A Predicative Processing Model of Categorical Perception

Spencer K. Lynn[1]

[1] Charles River Analytics, Inc., Cambridge, Massachusetts, USA
slynn@cra.com

Abstract. Prior knowledge influences perception, as evidenced by categorical perception phenomena, in which expectations create psychometric distortions of perceptual space. These distortions are nonetheless associated with categorization accuracy. The paradoxical association between strength of perceptual distortion (itself an inaccurate representation of reality) and accuracy of categorization judgments suggests that understanding the computational mechanism of categorical perception could lead to advances in machine learning and artificial intelligence. Here, a framework is presented that combines signal detection theory (SDT) and predictive processing. It instantiates the SDT expected value function in a Bayesian generative hierarchy, using the function's parameters as a priori expectations about the perceptual environment. These priors then weight sensor response profiles. This approach links prediction error minimization to the optimality of perceptual judgment. The framework's a posteriori predictions for incoming sensory signals model the distortions of perceptual space associated with categorical perception. The framework provides a computational mechanism by which SDT's decision criterion is emergent from sensor tuning, rather than determined by a "decision" stage, after "perception."

Keywords: predictive processing, signal detection theory, categorical perception, response bias, decision making

Introduction

Prior knowledge influences perception. For example, in the psychological phenomenon known as categorical perception, prior knowledge induces psychological distortion of perceptual space. Perception is said to be "categorical" when discriminability of stimuli that vary on a continuum is more pronounced for two stimuli from different categories that span the continuum than for two stimuli from within a particular category–despite uniform physical differences across stimulus pairs. Prior knowledge, made accessible via words used in task instructions, creates a salient perceptual difference between stimuli even when such a difference does not in fact physically exist among the stimuli themselves. This difference in discriminability, called the between-category discrimination advantage, is associated with efficient stimulus identification (e.g., [1]): categorization accuracy is associated with large between-category discrimination advantage (see, e.g., [2] for examples in the domain of social perceptual judgments). The paradoxical association between strength of perceptual distortion (itself an inaccurate representation of reality) and accuracy of categorization judgments suggests that understanding the computational mechanism of categorical perception could lead to advances in machine learning and artificial intelligence.

That prior knowledge influences perception suggests a blurring of the traditional distinction between "perception" and "decision," which posits that decision factors, such as the risk associated with alternative decision outcomes, are applied to perception after the act. However, while evidence accumulates documenting such influence (e.g., [3]), questions remain about what specific kind of knowledge influences what people see, and how that knowledge is related to the external world and behavior. To begin addressing these questions, a framework is being developed in which perception and the early, low-level classification of sensory information are themselves treated as decisions–processes that guide behavior pursuant to implicit knowledge about the environment.

Here, a framework is presented that combines insights from behavioral ecology and neuroscience to create a generative model of categorical perception. In the model, strength of perceptual distortion is linked to optimality of category judgments via the accuracy of the agent's knowledge of environmental contingencies. Signal detection approaches from behavioral ecology (e.g., [4]), provide three elements of the framework: a conceptual model of percept categorization, described at the phenomenological level; a mathematical model that sets forth requirements for optimal categorization, situated with respect to environmental contingencies; and a specification of what prior knowledge is needed for optimal categorization. Predictive processing approaches from neuroscience (e.g., [5]), provide two elements of the framework: a structural mechanism by which relatively abstract, high-level prior knowledge can come to influence relatively specific, low-level sensory processing; and a computational mechanism to transform a sensory signal into a percept while accounting for prior knowledge.

