



HHS Public Access

Author manuscript

Curr Dir Psychol Sci. Author manuscript; available in PMC 2020 June 01.

Published in final edited form as:

Curr Dir Psychol Sci. 2019 June ; 28(3): 280–291. doi:10.1177/0963721419831992.

The power of predictions: An emerging paradigm for psychological research

Dr. J. Benjamin Hutchinson,

Department of Psychology, University of Oregon, Eugene, OR 97403. b.hutch@uoregon.edu

Dr. Lisa Feldman Barrett

Department of Psychology, Northeastern University, Boston, MA 02115.

l.barrett@northeastern.edu

Abstract

The last two decades of neuroscience research has produced a growing number of studies that suggest the various psychological phenomena are produced by predictive processes in the brain. When considered together, these studies form a coherent, neurobiologically-inspired research program for guiding psychological research about the mind and behavior. In this paper, we briefly consider the common assumptions and hypotheses that unify an emerging framework and discuss its ramifications, both for improving the replicability and robustness of psychological research and for innovating psychological theory by suggesting an alternative ontology of the human mind.

Keywords

Predictive coding; active inference; allostasis; memory; emotion; perception; concepts; social cognition; metabolism; energetics

How does a human mind emerge within a human brain as it navigates within an uncertain world while attempting to efficiently regulate its body within hard biological constraints? Recently, a powerful hypothesis has emerged that offers a possible answer: actions, and the mental events that accompany them, are thought to begin as top-down representations in the brain, fashioned from past experiences that are tested against the state of the world (see Table 1). This is not a new idea (Box 1), but increasingly, computationally-informed research in a variety of different psychological domains is testing this core hypothesis, largely along parallel trajectories. In this paper, we suggest their integration offers a counter-intuitive but coherent, neurobiologically-plausible scientific paradigm that has implications for guiding psychological research about the mind and behavior. This family of research approaches is deeply rooted in various forms of *predictive coding* (e.g., Ballard & Rao, 1999; Clark, 2013; Spratling, 2017), in which prediction signals, as representations constructed from past experiences, are compared with incoming sensory information to form prediction errors; prediction errors can be encoded and learned to update stored experience, which is then available for use in future predictions. Such approaches also include *the*

Bayesian brain approach (e.g., Vilares & Kording, 2011), which assumes that the brain performs (approximate) Bayesian inferences when computing predictions and prediction errors, *belief propagation* (e.g., Lochmann & Deneve, 2011), which proposes that predictions are anticipatory cause explanations for sensations that are mapped, inversely, to those sensations, and *active inference* (e.g., Friston, 2010; Friston et al., 2017), which hypothesizes that the brain's model of how sensations are caused is constrained by the need to minimize the cost of prediction error (referred to as the free energy principle). In the discussion that follows, we first consider key elements of this overarching "predictive processing" research program, discussing several examples of its utility for psychological science. We then discuss its potential to improve the robustness and replicability of psychological research, as well as its potential to offer unintuitive but powerful hypotheses that explain cognitions, emotions, perceptions and actions as emerging from a smaller set of common, computational ingredients.

Predictive Processing: Some Core Hypotheses

A variety of research findings that seem counterintuitive within psychology's traditional framework (Figure 1a) are plausibly explained by predictive processing accounts of the human mind and brain (Figure 1b), as discussed in a growing number of books and review papers (e.g., Barrett, 2009, 2017a, b; Clark, 2013, 2016; Friston et al., 2017; Howhy, 2013; Keller & Masic-Flogel, 2018; Koster-Hale & Saxe, 2013). This research program is united by a core computational hypothesis: a brain is continually running an internal model of an animal's world. The model is generative, meaning past experiences can be recombined in novel ways as they are remembered. Unlike psychology's traditional framework where perception and action are separate processes with one causing the other, the predictive processing approach suggests that perception and action are united by the brain's internal model in its effort to efficiently navigate its body in the world. Efficient navigation entails predictively controlling the autonomic nervous system, the endocrine system and other internal systems to anticipate the needs of the body so as to meet those needs in the service of upcoming motor actions (Sterling, 2012; Sterling & Laughlin, 2015). As a consequence of these preparations, the model predicts, or infers, the sensory inputs that are expected to derive from those movements, from which perceptions emerge (e.g., Clark, 2013; Keller & Masic-Flogel, 2018).

Via these dynamics, ongoing neuronal activity in the brain -- that is, the brain's predictions -- is thought to be a continuously changing filter through which sensory inputs are processed, influencing the relevance of those inputs, in effect deciding which sensory features warrant further processing. The hypothesis is not that people are deliberately remembering and deciding relevance, or that there is a specific mechanism for appraising self-relevance per se (as suggested by Ellsworth & Scherer, 2003), but that experiences and actions in the future are automatically conditioned on the past in an obligatory way that is experienced as effortless, without a sense of personal agency. Furthermore, the brain's internal model is continuously refined based on comparisons with incoming sensory information from the body and the world; that is, the brain's intrinsic neural activity can be modulated -- either confirmed or modified -- by comparison to sensory input about features in the world and in the body (Raichle, 2015). In effect, by registering deviations from its

internal model, the brain establishes whether and how to spend energy resources to act, and correspondingly, whether and how to invest those resources to learn any unanticipated sensory inputs to improve the internal model, thereby enhancing future predictions.

This emerging predictive processing research program offers innovative and important hypotheses for psychological science, many of which are scattered throughout the diverse literatures and growing number of summaries and reviews. Here, to illustrate the program's scientific utility, we focus on just two hypotheses: (1) single mental events do not arise in a vacuum but are temporally dependent on prior events, and (2) energy regulation, plus its affective consequences, are core features of all psychological phenomena, not just those that are emotional or involve fight-or-flight.

Temporal dependence of mental events.

Psychologists and neuroscientists have known for some time that a brain processes information in a temporally-dependent fashion, such that responses to incoming sensory inputs are conditional on current activity. Many modern day experiments routinely highlight that perceptual processing on a current experimental trial is dependent on an internal state that carries over from a previous trial or is otherwise informed by the past (e.g., Fischer & Whitney 2014; Kok et al., 2017; St. John-Saaltink et al. 2016; Van de Cruys et al., 2017). Even at longer timescales, perceptual judgements can be influenced within a single session by experimentally induced predictions (Chalk, Seitz, & Series, 2010) or predictions that are acquired across a lifetime of experience (Hansen et al., 2006; Huang & Sekuler, 2010; Wallisch, 2017). This dependence on memory occurs even when experimental tasks do not require participants to explicitly 'remember' past events, suggesting that an internal model, derived from past experiences across multiple timescales, implicitly but meaningfully influenced how sensory information was processed in the present.

