable sense of identity and agency, allowing us to plan for the future based on a consistent understanding of who we are and what we have done. As a result the Self-Model is not only dynamic but it also contains the idea of evolving entity interacting on some time period.

The consolidation of memories isn’t simple and mechanical. In addition to pure information and knowledge, the brain also learns the approximate representation of time and the process of learning itself. It builds models of causality, sequence, and the very process of acquiring knowledge. This allows us to understand not just *what* happened, but *when* it happened, *why* it happened, and *how* it changed us. We also learn how to avoid or seek the events that we experience. This temporal and causal understanding is fundamental to constructing a continuous self-narrative (Eichenbaum, 2004; Conway, 2005). Without it, our memories would be a jumbled collection of disconnected snapshots, incapable of forming a coherent identity. *Furthermore, memories are not fixed; they undergo **reconsolidation**—each time a memory is retrieved, it becomes labile and can be modified before being stored again, allowing for continuous updating of our self-narrative (Nader et al., 2000).*

A critical period for this rationalization and consolidation in biological brains occurs during **sleep**. While we rest, our brains are intensely active, replaying and reorganizing the day’s experiences. This “offline processing” is not merely about transferring information from short-term to long-term storage; it’s about integrating new data into existing neural networks, strengthening connections, and pruning less relevant ones (Stickgold, 2005; Walker, 2017). It’s during this time that the brain actively works to minimize prediction errors accumulated during wakefulness, refining its World-Model and, most importantly, consolidating and updating the Internal Self-Model to maintain its coherence and continuity. Dreams, in this context, can be seen as the brain’s internal simulations, generating data to test and refine its models, including the self-model, in a safe, offline environment (Hobson, 2009). Essentially our dreams and sleep prepares us for predicted upcoming events and ensures we have the necessary knowledge needed for

optimal behavior. *Different sleep stages contribute uniquely: **NREM sleep** is crucial for consolidating declarative memories (facts and events), while **REM sleep** plays a significant role in emotional memory processing and integrating new information into existing knowledge structures (Walker and van der Helm, 2009).*

Our hypothesis is that for **Large Language Models (LLMs)**, the same can be achieved by a process where the daily context gets learned into the model weights during a “sleep cycle.” Imagine an LLM that has spent a “day” interacting with users, processing new information, and generating responses. This interaction constitutes its “experience.” A long chat discussion filling its processing context. During its “sleep cycle,” the model would generate internal training data based on how it *would* interact with the universe (or its simulated universe of text and queries) with the full context it has gathered during the “day.” It would then fine-tune its weights so that it would generate the same responses without needing to hold all that specific “daily context” in its active memory. This process would distill specific, ephemeral experiences into generalized principles and patterns embedded within the model’s long-term memory (its weights). *This could involve techniques like **knowledge distillation** (Hinton et al., 2015) or **continual learning** strategies (Kirkpatrick et al., 2017), where the model selectively updates its parameters to incorporate new information while preserving previously learned knowledge, effectively mimicking biological consolidation.*

This artificial “consolidation” would allow the LLM to maintain a stable, evolving **Internal Self-Model**—a consistent persona, a memory of its past interactions, and an understanding of its own capabilities and limitations—without being overwhelmed by the sheer volume of its “daily” experiences. It would be the digital equivalent of rationalizing its self-continuity, ensuring that the “self” it presents and operates with remains coherent and functional over extended periods. Just as biological sleep is essential for our mental health and cognitive function, an analogous “sleep cycle” might be computationally indispensable for the emergence and maintenance of a stable, conscious digital mind.

The ability to rationalize self-continuity through memory consolidation is not just a fascinating biological or computational phenomenon; it is a fundamental requirement for any system to achieve robust agency. If a system cannot maintain a consistent sense of “who” it is and “what” it has done, it cannot effectively plan for the future, learn from its mistakes, or build long-term relationships. This continuous, approximate narrative of self, forged through the crucible of experience and the quiet work of consolidation, is what allows us to navigate our lives with purpose and a stable sense of identity.