Optimal Categorization

A central piece of the framework is a mathematical model of optimal categorization based on signal detection theory (SDT; [6]). SDT provides three parameters that describe the environment and dictate how perceivers should optimally categorize stimuli. In, for example, the social perception of anger (Fig. 1A), signals (x-axis) comprise two categories: targets, defining what anger looks like, and foils, defining what not-anger looks like. Signals from either category vary over a perceptual domain, from weak to strong cues of anger. Any signal can arise from either category, with a likelihood given by the category distributions (bell-shaped probability density functions). Overlap of the distributions creates a risk of misclassification due to perceptual similarity of targets and foils. Three parameters describe the perceptual environment: target vs. foil perceptual similarity (described by distribution means and variances), payoffs (benefits and costs accrued for four possible decision outcomes: correct detections, false alarms, missed detections, and correct rejections), and the base rate of encountering targets relative to foils. d' (sensitivity) is a common measure of ability to discriminate targets from foils. *Beta* and *c* (response bias) are common measures of risk weighting.

The combination of perceived similarity of target vs. foil categories, payoffs, and the base rate determines a "line of optimal response" (LOR, Fig. 1B, curve on inset graph) that relates a perceiver's sensitivity (*x*-axis) to the amount of bias (*y*-axis) that will maximize the perceiver's expected net benefit. In order to optimize decisions, perceivers with lower perceptual sensitivity require more extreme response bias (here, a liberal-going bias, $c<0$).

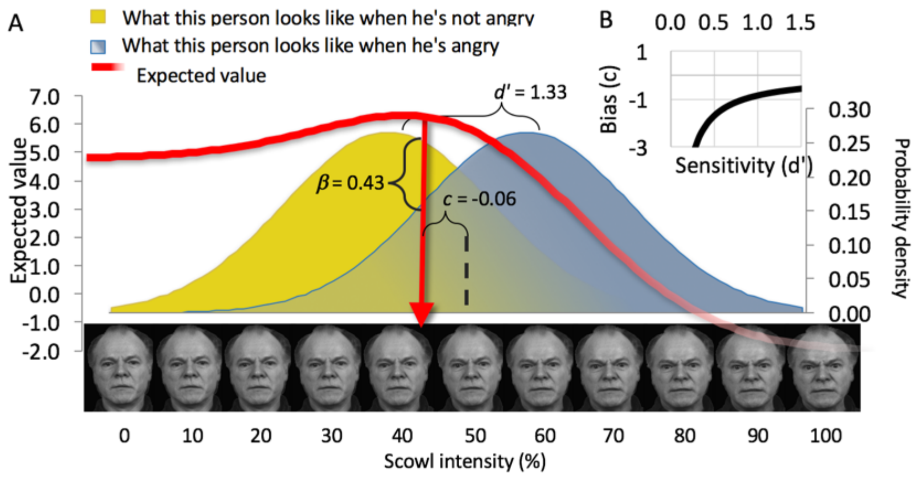

Fig. 1. Perception as a decision, after [7].

In this framework, activating category knowledge provides a set of expectations about the three parameters, which are used to generate predictions of forthcoming sensor data, influencing how a perceiver forms percepts from sensation. SDT models perceptual judgment, particularly when the same signal can be an exemplar of multiple categories and when there are costs and benefits to miscategorization; it provides well understood measures of performance; and the SDT expected value function [8], links categorization to utility of the behavior, in the world, that follows the categorization [7].

The Predictive Brain

An additional central piece of the framework is an emerging theory of how the brain works. In "predictive processing" models of cognition (see [9] for an overview), one's beliefs about the world (i.e., prior knowledge) yield predictions about incoming sensory signals (e.g., [10]). At the highest level of a neuronal processing hierarchy these predictions are abstract (e.g., concerned with physiological allostasis [11]). The predictions become increasingly specific as they descend the hierarchy to primary sensory cortex. There, they specify low-level sensory features [12]. On this account, prediction error is encoded by sensory cortex, not stimulus features, per se. The generation of predictions and processing of errors can be modeled as a hierarchy of Bayesian inference at each level of a hierarchy (Fig. 2). A prediction from level i, $\hat{x}_i$ (an estimate of the Bayesian posterior probability), becomes the prior expectation, μ_0, for the next deeper level, i+1. Level i+1 compares that prediction (now its own prior expectation) to an incoming signal, x_{i+1}, and passes the error (ε_{i+1}) back up to level i, where it is received as an incoming signal, x_i. Thus, the error of level i's prediction from time t = 0 is calculated by the next level down, i+1; the signal processed by level i at time t=0 is the error of that level's prediction from an earlier time point, t = -1; and the adjustment of level i's prior expectation in light of that error is performed by the next level up, at a later time, t = 1. A weighting term, w, weights the influence of a signal on a prediction by the variance of the signal relative to the prior expectation about it. See [13] for a deep probabilistic cognitive architecture for building more complex models.