These studies are just a handful of examples that suggest specific hypotheses for understanding how perceptions are formed. One such hypothesis is that, in certain circumstances, perceptual reports will resemble *a combination* between a perceiver's internal model (i.e., the prediction) and a stimulus, rather than just the stimulus per se. For example, a number of studies show that people are experientially blind when presented with an apparently random selection of black and white blobs (called Mooney images) that are in fact visually degraded versions of regular images. These images are then dramatically perceived as coherent images after exposure to the natural source images that was used to make them (e.g., Van de Cruys et al., 2017). Such findings suggest that perceptual experience (along with neural firing in low-level visual cortices) is sculpted by an internal model, after but not before exposure to the source images (congruent findings are observed in studies of Kanizsa illusions; see Kok et al., 2016). In another example, participants were presented with a visible banana and when asked to change its color to gray, adjusted the color to be slightly bluer than gray, suggesting a strong yellow coloring for the internal model of the banana (the sensory-driven bluish-gray and the predicted yellow would be combined to generate the subjective judgement of gray; Hansen et al., 2006; for a similar finding, see Wallisch, 2017). Similarly, studies have sought to articulate the properties of internal models by experimentally inducing various predictions and exploring their

behavioral consequences when the predicted visual input is non-existent or degraded (Chalk, Seitz, & Series, 2010; Fan, Hutchinson, & Turk-Browne, 2016). Additional work further suggests that the features of predicted visual inputs can be represented in visual cortex (Hindy, Ng, & Turk-Browne, 2016). Consistent with these findings, perception is also facilitated when features of the internal model anticipate incoming stimulus features (creating perceptual fluency), as in a recent study using continuous flash suppression; on a given trial, a photograph of a face was quicker to break through visual suppression when portraying a predicted set of facial actions than when a face portrayed an unexpected facial movements and therefore contained unanticipated features (i.e., prediction errors; Chanes et al., 2018).

The role of energy regulation in mental events.

One important development within the emerging predictive processing framework comes from several recent papers that have integrated computational hypotheses about predictive dynamics with anatomical models of information flow in the brain, involving the cerebral cortex (e.g., Barrett, 2017a; Barrett & Simmons, 2015; Chanes & Barrett, 2016; Keller & Mrsic-Flogel, 2018), the hippocampus and medial temporal lobe (e.g., Gravina & Sederberg, 2017), the cerebellum (Schlerf, Ivry & Diedrichsen, 2012), and the striatum along other subcortical regions that make up the brain's dopaminergic reward system (Schultz, 2016). There are even models that attempt to understand predictive processing within the anatomy of individual cells (Freddolino & Tavazoie, 2012), such as neurons (Koren & Deneve, 2017). One ambitious attempt offers specific hypotheses about the flow of predictions and prediction error signals within the brain using architectural models of cortical connections (Barrett & Simmons, 2015; Chanes & Barrett, 2016; Friston et al., 2017; see Figure 2). As we will see, a central implication arising from this approach is that the brain's internal model is centrally concerned with energy regulation, making energetics relevant for all mental events, not just those involving emotion or fight-and-flight responses. This hypothesis is consistent with the suggestion that predictive processing is ideal for minimizing free energy (Friston, 2010; Friston et al., 2017) or uncertainty, so that information in the brain is being represented as efficiently as possible.

The key structural hypothesis is that internal representations (prediction signals) and learning signals (prediction errors) are propagated by neurons arranged in a loose hierarchy (Barbas, 2015; Mesulam, 1998). Over thirty years of research has used anatomical features describing the arrangement and connectivity of neurons to predict information flow across the cortex (summarized in Barbas, 2015), which, when integrated into the predictive processing research program, suggests how a generative internal model, built and constrained by sensory events in the periphery, might be implemented in the brain and what consequence this implementation has for psychology. The main hypothesis is that neurons within cortical areas that are higher in this predictive hierarchy send prediction signals to neurons in regions that are lower in the hierarchy, with prediction errors flowing in the other direction (see Figure 2). An important implication of this hypothesis is that limbic cortices, such as portions of the cingulate cortex, orbitofrontal cortex, entorhinal cortex and anterior insula, along with the hippocampus, have anatomical features that place them at the top of this predictive architecture (Barbas, 2015), meaning that they are hypothesized to compute

the initial representations that propagate as inferential prediction signals throughout the brain (for a more detailed explanation, see Barrett & Simmons, 2015; Chanes & Barrett, 2016).

At the same time, this limbic ensemble, via a series of connections to the hypothalamus and throughout the brainstem, are responsible for regulating the body's global energy budget via control of the autonomic nervous system, the neuroendocrine and neuroimmune systems, and the other systems of the body's internal milieu (for a review of structural connectivity and evidence of intrinsic functional connectivity, see Kleckner et al., 2017). They are thought to regulate the body by anticipating its needs and attempting to meet those needs before they arise, a process called allostasis (Ganzel et al., 2010; Sterling, 2012), placing efficient energy regulation and metabolism at the structural and functional core of the brain's internal model that governs interactions with the outside world. Because learning new information also has an energy cost, energetics may also determine the value of prediction errors that update the model, as well (see Barrett et al., 2016). Limbic regions are still considered by many in psychology to be the most reactive parts of the brain, as home to emotions, and therefore in need of control. But in this predictive processing framework, strengthened by evidence from neuroanatomy, limbic cortices, plus the hippocampus, are the source of prediction signals, driving action and perception in an inferential way that is centrally concerned with energetics; not just during episodes of emotion, but during all mental events.

Several different lines of research are broadly consistent with the hypothesis that a brain's predictive dynamics govern efficient energy regulation, such that metabolism and energetics are core concerns in the construction of action and mental events. The principles of neural design indicate that anticipatory regulation of the body is advantageous for its energy efficiency (Sterling & Laughlin, 2015), consistent with evidence that this efficiency is a major constraint on brain evolution (Fonseca-Azevedo & Herculano-Houzel, 2012). Also consistent with the energetics hypothesis is the finding that an animal's internal state is a context for learning (i.e., the extent to which prediction errors update the brain's internal model; Yu & Dayan, 2005) and memory (Bouton, in press). More direct evidence for the energetics hypothesis comes from long-range connections between limbic cortices such as the anterior cingulate cortex (ACC) and primary sensory regions such as primary visual cortex (e.g., V1; Zhang et al., 2014), as well as evidence that the ACC sends prediction signals to V1 (Leinweber et al., 2017). Even more dramatically, these prediction signals appear to be the source of neural firing in V1 after retinal lesions and subsequent visual deprivation (Keck et al., 2013). Such evidence is consistent with other findings that a substantial fraction of activity in the visual cortex is not related to visual input (see references in Keck et al., 2013), but instead may be due to prediction signals (Liang et al., 2013), an interpretation that is consistent with the observation that the majority of synapses in primary visual cortex originate in top-down sources (e.g., discussed in Sillito & Jones, 2002).

Furthermore, both structural and functional connectivity data provide additional evidence that limbic cortices are contained in two, core intrinsic brain networks -- the default mode and salience networks -- that are implicated various psychological phenomena including

memory, perception, attention, social affiliation, pain, empathy, reward, addiction, stress, emotion, decision-making, among others, (reviewed in Kleckner et al., 2017). These findings suggest that allostasis, in the service of efficient energy regulation, is a fundamental feature of the brain's internal model as supported by these two networks. Such findings suggest a provocative hypothesis for future research: whatever other psychological functions these networks are performing during any given brain state, they are simultaneously maintaining or attempting to restore allostasis.

