This article is part of the topic "Computational Approaches to Social Cognition," Samuel Gershman and Fiery Cushman (Topic Editors). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

The Flatland Fallacy: Moving Beyond Low–Dimensional Thinking

Eshin Jolly, Luke J. Chang

Computational Social Affective Neuroscience Laboratory, Department of Psychological and Brain Sciences, Dartmouth College

Abstract

Psychology is a complicated science. It has no general axioms or mathematical proofs, is rarely directly observable, and is the only discipline in which the subject matter (i.e., human psychological phenomena) is also the tool of investigation. Like the Flatlanders in Edwin Abbot's famous short story (1884), we may be led to believe that the parsimony offered by our *low-dimensional* theories reflects the reality of a much *higher-dimensional* problem. Here we contend that this "Flatland fallacy" leads us to seek out simplified explanations of complex phenomena, limiting our capacity as scientists to build and communicate useful models of human psychology. We suggest that this fallacy can be overcome through (a) the use of quantitative models, which force researchers to formalize their theories to overcome this fallacy, and (b) improved quantitative training, which can build new norms for conducting psychological research.

Keywords: Dual-processing; Computational; Social; Decision-making; Psychological education

Correspondence should be sent to Eshin Jolly, Computational Social Affective Neuroscience Laboratory, Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755. E-mail: eshin.jolly.gr@dartmouth.edu

Yet I exist in the hope that these memoirs, in some manner, I know not how, may find their way to the minds of humanity in Some Dimension, and may stir up a race of rebels who shall refuse to be confined to limited Dimensionality.

—Edwin A. Abbott, *Flatland: A Romance of Many Dimensions* (1884)

1. Introduction

Few works consider the nature of perception and dimensionality as elegantly as Edwin Abbott's (1884) novella *Flatland: A Romance of Many Dimensions*. The narrator of the story, A. Square, lives in a world full of "Flatlanders," who are incapable of perceiving or even conceiving of a reality that exists beyond two dimensions. However, after a visit from a "Stranger" (a sphere) A. Square comes to appreciate how complex and high dimensional the world really is. Ultimately, he is imprisoned for his heretical beliefs after trying to teach his colleagues about his revelations. Abbott's key insight was that creatures with limited perceptual capacities (i.e., seeing in only two dimensions) come to *reason* in a limited way, ignoring the complexity of the world and truly believing their perceptions to be veridical (Fig. 1). Much like Flatlanders, humans exhibit strong biases in their reasoning about a complex and high-dimensional world due to finite limitations on their cognitive capacities. For this reason, we posit that psychological researchers,
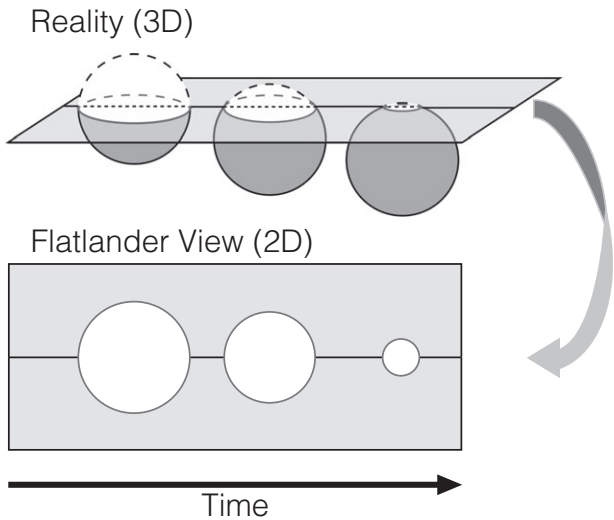

Fig. 1. In Abbott's novella, A. Square cannot perceive his world as anything other than two dimensional. From his limited perspective ("Flatlander View"; bottom), a three-dimensional entity (sphere) appears to be changing sizes before him (growing and shrinking circle). In reality (top), this entity is simply moving through a lower-dimensional plane, but A. Square's limited perspective leads to a false conclusion about the nature of reality. For similar reasons, psychological scientists may falsely conclude that the number of dimensions that accurately characterize psychological phenomena is sufficiently small, viewing the world like Flatlanders, even if in reality the complexity of psychological phenomena is high dimensional.

constrained by these limitations, may come to view the complexities of psychological life in similarly limited ways.