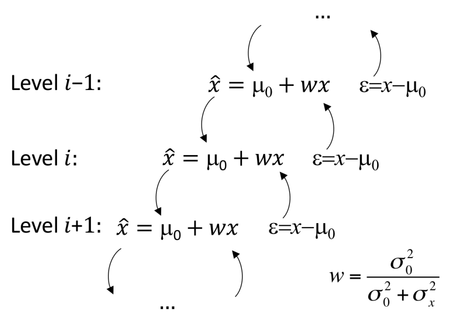

Fig. 2. Predictive processing hierarchy, after [13].

A Generative Model of Percept Formation

The SDT and predictive processing pieces of the framework combine to create a computational model of percept formation in five steps (see Fig. 3).

SDT's expected value function uses three parameters to describe the perceptual judgment task: similarity, payoffs, and base rate. In step 1, the model takes values for these parameters as prior knowledge. The parameters' values vary by context. In social perception, for example, the parameter values might differ depending on whether the man depicted in Fig. 3 is a friend or a supervisor. A perceiver with accurate estimates of these values will optimize its exposure to mistaken classifications, maximizing its net benefit over a series of judgments [7].

Fig. 3A depicts the similarity parameter as two Gaussians of a given mean and variance. Fig. 3B depicts the SDT expected value function itself for a given set of parameter values; the maximum of this function locates the optimal decision criterion (downward-pointing arrow).

The elements of the expected value function can be grouped into separate "yes" (or "go") and "no" (or "no-go") components (Fig 3C) [14]. These functions represent the expected value of categorizing a stimulus (x-axis) as a target (yes) or a foil (no), respectively.

At the framework's lowest predicative processing hierarchy level are sensors with receptive fields characterized by Gaussian probability density functions, spanning the stimulus continuum. In Fig. 3E, the sensors are represented along the bottom edge of the panel, as the source of the vertical lines. In step 2, the model uses the yes and no expected value functions to create separate yes- and no-weighted response profiles for each sensor (not shown in Fig. 3).

In step 3, for each sensor, the yes and no expected values are compared, and the profile with the greatest expected value is retained as the sensor's response profile (not shown in Fig. 3).

In step 4 (Fig. 3D), the model weights the variance of the response profile for each receptor by the difference in expected value between the yes and no expected values. This weighted variance serves to tune the response profile of a given sensor. Each sensor's contribution to responding to a given signal thereby optimizes the overall system's response to that signal (its value on the x-axis) and the system's expectations (prior knowledge) of the environmental contingencies associated with yes vs. no response to signals arising from each of the two categories.

In step 5, when an incoming signal, x, is received, it is processed across the array of weighted sensor response profiles, in a Kalman fusion-type manner. The Kalman process generates a predicted signal, or percept value, $\hat{x}$ (Fig. 3E), representing a psychological distortion of the objective signal value, x.