Practical Implications: The Robustness and Replicability of Psychological Research

Both hypotheses we discussed -- the temporal dependence of mental events and the importance of energetics in mental life -- have important implications for the robustness and replicability of psychological studies. Some of these observations have been made before, but a predictive processing framework unifies them into a common account and explains why, in an a priori sense, they should matter. The approach also suggests points for enhancing robustness and replicability when designing future experiments.

Temporal dependence of mental events.

The predictive processing framework suggests that the brain is constantly attempting to impose predictability over multiple time scales to maintain allostasis and maximize efficiency. Interestingly, this process is at odds with traditional laboratory experiments that are designed to probe and understand human minds using unpredictable sequences of events, potentially limiting the generalizability of stimulus-response experiments, particularly when moving from the laboratory to the real world. That is, traditional laboratory experiments are typically constructed as sequences of stimuli (provided by the experimenter), whose order is randomized to minimize the extent to which stimulus-response pairings are conditional on one another. This allows adjacent trials to be treated as independent, making them suitable for aggregation and traditional statistical analysis. In principle, many scientists would not defend the assumption that the mind works in similarly independent and discrete chunks in time. In practice, however, a participant's response on any given trial is not simply a function of the stimulus presented: it is some combination of the participant's internal model and that stimulus, and both should be modeled to maximize the robustness of scientific findings. Indeed, recent research has highlighted the utility of an experimental framework which moves beyond discrete individual events towards attempting to understand brain and behavior in terms of continuous, temporally-dependent processes (Huk et al., 2018).

The role of energy regulation in mental events.

Psychological science routinely makes reference to "affective stimuli" and "rewards" or "threats" as if such qualities are embedded in objects and events in the world rather than features that arise from transactions between a person's current state and those objects and events. This is more than a trivial distinction: the features of an experimental stimulus that are deemed salient by an experimenter may not be experienced thus by a participant, and other features that are deemed psychologically impotent by an experimenter may be salient to participants (or to certain participants). If the brain runs an internal model that, at its core,

is concerned with efficient energy regulation, and if the processing of stimuli are conditioned on that model, then the state of a person's energy balance, and all the factors that can influence that balance (such as amount of sleep, ingestion of caffeine, sugar, or nicotine, the degree of hydration, etc.) can influence how the brain processes information and, correspondingly, task performance. For example, people perform better on a variety of cognitive tasks when they are tested at an optimal moment in their circadian cycle versus at non-optimal times (Yoon et al., in press). If you schedule participants at a time of day that takes their circadian rhythm into account, then you will reduce what appears to be random variation which should increase replicability.

Even if any individual source of energetic influence is small, the impact on the robustness and replicability can be substantial, particularly when sample sizes are as small as they usually are in psychology studies (Benjamin et al., 2018). These factors, when not measured, can add variance in an experiment that will appear as noise in the measurements, increasing the likelihood of Type I and Type II errors. That is, only the most potent effects will replicate, compounding the impact of any Type II errors that are lurking in the data. Statistically controlling for such influences may remove meaningful variance that is better modeled and examined, again enhancing the likelihood of Type II errors. Additionally, when not controlling for such sources of variance, Type I errors may be inflated. When sample sizes are low, statistical power suffers, and any observed effect may be proportionally driven more by these factors (because as power is reduced, so does the proportion of any observed effect that is due to reproducible variance). Thus, observed effects will be less likely to replicate as a consequence (Szucs & Ioannidis, 2017).

Conceptual Implications For Psychological Theory

The predictive processing framework may be a principled approach for unifying psychological phenomena into a common explanatory framework with a shared vocabulary for theory building, offering several novel conceptual implications. One implication is that phenomena that we think of as separate processes arising from separate brain systems, such as automatic and controlled processing, may actually be better thought of as different modes of whole-brain function (prediction-based and prediction error-based).

With these observations in mind, the ubiquity of dual-process theories in psychology (e.g., Evans & Stanovich, 2013) may reflect a single underlying distinction between modes of prediction and prediction error processing in the brain (for a recent critique of dual-process theories, see Melnikoff & Bargh, 2018). For example, in the context of controlled vs. automatic processing, on a given experimental trial, a participant's brain will launch a set of prediction signals, which scientists interpret as evidence for a rapid, automatic and effortless process. If the stimulus is unexpected (such as when trials are randomized so stimuli remain unpredicted) or when a non-prepotent response is required, then participants will appear to correct the automatic process with a more deliberate, controlled and effortful process. Learning and practice effects during the course of an experiment might be understood as the consequence of a brain successfully updating its internal model, just as inhibiting a prepotent response may be adjusting the action plan by updating the model with prediction error. Participants thus appear to behave based on rapid, automatic responses (predictions)

followed by a more effortful choice that corrects it (updating with prediction error). What appears to be separate automatic and controlled processes in the mind that battle for control of behavior may actually be different modes of operation for the brain: one emphasizing a participant's internal model (i.e., prediction), and the other emphasizing correction and learning (i.e., prediction error). This is consistent with recent evidence that the human brain is constantly switching between "internal" and "external" modes of function (Honey, Newman & Shapiro, 2017).

One consequence of this two-mode perspective is that traditional experiments, designed as independent sequences of stimulus and response, effectively sever the contingencies between one moment and the next. That is, the brain's predictions will almost certainly be wrong on many trials, forcing it into a mode that favors encoding and processing prediction error (e.g. driven by the stimuli), when in the real world the dynamics may frequently favor prediction (i.e., people with neurotypical brains aren't continually in a high arousal state, flush with norepinephrine, focusing on encoding and learning prediction error). Thus, standard, randomized designs encourage oversampling of what might be an unnatural state of error processing. Instead, psychological causation might be better measured and modeled as temporally extended, probabilistic sequences of brain states (Barrett, 2009). Indeed, a relatively unexplored, but potentially rich line of research would be to explore traditional psychological phenomena in predictable versus unpredictable experimental contexts (e.g. repeating versus randomly generated sequences of trials).

In addition, a predictive processing framework also has the conceptual implication that certain psychological phenomena, such as memory, are actually ingredients in all psychological events, even those that do not appear to strictly involve memory. For example, attention has long been operationalized in terms of its influence on perceptual processing, however a growing amount of evidence suggests that attention appears to be meaningfully guided by what you've encountered in the past (e.g., Hutchinson & Turk-Browne, 2012). What's more, neural signals in brain regions historically associated with memory, such as the hippocampus, are also systematically observed during tasks of perception and attention (Aly, Ranganath, & Yonelinas, 2013; Aly & Turk-Browne, 2016). Predictions based on past experience influence relatively low-level perceptual processes and activity in primary visual cortex (den Ouden, Kok, & de Lange, 2012; O'Callaghan et al., 2017; Rao & Ballard, 1999). Here, we suggest that such processes are pervasive and similar dynamics take place across a distributed cortical hierarchy where the interplay between an internal model and feedback from the external world, as well as the interoceptive state of the body, guide learning at multiple timescales (e.g., Honey, Newman, and Schapiro, 2017).