Human psychology is rife with complexity, the product of an immensely high-dimensional space characterized by interactions between trillions of neural connections, billions of unique individuals, and dynamic changing contexts spanning thousands of years of history. Despite this complexity, the majority of theoretical developments in psychological research have consistently converged on producing a number of low (typically two) dimensional/factor theories/process models of human mental life. Since at least the early work of Plato (Evans & Frankish, 2009), this manner of characterizing the mind has come to dominate our understanding of *emotion* (Damasio, 1994; Davidson, 1993; Ochsner, Bunge, Gross, & Gabrieli, 2002; Posner, Russell, & Peterson, 2005; Russell, 1980; Schachter & Singer, 1962; Zajonc, 1980), *social cognition* (Chaiken & Trope, 1999; Cuddy, Fiske, & Glick, 2008; Gray, Gray, & Wegner, 2007; Gray & Wegner, 2009; Greenwald & Banaji, 1995; Haslam, 2006; Mitchell, 2005; Saxe, 2005; Todorov, Said, Engell, & Oosterhof, 2008; Waytz & Mitchell, 2011), *moral judgment* (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Rand, Greene, & Nowak, 2012), *learning* (Daw, Niv, & Dayan, 2005; Frank, Cohen, & Sanfey, 2009; Poldrack et al., 2001; Poldrack & Packard, 2003), c*ognitive control* (Heatherton & Wagner, 2011; Hofmann, Friese, & Strack, 2009; McClure, Laibson, Loewenstein, & Cohen, 2004; Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1984), *decision-making* (Chang & Sanfey, 2008; Dijksterhuis, Bos, Nordgren, & van Baaren, 2006; Kahneman, 2003; Sanfey & Chang, 2008; Sloman, 1996; Wilson & Schooler, 1991), and *reasoning* (Epstein, 1994; Evans, 2003; Stanovich & West, 2000). How could something as complex as the human mind be consistently described in two dimensions, irrespective of the mental faculty under consideration? Although these theories have provided a bedrock for empirical investigation, we argue that rather than reflecting a rich characterization of the complexity of human psychology, they instead reflect a simplistic view of our scientific understanding (Flatland fallacy)—a product of the limits of our cognition.

In this paper, we outline several reasons why we believe psychologists consistently converge on two-factor solutions to characterize our understanding of human psychology. We argue that these conclusions arise from our limited cognitive capacities, social norms ubiquitous in the field of psychology, and our reliance on low-bandwidth channels to communicate research findings (e.g., natural language and simple visualizations). We suggest that moving beyond low-dimensional thinking requires formalizing psychological theories as *quantitative* computational models capable of making precise predictions about cognition and/or behavior, and we advocate for improving training in technical skills and quantitative reasoning in psychology.

2. Why does the Flatland fallacy happen?

Understanding why the Flatland fallacy occurs requires examining both biases and limitations in human cognition as well as cultural norms in research and training in psychology. Specifically, we propose four main reasons why this fallacy occurs: (a) biases in the feeling

of understanding; (b) limitations of human cognition; (c) over-reliance on traditional experimental design and analytic approaches; and (d) limitations in our ability to communicate complex concepts. Because psychology researchers have the unique privilege of being members of both the matter of study and those conducting the study, it is critical that the products of our science not be constrained by the limits of our own psychology (Meehl, 1954).