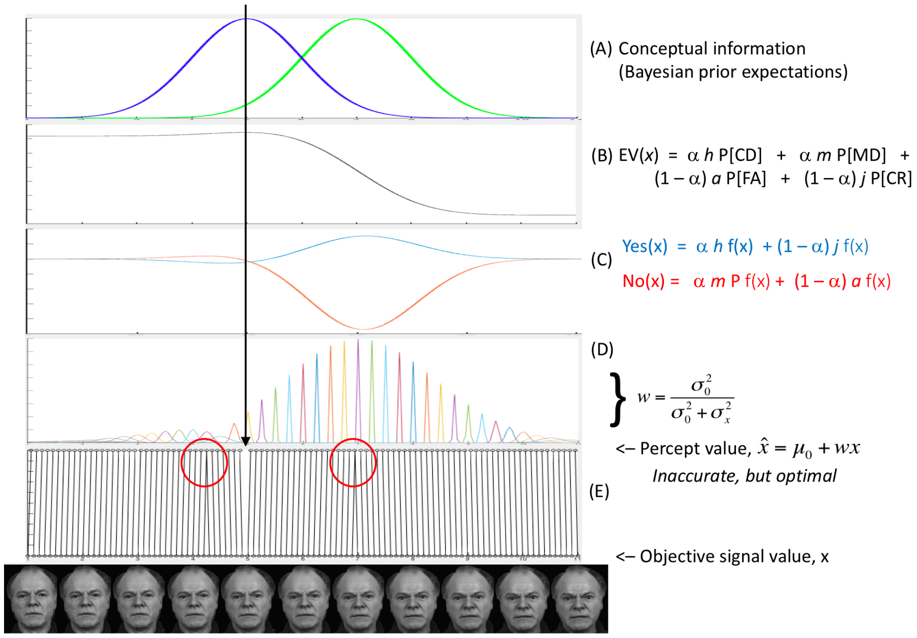

Fig. 3. Elements of the framework combining signal detection theory and predictive processing, and the resultant model of signal reception and percept formation. The vertical arrow demarcates signal detection theory's optimal criterion location across panels A-C, which aligns with the maximal generative perceptual distortion at panel E.

This distortion exhibits the hallmarks categorical perception: acquired equivalence of within category stimuli (Fig. 3E, red circles) and acquired distinctiveness of between category stimuli (Fig. 3E, arrow) [1]. Moreover, the maximal distortion of percepts occurs at the traditional decision criterion location (maximum of the SDT expected value function (Fig. 3B). The model thus provides a computational mechanism that generates SDT's optimal decision criterion from putatively Bayesian first principals of perceptual discrimination and the modulation of perceptual processes by conceptual knowledge.

Discussion

The framework presented here uses SDT as a conceptual model of perception, and instantiates the SDT expected value function in a predictive processing hierarchy. The model's a posteriori predictions for incoming sensory signals resemble the distortions of perceptual space associated with categorical perception. In this model, SDT's decision criterion, which is typically measured as a behavioral average over many trials, emerges from sensor receptive field tuning. Thus, response bias is encoded near the sensors, not applied to percepts in a separate "decision" step or module.

This model uses conventional SDT to describe perceptual judgment at a functional level and predictive processing to inspire the computational mechanism for perceptual judgment. Nonetheless, more process-oriented approaches to perceptual judgment might substitute for traditional SDT, bringing their own strengths (e.g., general recognition theory [15] or dynamic SDT [16]). Likewise, additional algorithms could potentially be applied to SDT's probabilities and payoffs to create common cognitive biases (e.g., [17-19]).

The model makes predictions for behavioral studies. In the model, the strength of between-category discrimination advantage (BCA, the amount of distortion at the downward arrow in Fig. 3) is a function of the accuracy of the prior expectations about the three SDT parameters. In simulation, less accurate parameters (relative to the simulation's ground-truth) result in weaker BCA. In studies, when environmental contingencies dictate that some amount of response bias is optimal, then response bias should be encoded in the distortion of perceptual space such that optimality of bias should correlate with reduced prediction error and stronger distortion of perceptual space.