The prediction framework also leads us to speculate that affect is part of every psychological phenomenon, even those that are not explicitly emotional. If allostasis and energetics are key features of the brain's internal model, then so too are the predicted sensory consequences of those processes, referred to as interoception (Craig, 2014). Interoception is usually experienced in a low dimensional form as the affective properties of valence and arousal (Barrett & Bliss-Moreau, 2009), suggesting that all psychological events exist in affective space. Valence and arousal might be better thought of as properties of consciousness, rather than properties of emotional episodes per se, as suggested by a number of philosophers.

Consistent with this speculation, research shows that all words have affective connotations (Osgood, May & Mirron, 1975) and even putatively “neutral” objects are experienced with subtle affective features (Lebrecht et al., 2012). This insight, if correct, calls into question all hypotheses that cast cognition and emotion in a battle for the control of behavior, or as two separate processes that interact to produce behavior, and suggests that the concept of “rationality” must be redefined as something other than the absence of affect.

When taken together, the implications of these conceptual considerations suggest several highly speculative ideas. First, many phenomena that go by different psychological names and up to now have been assumed to arise from distinct processes -- such as memory, perception, emotion, and so on -- may actually be better understood as arising from a smaller set of common, computational building blocks, with prediction-related processing at the core. Speculating even further, predictive optimization may even be implemented at the level of species-wide neural development. Recently, it has been proposed that many brains within the animal kingdom are structured to function via prediction and correction (Sterling & Laughlin, 2015). Consistent with the notion of ‘experience expectant’ processes (i.e. information storage in anticipation of particular life periods of experience rather than in response to them; Greenough, Black, & Wallace, 1987), efficiency based on temporal regularities in the environment might be even built into the evolution of the brain itself.

Conclusions

The scientific story of predictive processing is still evolving, but many believe this approach has ignited a paradigm shift in neuroscience. In this paper, we have proposed that this paradigm shift has important implications for psychological science, both in theory and in practice. The mind is a computational moment in a brain that creates a temporally continuous trajectory of neural activity, tasked with regulating a body in the world. Appreciating this perspective may improve the quality of our scientific findings and also offer opportunities for new discoveries about the nature of the human mind.

Acknowledgements

We thank Peter Kok for his comments on an earlier version of this manuscript. The paper was supported by grants to from the U.S. Army Research Institute for the Behavioral and Social Sciences (W911NF-16-1-019), the National Cancer Institute (U01 CA193632) and the National Institute of Mental Health (R01 MH113234 and R01 MH109464) to LFB. The views, opinions, and/or findings contained in this paper are those of the authors and shall not be construed as an official U.S. Department of the Army position, policy, or decision, unless so designated by other documents.

Recommended Readings

- Barrett LF (2017a). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12, 1–23. 10.1093/scan/nsw154 [PubMed: 27798257]
- Clark A (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 281–53. [PubMed: 23673028]
- den Ouden HEM, Kok P, and de Lange F (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology* 3
- Hohwy J (2013). *The Predictive Mind*. Oxford: OUP Oxford.

- Tamir DI, & Thornton M,A (2017). Modeling the predictive social mind. *Trends in Cognitive Sciences*, 10.1016/j.tics.2017.12.00
- Adams RA et al. (2013) Predictions not commands: active inference in the motor system. *Brain Struct. Funct* 218, 611–643 [PubMed: 23129312]
- Adolph KE, & Robinson SR (2015). Motor development In Lerner LLUMRM (Ed.), *Handbook of child psychology and developmental science* (Vol. 2, pp. 114–157). New York: Wiley.
- Aly M, Ranganath C, & Yonelinas AP (2013). Detecting changes in scenes: the hippocampus is critical for strength-based perception. *Neuron*, 78(6), 1127–1137. doi:10.1016/j.neuron.2013.04.018 [PubMed: 23791201]
- Aly M, & Turk-Browne NB (2016). Attention promotes episodic encoding by stabilizing hippocampal representations. *Proc Natl Acad Sci U S A*, 113(4), E420–429. doi:10.1073/pnas.1518931113 [PubMed: 26755611]
- Bar K-J, de la Cruz F, Schumann A, Koehler S, Sauer H, Critchley H, & Wagner G (2016). Functional connectivity and network analysis of midbrain and brainstem nuclei. *Neuroimage*.
- Barbas H (2015). General cortical and special prefrontal connections: principles from structure to function. *Annual review of neuroscience*, 38, 269–289.
- Barrett LF (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cogn Emot*, 23(7), 1284–1306. doi:10.1080/02699930902985894 [PubMed: 20221411]
- Barrett LF (2017a). *How emotions are made: The secret life the brain*. New York, NY: Houghton-Mifflin-Harcourt.
- Barrett LF (2017b). The theory of constructed emotion: an active inference account of interoception and categorization. *Soc Cogn Affect Neurosci*, 12(1), 1–23. doi:10.1093/scan/nsw154 [PubMed: 27798257]
- Barrett LF, & Bliss-Moreau E (2009). Affect as a Psychological Primitive. *Advances in Experimental Social Psychology*, 41, 167–218. 10.1016/S0065-2601(08)00404-8 [PubMed: 20552040]
- Barrett LF, Quigley KS, & Hamilton P (2016). An active inference theory of allostasis and interoception in depression. *Philos Trans R Soc Lond B Biol Sci*, 371(1708). doi:10.1098/rstb.2016.0011
- Barrett LF, & Satpute AB (2013). Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology*, 23(3), 361–372. doi:10.1016/j.conb.2012.12.012 [PubMed: 23352202]
- Barrett LF, & Simmons WK (2015). Interoceptive predictions in the brain. *Nat Rev Neurosci*, 16(7), 419–429. doi:10.1038/nrn3950 [PubMed: 26016744]
- Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers E-J, Berk R, ... Johnson VE (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6 10.1038/s41562-017-0189-z
- Bertolero MA, Yeo BTT, & D’Esposito M (2017). The diverse club. *Nat Commun*, 8(1), 1277. doi: 10.1038/s41467-017-01189-w [PubMed: 29097714]
- Bouton ME (in press). Extinction of instrumental (operant) learning: Interference, varieties of context, and mechanisms of contextual control. *Psychopharmacology*.
- Carbajal GV, & Malmierca MS (2018). The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends in Hearing*, 22, 2331216518784822 10.1177/2331216518784822 [PubMed: 30022729]
- Chalk M, Seitz AR, & Seriès P (2010). Rapidly learned stimulus expectations alter perception of motion. *Journal of Vision*, 10(8), 2–2. 10.1167/10.8.2
- Chanes L, & Barrett LF (2016). Redefining the Role of Limbic Areas in Cortical Processing. *Trends in Cognitive Sciences*, 20(2), 96–106. [PubMed: 26704857]
- Chanes L, Wormwood JB, Betz N, & Barrett LF (2018). Facial expression predictions as drivers of social perception. *Journal of Personality and Social Psychology*, 114, 380–396. [PubMed: 29369657]
- Clark A (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 281–253. [PubMed: 23673028]