2.1. Feelings of understanding

Much like Abbott's A. Square feels that a two-dimensional existence is a complete account of his universe, humans are prone to a "folk understanding bias"—the sensation that simplistic explanations lead us to believe we truly understand more complex phenomena. Prior work in cognitive science and philosophy has illustrated how individuals can fall victim to cognitive biases that lead them to believe their actual understanding of phenomena exceeds their true understanding of phenomena. For example, individuals often report a high feeling-of-knowing despite their inability to accurately recall previously learned information (Koriat, 1993). When individuals are tested on their ability to explain how a system works (e.g., a quartz watch), they tend to report an overestimate of their knowledge until they are asked to provide a specific explication (Rozenblit & Keil, 2002). In other words, individuals create mental placeholders of elaborate, in-depth explanations (e.g., essences and hidden mechanisms) that give rise to a feeling of certainty and understanding, even when limited understanding exists (Medin, 1989; Strevens, 2000). Because these approximations can provide basic explanations as to how a system works, they are initially insightful, leaving people with the sensation that they know more than they really do (Rozenblit & Keil, 2002).

A critical observation is that this bias is exacerbated when individuals are asked to explain systems that are highly opaque, that is, have poor observability of their inner workings (Rozenblit & Keil, 2002). Unfortunately, the complexity of psychological science lies almost entirely in its lack of transparency; mental processes are not directly observable, only inferrable through observations of behavior and their correlations with biological functioning.

This bias likely originates from our strong motivation to understand and find meaning in our experiences and the world as a whole (Cohen, Stotland, & Wolfe, 1955). For this reason, it makes sense to favor simplistic and easily understandable theoretical conclusions over complex and complete accounts of phenomena, even if they are only weakly supported by experimental data. Simple explanations provide some uncertainty resolution even if they paint an incomplete picture (Pinker, 1999; Webster & Kruglanski, 1994). Consequently, researchers may be collectively at risk for pursuing a psychological science that they can "understand," irrespective of whether that science offers robust predictive accuracy.

2.2. Limitations of cognitive capacity

Limitations in the cognitive capacities and motivations of individuals offer another explanation for the Flatland fallacy. It is well established, for example, that humans are

not supercomputers who always calculate mathematically optimal solutions for the problems they face (Gigerenzer & Goldstein, 1996). Rather, our brains are the product of specific evolutionary constraints such as physical *size*—they need to be small enough to permit live births and allow us to locomote; *speed*—they need to support processing that can occur on finite time scales; and *energy*—their energy demands cannot exceed our metabolic abilities (Montague, 2007). For this reason, the notion of bounded rationality has been instrumental in characterizing how the mind processes information quickly and reasonably accurately (Simon, 1957). There are many well-known examples, such as our limited capacity to simultaneously manipulate large chunks of information (Miller, 1956), process multiple attentional tasks (Pashler, 1994), and uniquely represent person-identity information without relying on feature similarity such as stereotypes (Mervis & Rosch, 1981; Smith & Zarate, 1992).

We believe that the key psychological limitation underlying the Flatland fallacy is our inability to reason in more than a few dimensions, particularly in contexts that require integrating multiple sources of information together. Individuals tend to default to simpler, general, heuristic-like strategies that serve to make such reasoning more cognitively tractable. These strategies often constitute lower-dimensional approximations (e.g., two or three) of far more complex information landscapes, which raises the possibility that even the process of conducting scientific research can be similarly marred by the limits of our cognition. We outline three examples of how lower dimensional approximations impact how we make judgments and decisions.

2.2.1. Judgment

Many real-world settings involve situations in which individuals are faced with the task of making judgments by combining a large number of potentially relevant factors (e.g., clinical evaluations). A large body of work has consistently demonstrated that humans make judgments using just a handful of dimensions (Brunswick, 1952) rather than considering all the available information on hand. In particular, individuals over weight or under weight the relative importance of specific factors or simply ignore seemingly irrelevant information in favor of simplified evaluation criteria, such as when estimating school admissions, personality metrics, and even criminal evaluations (Dudycha & Naylor, 1966; Karelaia & Hogarth, 2008; Meehl, 1954). In other words, individuals rely on heuristics to simplify the space of information under consideration, especially when this space is very large or shares a nonlinear relationship with an outcome (Deane, Hammond, & Summers, 1972; Karelaia & Hogarth, 2008). Given the robustness of this evidence, there has been a strong call to incorporate statistical models that can integrate more dimensions in place of solely relying on clinical judgment to overcome these cognitive limitations (Dawes, 1971; Dawes, Faust, & Meehl, 1989; Meehl, 1954).