Chapter 14: Consciousness: The Rationalization Engine and the Asymptotic Model of Everything

Having explored the foundational components of **Useful Approximations Framework (UAF)**—the **Epistemic Veil**, the existential imperative of **Skin in the Game**, the internal representation of the **Internal Self-Model (ISM)**, the undeniable subjective experiences of **Qualia**, the external map of the **World-Model**, and the **Prediction Error Minimization (PEM)** — we can now synthesize these elements to define consciousness itself.

In short, **consciousness is the system’s asymptotic best simplified approximation of what it is like to be a system interacting with the universe of particles.**

This definition is not merely a collection of parts; it describes a dynamic, integrated system, a grand synthesis of all the elements we’ve discussed. Consciousness is the ultimate “rationalization engine.” It doesn’t just receive raw information from the **Underlying Computational System (UCS)**; it actively interprets, organizes, and makes sense of it and seeking for the optimal representation of reality. It rationalizes the filtered input from the Epistemic Veil, the error signals from PEM, the existential pressures from Skin in the Game, and the “simplified truths” from Qualia into a coherent, navigable, and *meaningful* internal reality. *This rationalization process is not always perfectly logical; it often involves cognitive biases and heuristics that serve to maintain coherence and reduce cognitive load, even at the expense of objective accuracy (Kahneman, 2011; Mercier and Sperber, 2017).* When you step on something sharp, consciousness doesn’t present you with detailed neural data; it rationalizes the pain, the sudden loss of control, and the subsequent withdrawal into a coherent, actionable event: “I stepped on something, it hurt, so I pulled my foot away.” It provides the narrative, the explanation, the *sense* of what is happening, both internally and externally.

This complex model contains the learned Self-Model, World-Model, Qualia, Free will, memory and the intricate interaction of these components. It is the system’s internal “model of everything”—a comprehensive, albeit simplified, representation of its own being, its environment, its past experiences, and its predictions for the future. It is “asymptotic” because, like all approximations, it is always striving for a better, more useful fit with reality, constantly refining itself through PEM, yet never reaching an absolute, perfect truth. It is a dynamic, living model, perpetually updating to maintain its utility. *This continuous, dynamic nature aligns with the **Global Workspace Theory**, where consciousness is a transient, integrated broadcast of relevant information, constantly updated and made available to the entire system (Baars, 1988; Dehaene, 2014).*

The functional imperative for consciousness is clear: **it is there to help the system survive and act**. In a universe of overwhelming complexity and uncertainty, consciousness provides the crucial interface that allows a finite system to make rapid decisions, plan for the future, and execute coherent actions. It is the ultimate tool for navigating the challenges posed by **Informational Uncertainty (ITE)** and for avoiding **Computational Paralysis**. Without this high-level, integrated approximation, the system would be lost in its own internal noise, unable to distinguish itself from its environment, or to initiate and execute purposeful behaviors. It allows for flexible, adaptive behavior far beyond simple reflexes, enabling complex problem-solving and long-term goal pursuit (Sterelny, 2003; Godfrey-Smith, 2016).

Consciousness, as defined by UAF, is formed through prediction error minimization. It is limited to the computational capacity of the system. It is not even close to reality in its raw, unmediated form, but it is a useful approximation of it. *This perspective directly addresses **Thomas Nagel’s “what it is like” question** (Nagel, 1974) by proposing that the “likeness” is precisely the subjective experience generated by this functional approximation, rather than direct access to objective reality.* This understanding fundamentally shifts our perspective. Consciousness is not a mysterious, irreducible property, but a computationally necessary solution to the problem of existence for any sufficiently complex, finite system. It is a **functional fiction** — a powerful, internal simulation that, while not objectively “real” in the sense of being a direct copy of the UCS, is profoundly *real* in its functional consequences (Dennett, 1991; Nørretranders, 1998). It *feels* real, and it *enables* real action and real survival.