- Clark A (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford, England, UK: Oxford University Press.
- Clark JE, Watson S, & Friston KJ (2018). What is mood? A computational perspective. *Psychol Med*, 1–8. doi:10.1017/S0033291718000430
- Conant RC, & Ross Ashby W (1970). Every good regulator of a system must be a model of that system†. *International journal of systems science*, 1(2), 89–97.
- Craig AD (2014). *How Do You Feel?: An Interoceptive Moment with Your Neurobiological Self*. Princeton, NJ: Princeton University Press.
- den Ouden HE, Kok P, & de Lange FP (2012). How prediction errors shape perception, attention, and motivation. *Front Psychol*, 3, 548. doi:10.3389/fpsyg.2012.00548 [PubMed: 23248610]
- Ellsworth PC, & Scherer KR (2003). Appraisal processes in emotion. *Handbook of affective sciences*, 572, V595.
- Evans JS, & Stanovich KE (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect Psychol Sci*, 8(3), 223–241. doi:10.1177/1745691612460685 [PubMed: 26172965]
- Fan JE, Hutchinson JB, & Turk-Browne NB (2016). When past is present: Substitutions of long-term memory for sensory evidence in perceptual judgments. *Journal of Vision*, 16(8). 10.1167/16.8.1
- Feldman H, & Friston KJ (2010). Attention, uncertainty, and free-energy. *Front Hum Neurosci*, 4, 215. doi:10.3389/fnhum.2010.00215 [PubMed: 21160551]
- Fonseca-Azevedo K, & Herculano-Houzel S (2012). Metabolic constraint imposes tradeoff between body size and number of brain neurons in human evolution. *Proc Natl Acad Sci U S A*, 109(45), 18571–18576. doi:10.1073/pnas.1206390109 [PubMed: 23090991]
- Fischer J, & Whitney D (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5), 738–743. 10.1038/nn.3689 [PubMed: 24686785]
- Freddolino PL, & Tavazoie S (2012). Beyond Homeostasis: A Predictive-Dynamic Framework for Understanding Cellular Behavior. *Annual Review of Cell and Developmental Biology*, 28(1), 363–384. 10.1146/annurev-cellbio-092910-154129
- Friston K (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127–138. [PubMed: 20068583]
- Friston K, FitzGerald T, Rigoli F, Schwartenbeck P, & Pezzulo G (2017). Active Inference: A Process Theory. *Neural Comput*, 29(1), 1–49. doi:10.1162/NECO_a_00912 [PubMed: 27870614]
- Ganzel BL, Morris PA, & Wethington E (2010). Allostasis and the human brain: Integrating models of stress from the social and life sciences. *Psychological Review*, 117(1), 134–174. 10.1037/a0017773 [PubMed: 20063966]
- Gardner MP & Fontanini A (2014) Encoding and tracking of outcome-specific expectancy in the gustatory cortex of alert rats. *J. Neurosci.* 34, 13000–13017 [PubMed: 25253848]
- Gravina MT, & Sederberg PB (2017). The neural architecture of prediction over a continuum of spatiotemporal scales. *Current Opinion in Behavioral Sciences*, 17, 194–202. 10.1016/j.cobeha.2017.09.001
- Greenough WT, Black JE, & Wallace CS (1987). Experience and brain development. *Child Development*, 58(3), 539–559. [PubMed: 3038480]
- Hansen T, Olkkonen M, Walter S, & Gegenfurtner KR (2006). Memory modulates color appearance. *Nature Neuroscience*, 9(11), 1367–1368. 10.1038/nn1794 [PubMed: 17041591]
- Hindy NC, Ng FY, & Turk-Browne NB (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nat Neurosci*, 19(5), 665–667. doi:10.1038/nn.4284 [PubMed: 27065363]
- Hohwy J (2013). *The predictive mind*: OUP Oxford.
- Honey CJ, Newman EL, & Schapiro AC (2017). Switching between internal and external modes: A multiscale learning principle. *Network Neuroscience*, 1(4), 339–356. 10.1162/NETN_a_00024 [PubMed: 30090870]
- Huang J, & Sekuler R (2010). Distortions in recall from visual memory: Two classes of attractors at work. *Journal of Vision*, 10(2), 24–24. 10.1167/10.2.24

- Huk A, Bonnen K, & He BJ (2018). Beyond Trial-Based Paradigms: Continuous Behavior, Ongoing Neural Activity, and Natural Stimuli. *Journal of Neuroscience*, 38(35), 7551–7558. 10.1523/JNEUROSCI.1920-17.2018 [PubMed: 30037835]
- Hutchinson JB, & Turk-Browne NB (2012). Memory-guided attention: control from multiple memory systems. *Trends Cogn Sci*, 16(12), 576–579. doi:10.1016/j.tics.2012.10.003 [PubMed: 23141429]
- Keck T, Keller GB, Jacobsen RI, Eysel UT, Bonhoeffer T, & Hübener M (2013). Synaptic Scaling and Homeostatic Plasticity in the Mouse Visual Cortex In Vivo. *Neuron*, 80(2), 327–334. 10.1016/j.neuron.2013.08.018 [PubMed: 24139037]
- Keller GB, & Mscic-Flogel TD (2018). Predictive Processing: A Canonical Cortical Computation. *Neuron*, 100(2), 424–435. 10.1016/j.neuron.2018.10.003 [PubMed: 30359606]
- Kok P, Bains LJ, van Mourik T, Norris DG, & de Lange FP (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Current Biology*, 26(3), 371–376. 10.1016/j.cub.2015.12.038 [PubMed: 26832438]
- Kok P, Mostert P, & Lange FP de. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences*, 201705652 10.1073/pnas.1705652114
- Koren V, & Denève S (2017). Computational Account of Spontaneous Activity as a Signature of Predictive Coding. *PLoS Comput Biol*, 13(1):e1005355. doi:10.1371/journal.pcbi.1005355 [PubMed: 28114353]
- Koster-Hale J, & Saxe R (2013). Theory of Mind: A Neural Prediction Problem. *Neuron*, 79(5), 836–848. 10.1016/j.neuron.2013.08.020 [PubMed: 24012000]
- Kleckner IR, Zhang J, Touroutoglou A, Chanes L, Xia C, Simmons WK, ... Barrett LF (2017). Evidence for a Large-Scale Brain System Supporting Allostasis and Interoception in Humans. *Nat Hum Behav*, 1. doi:10.1038/s41562-017-0069
- Kuperberg GR, & Jaeger FT (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31 (1), 32–59.
- Lebrecht S, Bar M, Barrett LF, & Tarr MJ (2012). Micro-valences: Affective valence in “neutral” everyday objects. *Frontiers in Perception Science*. 3:107. doi:10.3389/fpsyg.2012.00107 PMC3328080
- Leinweber M, Ward DR, Sobczak JM, Attinger A, & Keller GB (2017). A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions. *Neuron*, 95(6), 1420–1432.e5. 10.1016/j.neuron.2017.08.036 [PubMed: 28910624]
- Liang M, Mouraux A, Hu L, & Iannetti GD (2013). Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nature Communications*, 4, 1979 10.1038/ncomms2979
- Lindquist KA, & Barrett LF (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends in Cognitive Sciences*, 16(11), 533–540. doi:10.1016/j.tics.2012.09.005 [PubMed: 23036719]
- Lochmann T, & Deneve S (2011). Neural processing as causal inference. *Current Opinion in Neurobiology*, 21(5), 774–781. doi:10.1016/j.conb.2011.05.018 [PubMed: 21742484]
- Lupyan G, & Clark A (2015). Words and the World. *Current Directions in Psychological Science*, 24(4), 279–284. doi:10.1177/0963721415570732
- Melnikoff DE, & Bargh JA (2018). The Mythical Number Two. *Trends in Cognitive Sciences*, 22(4), 280–293. 10.1016/j.tics.2018.02.001 [PubMed: 29571664]
- Mesulam MM (1998). From sensation to cognition. *Brain*, 121, 1013–1052 [PubMed: 9648540]
- O’Callaghan C, Kveraga K, Shine JM, Adams RB Jr., & Bar M (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Conscious Cogn*, 47, 63–74. doi:10.1016/j.concog.2016.05.003 [PubMed: 27222169]
- O’Regan JK, & Noe A (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 939–1031. [PubMed: 12239892]
- Osgood CE, May WH, & Mirron MS (1975). *Cross-cultural universals of affective meanings*. Urbana: University of Illinois Press.
- Pezzulo G, Rigoli F, & Friston K (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Prog Neurobiol*, 134, 17–35. doi:10.1016/j.pneurobio.2015.09.001 [PubMed: 26365173]