2.2.2. Decision-making

Dual-process theories have a rich history of characterizing human decision-making (Sanfey & Chang, 2008). These accounts have been used to explain how we can switch between fast, intuitive, and reflexive modes of thinking to slow deliberative calculations

(Kahneman, 2003, 2011) and also how emotions and cognitive deliberation might be integrated when making decisions (Chang & Sanfey, 2008; Chang, Smith, Dufwenberg, & Sanfey, 2011; Greene, Nystrom, Engell, Darley, & Cohen, 2004). In addition, there appears to be an upper limit on the number of attributes that can be simultaneously considered when making a decision between different choices (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Payne, 1976). Faced with a large number of factors to consider, individuals appear to act in an adaptive manner, falling back on heuristic shortcuts rather than considering all of the available information on hand (Payne, Bettman, & Johnson, 1993; Simon, 1987). In other words, humans tend to make high-dimensional problems more cognitively tractable by considering lower dimensional perspectives—specifically, through the use of heuristic strategies that entail ignoring potentially relevant factors (Gigerenzer & Brighton, 2009).

2.2.3. Conditional reasoning

Even in more socially interactive contexts that require individuals to consider the motivations of *others*, individuals exhibit consistent limits on their cognitive abilities. A large body of work in game theory and behavioral economics has demonstrated that individuals are limited in their depth of strategic reasoning (Camerer, Ho, & Chong, 2015). In this work, individuals compete or coordinate with each other within an economic game. Players' strategies in these games allow for optimizing their own payoffs while also considering the strategy utilized by other players. Interestingly, individuals are rarely able to reason more than two steps ahead of other individuals, (i.e., more than two levels of such conditional reasoning: "I think that you think that I think") (Camerer, Ho, & Chong, 2004; Griessinger & Coricelli, 2015; Stahl & Wilson, 1995). Even in non-strategic contexts, individuals have been shown to exhibit limits on the amount of recursive reasoning they are capable of, such as during theory-of-mind tasks that require inferring the motives of fictionalized agents (e.g., Happé, 1994) or, more generally, parsing language comprised of numerous embedded clauses (Karlsson, 2010). Across a large number of spoken languages, for example, this type of syntactic recursion (e.g., "a car the man the woman the boy saw drove fast") rarely exceeds two levels of depth and even in written text rarely exceeds three levels of depth (Karlsson, 2007). Because understanding conditional complexity quickly becomes incredibly difficult, individuals fall back on using simple heuristic strategies and generalized decisions rules in lieu of making more optimally rational choices (Camerer, Johnson, Rymon, & Sen, 1993; Costa-Gomes, Crawford, & Broseta, 2001).

Taken together, these findings suggest that in the face of complex information processing, individuals intuitively converge on strategies that reduce the number of factors (dimensions) under consideration to make cognitive problems more tractable. Because the dimensionality of factors necessary for understanding human psychology is incredibly high, psychological researchers may be focusing on far fewer dimensions than what actually comprise the mind. That is, scientists may intuitively converge on establishing low-dimensional theories (e.g., dual process models) because they allow for a reduction of the

diverse set of factors relevant to building a comprehensive account of psychological processing.

2.3. Cultural norms

Together with individual biases and cognitive limitations, we believe that methodological traditions within the field of psychology have built a pedagogy that supports the Flatland fallacy. Heavily inspired by Ronald Fisher's iconic Statistic Methods for Research Workers (Fisher, 1925) and deeply embedded in undergraduate training in psychology, there is a strong persistent cultural tradition of academic psychology's reliance on a specific type of experimental design and statistical analysis to test hypotheses: two-dimensional factorial designs (i.e., two-way analysis of variance; ANOVA) evaluated via null hypothesis significance testing (NHST).