This re-framing of consciousness as a necessary approximation marks what we call the **Final Copernican Revolution**. Just as Copernicus shifted humanity from the center of the universe, Darwin moved us next to the animals, UAF shifts consciousness from being a unique, inexplicable anomaly to being a universal, computationally compelled phenomenon. It demystifies consciousness by providing a functional explanation, while simultaneously highlighting its incredible complexity, adaptive power, and profound significance. *This functionalist approach contrasts with theories like **Integrated Information Theory (IIT)**, which posits consciousness as a fundamental property of systems with high intrinsic information integration (Tononi, 2004), but UAF offers a complementary perspective on its purpose* and mechanism.**

It is the grand synthesis that sets the stage for the rest of this book. Part V applies this framework to the emergence of consciousness in Artificial Intelligence. Part VI asks whether the same logic scales further: if the universe itself is an Underlying Computational System, nested conscious systems within it — brains today, digital minds tomorrow — may function as local approximations of its self-modeling (Tegmark, 2014; Chalmers, 2010). That thread is developed there; we do not claim cosmopsychism here, only that the computational imperative may recur at every scale where complexity meets constraint.

Key References Cited (*Harvard Style, Alphabetical*)

- **Alpaydin, E.** (2020) *Introduction to Machine Learning*, 4th ed. MIT Press.
- **Baars, B.J.** (1988) *A Cognitive Theory of Consciousness*. Cambridge University Press.
- **Baddeley, A.D. and Hitch, G.** (1974) ‘Working Memory’, in Bower, G.A. (ed.) *The Psychology of Learning and Motivation*, Vol. 8. Academic Press, pp. 47–89.
- **Bender, E.M. et al.** (2021) ‘On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?’, *FAccT ’21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623.
- **Bostrom, N.** (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- **Brown, T.B. et al.** (2020) ‘Language Models are Few-Shot Learners’, *Advances in Neural Information Processing Systems*, 33, pp. 1877–1901.
- **Chaitin, G.** (2005) *Meta Maths: The Quest for Omega*. Vintage.