Several elements of the model suggest possible avenues for advances in machine learning and AI. It has already been recognized that the predictive processing architecture is efficient once priors are established; only errors are transmitted, not increasing complex models of the world. The model presented here additionally suggests that categorical perception could reflect a novel tactic for management of signal uncertainty close to the sensor, by weighting the sensor response profiles as information ascends the hierarchy rather than as a separate, monolithic "decision" stage following "perception." One can speculate that, at scale (e.g., for dense multimodal sensors supporting rich conceptual/semantic systems), the distortions associated with categorical perception may yield more efficient and/or optimal classification than veridical perception followed by a subsequent decision stage.

References

1 Goldstone, R.L.: 'Influences of categorization on perceptual discrimination', Journal of Experimental Psychology: General, 1994, 123, (2), pp. 178-200

2 Fugate, J.M.B.: 'Categorical Perception for Emotional Faces', Emotion Review, 2013, 5, (1), pp. 84-89

3 Gendron, M., Lindquist, K., Barsalou, L., and Barrett, L.F.: 'Emotion words shape emotion percepts', Emotion, 2012, 12, (2), pp. 314-325

4 Lynn, S.K., Cnaani, J., and Papaj, D.R.: 'Peak shift discrimination learning as a mechanism of signal evolution', Evolution, 2005, 59, (6), pp. 1300-1305

5 Denève, S., and Jardri, R.: 'Circular inference: mistaken belief, misplaced trust', Current Opinion in Behavioral Sciences, 2016, 11, pp. 40-48

6 Green, D.M., and Swets, J.A.: 'Signal Detection Theory and Psychophysics' (Wiley, 1966. 1966)

7 Lynn, S.K., and Barrett, L.F.: '"Utilizing" signal detection theory', Psychological Science, 2014, 25, (9), pp. 1663-1673

8 Swets, J.A., Tanner, W.P., Jr., and Birdsall, T.G.: 'Decision processes in perception', Psychological Review, 1961, 68, (5), pp. 301-340

9 Allen, M., and Friston, K.J.: 'From cognitivism to autopoiesis: towards a computational framework for the embodied mind', Synthese, 2016

10 Hohwy, J.: 'The predictive mind' (Oxford University Press, 2013. 2013)

11 Chanes, L., and Barrett, L.F.: 'Redefining the Role of Limbic Areas in Cortical Processing', Trends in Cognitive Sciences, 2015

12 Rao, R.P.N., and Ballard, D.H.: 'Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects', Nature Neuroscience, 1999, 2, pp. 79

13 Pfeffer, A., and Lynn, S.K.: 'Scruff: A Deep Probabilistic Cognitive Architecture for Predictive Processing', Proceedings of 2018 International Conference on Biologically Inspired Cognitive Architectures, Forthcoming

14 Lynn, S.K.: 'Cognition and evolution: Learning and the evolution of sex traits', Current Biology, 2006, 16, (11), pp. R421-R423

15 Ashby, F.G.: 'A stochastic version of general recognition theory', Journal of Mathematical Psychology, 2000, 44, (2), pp. 310-329

16 Pleskac, T.J., and Busemeyer, J.R.: 'Two-stage dynamic signal detection: A theory of choice, decision time, and confidence', Psychological Review, 2010, 117, (3), pp. 864-901

17 Baucells, M., and Heukamp, F.H.: 'Reevaluation of the results of Levy and Levy (2002a)', Organizational Behavior and Human Decision Processes, 2004, 94, (1), pp. 15-21

18 Zhang, H., and Maloney, L.T.: 'Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action and cognition', Frontiers in Neuroscience, 2012, 6, pp. Article 1. Retrieved from journal.frontiersin.org/article/10.3389/fnins.2012.00001/

19 Lynn, S.K., Wormwood, J.B., Barrett, L.F., and Quigley, K.S.: 'Decision making from economic and signal detection perspectives: Development of an integrated framework', Frontiers in Psychology, 2015, pp. Article 952. Retrieved from journal.frontiersin.org/article/910.3389/fpsyg.2015.00952/