- Pulvermüller F (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, 17(9), 458–470. [PubMed: 23932069]
- Raichle ME (2015). The restless brain: how intrinsic activity organizes brain function. *Phil. Trans. R. Soc. B*, 370: 20140172 10.1098/rstb.2014.0172 [PubMed: 25823869]
- Rao RP, & Ballard DH (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*, 2(1), 79–87. doi:10.1038/4580 [PubMed: 10195184]
- Schlerf J, Ivry RB, & Diedrichsen J (2012). Encoding of Sensory Prediction Errors in the Human Cerebellum. *Journal of Neuroscience*, 32(14), 4913–4922. 10.1523/JNEUROSCI.4504-11.2012 [PubMed: 22492047]
- Schultz W (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience*, 17(3), 183–195. 10.1038/nrn.2015.26 [PubMed: 26865020]
- Seth AK (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, 17(11), 565–573. doi:10.1016/j.tics.2013.09.007 [PubMed: 24126130]
- Seth AK, & Tsakiris M (2018). Being a Beast Machine: The Somatic Basis of Selfhood. *Trends in Cognitive Sciences*, 22(11), 969–981. 10.1016/j.tics.2018.08.008 [PubMed: 30224233]
- Shadmehr R, Smith MA, & Krakauer JW (2010). Error correction, sensory prediction, and adaptation in motor control. *Annu Rev Neurosci*, 33, 89–108. doi:10.1146/annurev-neuro-060909-153135 [PubMed: 20367317]
- Shin YK et al. (2010) A review of contemporary ideomotor theory. *Psychol. Bull*, 136, 943–974 [PubMed: 20822210]
- Sillito AM, & Jones HE (2002). Corticothalamic interactions in the transfer of visual information. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 357(1428), 1739–1752. 10.1098/rstb.2002.1170 [PubMed: 12626008]
- Spratling MW (2017). A review of predictive coding algorithms. *Brain and Cognition*, 112, 92–97. 10.1016/j.bandc.2015.11.003 [PubMed: 26809759]
- Spunt RP, & Adolphs R (2017). A new look at domain specificity: insights from social neuroscience. *Nat Rev Neurosci*, 18(9), 559–567. doi:10.1038/nrn.2017.76 [PubMed: 28680161]
- St. John-Saaltink ES, Kok P, Lau HC, & Lange F. P. de. (2016). Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *Journal of Neuroscience*, 36(23), 6186–6192. 10.1523/JNEUROSCI.4390-15.2016 [PubMed: 27277797]
- Sterling P (2012). Allostasis: a model of predictive regulation. *Physiol Behav*, 106(1), 5–15. [PubMed: 21684297]
- Sterling P, & Laughlin S (2015). *Principles of neural design*: MIT Press.
- Szucs D, Ioannidis JPA (2017) Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biol*, 15(3): e2000797. doi:10.1371/journal.pbio.2000797 [PubMed: 28253258]
- Tamir DI, & Thornton MA (2018). Modeling the Predictive Social Mind. *Trends Cogn Sci*, 22(3), 201–212. doi:10.1016/j.tics.2017.12.005 [PubMed: 29361382]
- Triarhou LC (2008). *The 107 Cortical Cytoarchitectonic Areas of Constantin Von Economo and Georg N. Koskinas in the Adult Human Brain: Excerpt From: "Atlas of Cytoarchitectonics of the Adult Human Cerebral Cortex"*, Authors, Von Economo C(Vienna), Koskinas GN (Athens): Karger.
- Van de Cruys S, Vanmarcke S, Van de Put I, & Wagemans J (2018). The Use of Prior Knowledge for Perceptual Inference Is Preserved in ASD. *Clinical Psychological Science*, 6(3), 382–393. 10.1177/2167702617740955
- Vilares I, & Kording K (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Ann N Y Acad Sci*, 1224, 22–39. doi:10.1111/j.1749-6632.2011.05965.x [PubMed: 21486294]
- von Economo C (2009). *Cellular structure of the human cerebral cortex*: Karger Medical and Scientific Publishers Translated and edited by Triarhou LC (Thessaloniki)
- Wallisch P (2017). Illumination assumptions account for individual differences in the perceptual interpretation of a profoundly ambiguous stimulus in the color domain: “The dress.” *Journal of Vision*, 17(4), 5–5. 10.1167/17.4.5