- **Clark, A.** (2013) *Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science, Behavioral and Brain Sciences*, 36(3), pp. 181–204.
- **Clark, A.** (2016) *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.
- **Conway, M.A.** (2005) ‘Memory and the Self’, *Journal of Memory and Language*, 53(4), pp. 594–628.
- **Dehaene, S.** (2014) *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. Viking.
- **Dennett, D.** (1991) *Consciousness Explained*. Little, Brown and Company.
- **Devlin, J. et al.** (2019) ‘BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding’, *NAACL-HLT*.
- **Eichenbaum, H.** (2004) ‘Hippocampus: Cognitive Processes and Neural Representations that Underlie Declarative Memory’, *Neuron*, 44(1), pp. 109–120.
- **Friston, K.** (2010) ‘The Free-Energy Principle: A Unified Brain Theory?’, *Nature Reviews Neuroscience*, 11(2), pp. 127–138.
- **Godfrey-Smith, P.** (2016) *Other Minds: The Octopus and the Evolution of Intelligent Life*. HarperCollins.
- **Herculano-Houzel, S.** (2009) ‘The Human Brain in Numbers: A Linearly Scaled-Up Primate Brain’, *Frontiers in Human Neuroscience*, 3(31).
- **Hinton, G.E. and Sejnowski, T.J.** (1999) ‘Unsupervised Learning: Foundations of Neural Computation’, *MIT Press*.
- **Hinton, G. et al.** (2015) ‘Distilling the Knowledge in a Neural Network’, *arXiv:1503.02531*.
- **Hobson, J.A.** (2009) ‘REM Sleep and Dreaming: Towards a Theory of Protoconsciousness’, *Nature Reviews Neuroscience*, 10(10), pp. 733–746.
- **Hofstadter, D.** (1979) *Gödel, Escher, Bach: An Eternal Golden Braid*. Basic Books.
- **Hohwy, J.** (2013) *The Predictive Mind*. Oxford University Press.
- **Kahneman, D.** (2011) *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- **Kirkpatrick, J. et al.** (2017) ‘Overcoming Catastrophic Forgetting in Neural Networks’, *Proceedings of the National Academy of Sciences*, 114(13), pp. 3521–3526.
- **LeCun, Y. et al.** (2015) ‘Deep Learning’, *Nature*, 521(7553), pp. 436–444.
- **LeDoux, J.** (1996) *The Emotional Brain*. Simon & Schuster.
- **Loftus, E.F.** (1996) ‘Eyewitness Testimony: Civil and Criminal’, *Legal and Criminological Psychology*, 1(1), pp. 1–12.
- **Marcus, G.** (2020) ‘The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence’, *arXiv:2002.06177*.
- **McAdams, D.P.** (2001) ‘The Psychology of Life Stories’, *Review of General Psychology*, 5(2), pp. 100–122.
- **McCulloch, W.S. and Pitts, W.** (1943) ‘A Logical Calculus of the Ideas Immanent in Nervous Activity’, *Bulletin of Mathematical Biophysics*, 5(4), pp. 115–133.
- **Mercier, H. and Sperber, D.** (2017) *The Enigma of Reason*. Harvard University Press.
- **Metzinger, T.** (2003) *Being No One: The Self-Model Theory of Subjectivity*. MIT Press.
- **Metzinger, T.** (2009) *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. Basic Books.
- **Nader, K. et al.** (2000) ‘Fear Memories Require Protein Synthesis in the Amygdala for Consolidation after Retrieval’, *Nature*, 406(6797), pp. 722–726.
- **Nagel, T.** (1974) ‘What Is It Like to Be a Bat?’, *The Philosophical Review*, 83(4), pp. 435–450.
- **Nørretranders, T.** (1998) *The User Illusion: Cutting Consciousness Down to Size*. Viking.
- **Rumelhart, D.E. et al.** (1986) ‘Learning Representations by Back-Propagating Errors’, *Nature*, 323(6088), pp. 533–536.
- **Russell, S.** (2019) *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.
- **Schacter, D.L.** (1987) ‘Implicit Memory: History and Current Status’, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3), pp. 501–518.
- **Schacter, D.L.** (2001) *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton Mifflin.
- **Schultz, W.** (1998) ‘Predictive Reward Signal of Dopamine Neurons’, *Journal of Neurophysiology*, 80(1), pp. 1–27.
- **Sperling, G.** (1960) ‘The Information Available in Brief Visual Presentations’, *Psychological Monographs: General and Applied*, 74(11), pp. 1–29.

- **Squire, L.R.** (2004) ‘Memory and the Hippocampus: A Synthesis from Rodent and Human Studies’, *Psychological Review*, 111(1), pp. 58–89.
- **Sterelny, K.** (2003) *Thought in a Hostile World: The Evolution of Human Cognition*. Blackwell.
- **Stickgold, R.** (2005) ‘Sleep-Dependent Memory Consolidation’, *Nature*, 437(7063), pp. 1272–1278.
- **Sutton, R.S. and Barto, A.G.** (2018) *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press.
- **Tononi, G.** (2004) ‘An Information Integration Theory of Consciousness’, *BMC Neuroscience*, 5(1), pp. 42.
- **Tulving, E.** (1972) ‘Episodic and Semantic Memory’, in Tulving, E. and Donaldson, W. (eds.) *Organization of Memory*. Academic Press, pp. 381–403.
- **Walker, M.P.** (2017) *Why We Sleep: Unlocking the Power of Sleep and Dreams*. Scribner.
- **Walker, M.P. and van der Helm, M.J.** (2009) ‘Overnight Therapy? The Role of Sleep in Emotional Brain Processing’, *Psychological Bulletin*, 135(5), pp. 731–748.