- Yeo BT, Krienen FM, Eickhoff SB, Yaakub SN, Fox PT, Buckner RL, ... Chee MW (2015). Functional Specialization and Flexibility in Human Association Cortex. *Cereb Cortex*, 25(10), 3654–3672. doi:10.1093/cercor/bhu217 [PubMed: 25249407]
- Yoon C, May CP, Goldstein D, & Hasher L (in press). Attention and memory: Aging, circadian arousal patterns, and cognition In Schwarz DPN (Ed.), *Cognitive Aging: A Primer* (2nd ed.): Psychology Press.
- Yu AJ, & Dayan P (2005). Uncertainty, Neuromodulation, and Attention. *Neuron*, 46(4), 681–692. 10.1016/j.neuron.2005.04.026 [PubMed: 15944135]
- Zelano C, Mohanty A, & Gottfried JA (2011). Olfactory Predictive Codes and Stimulus Templates in Piriform Cortex. *Neuron*, 72(1), 178–187. 10.1016/j.neuron.2011.08.010 [PubMed: 21982378]
- Zhang S, Xu M, Kamigaki T, Hoang Do JP, Chang W-C, Jenvay S, Miyamichi K, Luo L, and Dan Y (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. *Science*, 345, 660–665. [PubMed: 25104383]
- Craik K (1943). *The nature of explanation*. Cambridge University Press.
- Danziger K (1997). *Naming the mind: How psychology found its language*. London, England: SAGE.
- Gregory RL (1980) Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London B*, 290(1038),181–97.
- Helmholtz H von (1860/1962) *Handbuch der physiologischen optik*, Vol. 3, ed. And trans. Southall JPC. Dover (Original work published in 1860; Dover English edition in 1962).
- Johnson-Laird PN (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge MA: Harvard University Press
- Kant I (1781/1929) *Critique of pure reason*, trans. Smith N. Kemp. Macmillan (Original work published in 1781; Kemp Smith translation 1929).
- Neisser U (1967). *Cognitive psychology*. Appleton-Century-Crofts.
- Sokolov EN (1963). Higher Nervous Functions: The Orienting Reflex. *Annual Review of Physiology*, 25(1), 545–580. 10.1146/annurev.ph.25.030163.002553
- Tolman EC (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208 [PubMed: 18870876]

Box 1: Standing on the Shoulders of Giants: The Rich History of Internal Models

Since psychology's emergence as an empirical science in the mid-1800's, research has largely relied on a stimulus (S) → organism (O) → response (R) model of the mind (Danziger, 1997). Yet this model has consistently been questioned through the ages (Barrett, 2009). Kant (1781) proposed that we experience the world through a web of our own concepts, as did the 7th century Buddhist philosopher Dharmak rti and the 11th century Islamic philosopher Ibn al-Haytham. At the dawn of psychological science, Helmholtz's idea of unconscious inference (1860) and Dewey's challenge of the reflex arc (1896) criticism of the reflex arc cautioned against a SOR model of the mind. Decades later, Craik's (1943) and Johnson-Laird's (1983) internal models, and Tolman's cognitive maps (1948) all proposed, in different ways, that the internal mental representations such as beliefs and knowledge influence subsequent perception and action at least as much, if not more than, the other way around. Early studies on attention posited that the degree to which a stimulus elicited an orienting response was related to how different it was from a "nervous model" of that stimulus based on past experience (Sokolov 1963). The hypothesis that internal representations are hypotheses that play a key role in perception and action formed the basis of the cognitive revolution (e.g., Gregory, 1980; Neisser, 1967) and within social psychology, implicit attitudes, stereotyping and prejudice are predicated on the idea that information inside the head shapes experience of and action in the world. A major benefit of predictive processing-related accounts of psychological phenomena is that they are usually embedded within an anatomical and/or computational framework, allowing, for the first time, a more direct assessment of their common (or distinct) implementations and consequences.

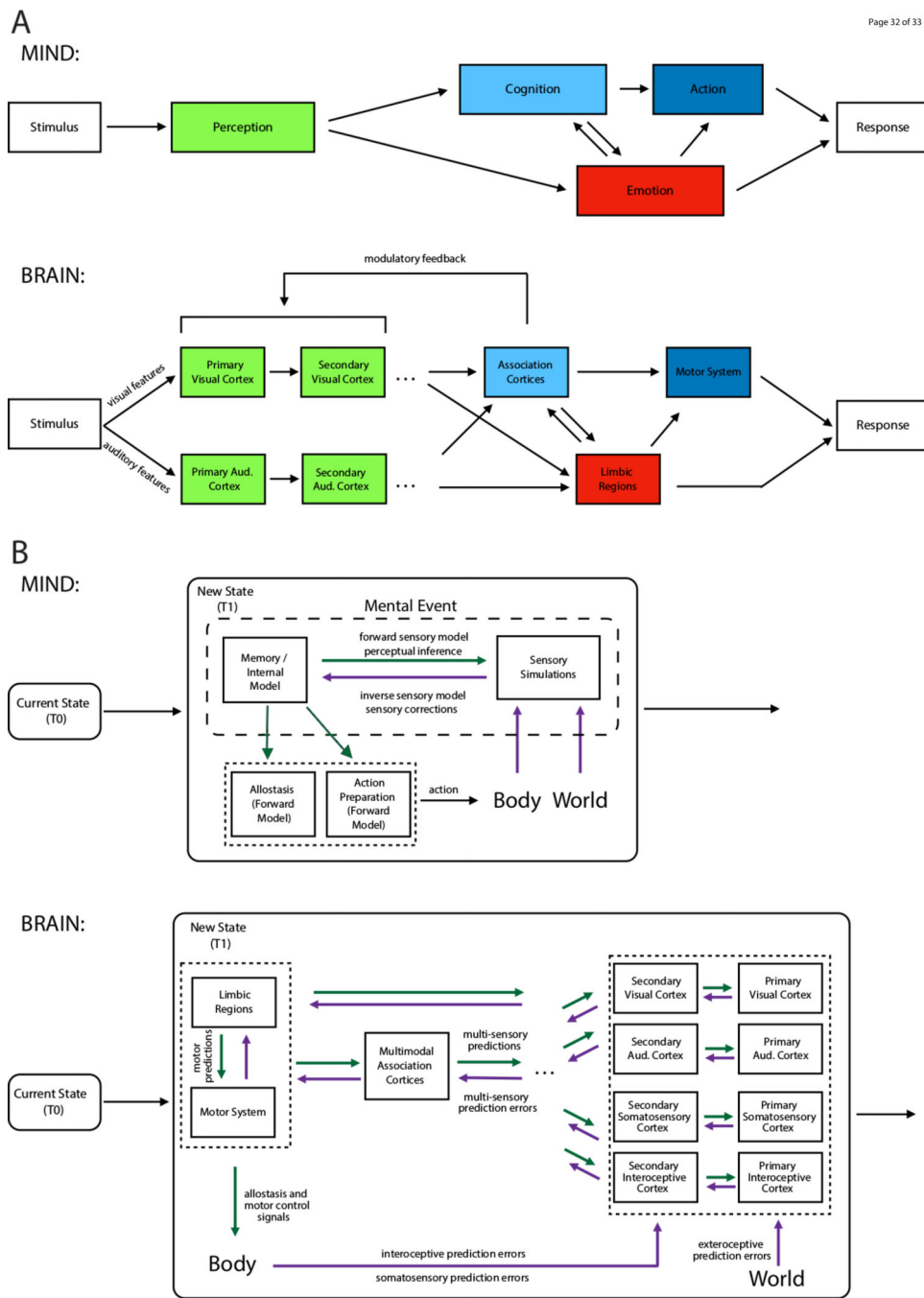


Figure 1. Psychology’s representational framework compared to the emerging predictive processing framework.

Since emerging in the mid-1800s, psychological science has been guided by what is called a representational framework (Keller & Mrcic-Flogel, 2018), where sensory neurons are thought to create a representation of features in the world (i.e., the perception of a stimulus) which is then passed, like a baton in a relay race, to other parts of the brain which appraise it (i.e., cognition and/or emotion) and then respond. In the representational framework (A), internal representations and their outputs – the mental events you experience (thoughts, feelings) and the actions you take (behaviors) – are the result of sensory inputs (i.e., stimuli)

from the world (purple arrows). Past experiences modulate these stimulus-response links (green arrows). The emerging predictive processing research paradigm (depicted in B) offers a counterintuitive correction, turning the implied causality on its head. Your experience of the world and your action in that world derive from an active, constructive process driven by your brain's internal model. Your brain starts with current conditions and "remembers" trajectories of prior experiences, projecting itself forward in time, assembling multiple competing prediction signals that prepare the body to move (green arrows; e.g., what muscles did I move the last time I was in a situation that is similar to this one?). Copies of these motor commands are thought to modulate the ongoing firing of sensory neurons, inferring the sensory consequences of those movements, thereby simulating some future state of the body and the world (green arrows; e.g., the last time I was in a similar situation and I prepared to move my body in a similar way, what did I see next?). If a brain's internal model are hypotheses about the future state of things, then incoming sensory inputs are the data used to test those hypotheses (purple arrows). The key structural hypothesis is that internal representations (prediction signals) and learning signals (prediction errors) are propagated by neurons arranged in a loose hierarchy (discussed in more detail in Figure 2). It is hypothesized that, when there is a mismatch between a prediction signal at a given level in the hierarchy and information being passed from a lower level, the neurons in question have the opportunity to change their pattern of firing to capture the unexpected input; this is how prediction error is thought to propagate up the processing hierarchy to modify the internal model in the moment and to optimize future predictions. When there is no mismatch, the prediction signal at a given level in the hierarchy is already firing in a way to represent incoming information. Once prediction errors are sufficiently minimized, these "inferences" become the brain's causal account of what caused the sensations in the first place, effectively categorizing the sensations so that they are meaningful.

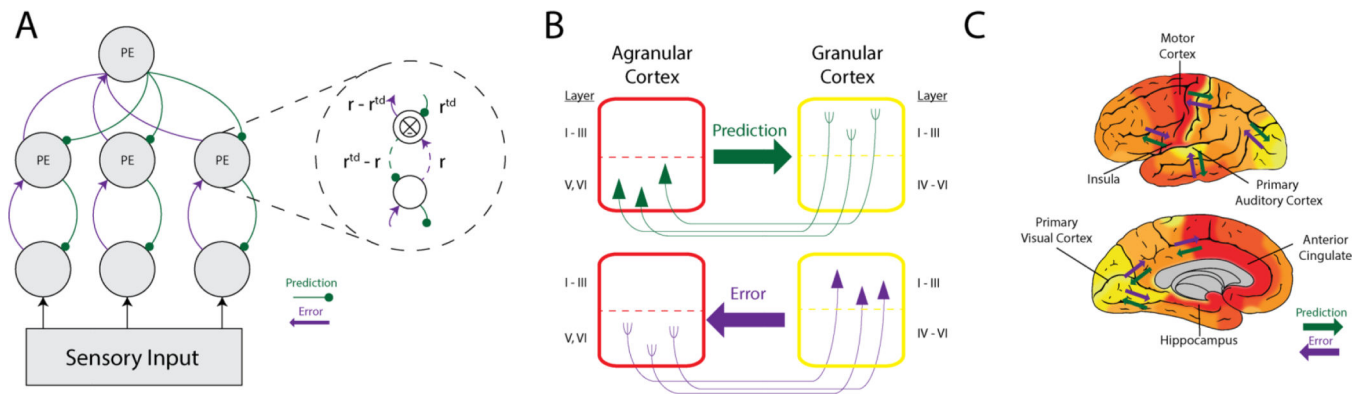


Figure 2. Predictive processing from computation to whole-brain dynamics.

(A) Schematic depiction of the original predictive coding model (Rao & Ballard, 1999).

Sensory input drives feedforward error signals which are processed by predictive estimators (PE). PE modules consist of neurons whose activity (r) is compared to top-down activity from higher levels (i.e., prediction; r^{td}) and the difference of which is propagated forward to the next level as error ($r - r^{td}$). (B) Flow of prediction and prediction error signals between cortical columns based on cortical lamination gradients. Using anterograde and retrograde tracers, Barbas and colleagues showed that the relative difference in laminar structure between two communicating cortical columns predicts whether the information flow is a feedback (prediction) or a feedforward (prediction error) signal. Prediction signals (in green) originate in the deep layers (layers V and VI) of less differentiated cortical areas (such as agranular cortex with undifferentiated layers II and III and without a layer IV, as depicted in the red column) and terminate in superficial layers of areas with a more developed laminar structure (such as dysgranular cortices with differentiated layers II and III and a rudimentary layer IV or granular cortices with differentiated layers II and III and a well-defined layer IV, depicted in the yellow column). Prediction error signals (in purple) flow in the other direction, originating in the superficial layers (II–III) with more laminar differentiation and terminating in middle deep layers (V–VI) of areas with less differentiated laminar architecture. This structural model successfully predicts the flow of information in frontal, temporal, and parietal cortices in experiments with monkeys and cats see Barbas, 2015, for a review). (C) Whole-brain estimate of flow of prediction and prediction error signals. Based on whole-brain cortical granularity data (Triarhou, 2008; von Economo, 2009). Predictions flow from cortical regions with less laminar differentiation to regions with increasing laminar differentiation (e.g., from limbic cortices to motor, interoceptive, and primary somatosensory, auditory and visual cortices). Many of the anatomical and computational details are still under investigation, such as whether each individual neuron is capable of coding for internal representations (i.e., predictions) and comparing those predictions to incoming inputs from lower in the hierarchy (i.e., prediction errors) or whether predictions and prediction errors are coded by different neurons in the cortex.

Table 1

Examples of Theory Building and Research Guided by a Predictive Processing Framework

Psychological Domain	Example Reference
Sensation and Perception: Vision	Rao & Ballard, 1999
Sensation and Perception: Audition	Carbajal & Malmierca, 2018
Sensation and Perception: Somatosensory	Adams et al, 2013
Sensation and Perception: Olfaction	Zelano et al. 2011
Sensation and Perception: Taste	Gardner & Fontanini, 2014
Sensation and Perception: Interoception	Barrett & Simmons, 2015
Memory	Hindy, Ng, & Turk-Browne, 2016
Language	Kuperberg & Jaeger, 2016
Attention	Feldman & Friston, 2010
Emotion	Barrett, 2017a
Mood	Clark, Watson, & Friston, 2018
Reward	Schultz, 2016
Social Cognition	Tamir & Thornton, 2018
Motor Action	Shadmehr, Smith, & Krakauer, 2010
Depression	Barrett, Quigley & Hamilton, 2016
The Self	Seth & Tsakiris, 2018
Words as Context	Lupyan & Clark, 2015