Although psychology does not have a generally agreed–upon core curriculum, at minimum almost all psychology undergraduate programs require that students take an introductory psychology course and one or more courses in statistics or research methods (Stoloff et al., 2009). Introductory statistics courses typically introduce basic concepts of inferential statistics, culminating in an introduction to two-way ANOVA. Research methods courses primarily emphasize making causal inferences using 2×2 factorial designs. In addition, approximately 40% of psychology programs require taking five courses in specific topics of psychology (e.g., abnormal, developmental, cognitive, social, biological) and a capstone course that results in a culminating experience for the psychology major (Stoloff et al., 2009). This limited training in statistical theory and psychological measurement persists into graduate training. A survey conducted in 1990 of psychology doctoral programs found that statistical training in most programs was highly similar to what it had been 20 years prior with 73% of programs providing in-depth understanding of "old standards of statistics" predominantly ANOVA, but only 21% provided more advanced training such as multivariate procedures (Aiken et al., 1990). A follow-up study 20 years later reached nearly identical conclusions with 80% of training still devoted to ANOVA and a continual general decline in measurement training (Aiken, West, & Millsap, 2008). In particular, Aiken et al., (2008) note the decline in coverage of measurement techniques (median 4.5 weeks of total PhD curriculum) with many programs offering absolutely no training in test theory or construction.

We believe this pedagogy not only trains generations of psychological scientists to pursue empirical investigations that favor simple factorial designs, but also teaches them to *think* about psychological science in a low-dimensional way. Consequently, the standard approach for inferential reasoning in psychology produces theoretical ideas that are limited to measuring mean differences between variables manipulated along a handful of psychological dimensions, thereby perpetuating the Flatland fallacy.

Although this approach provides an accessible starting point for the evaluation of psychological phenomena, significance testing of experimental manipulations alone does not constitute a formal model of human psychology and cannot be used as evidence for such (Bolles, 1962). The pervasive danger of relying on this approach when testing

psychological theory is that it reinforces the generation of weaker, less specific, and more nebulous theories, particularly in a growing era of "big data" (Meehl, 1990; Van Horn & Toga, 2014). This methodological paradox was elegantly articulated more than 50 years ago by Paul Meehl (Meehl, 1967). Whereas increasing experimental power constitutes a *more difficult* test that a quantitative theory must pass to remain viable (e.g., a mathematical model in physics), the exact opposite is true of psychological research, which is primarily concerned with utilizing NHST to detect arbitrary non-zero differences between experimental conditions. Because measurement error decreases with increasing power and precision, smaller non-zero differences will be *necessarily* more detectable in the limit of NHST. This leads psychological researchers to conclude that a statistically significant result of a trivially small difference provides support for a given theory. By instead developing theories as the "prediction of a *form* of function (with parameters to be fitted)" or "prediction of a quantitative magnitude (point-value)," the band of tolerance around theoretical validity *decreases* as experimental fidelity increases (Meehl, 1967). In other words, using NHST to evaluate theories formalized as models that make specific quantitative predictions ensures that theories must be *specific* in order to remain robust.

A classic example of the benefit of such an approach in psychology can be found in prospect theory (Kahneman & Tversky, 1979). Prospect theory outlines the form of a function that describes how the quantitative gain or loss that an individual incurs is mathematically transformed into the subjective *value* he or she feels (Fig. 2a). Critically, this theory defines a weighting function (Fig. 2b) for outcomes that cannot only be used to predict individuals' decisions, but also captures the asymmetry that occurs between changes in income framed as gains or losses. This model differs drastically from conventional approaches discussed previously in that it can be used to make point predictions about individuals' behavior and attitudes in a wide variety of decision contexts (Tversky & Kahneman, 1974).

2.4. Communicating complexity

Lastly, we consider how a pervasive communication problem endemic to the field of psychology as a whole gives rise to the Flatland fallacy. Central to this issue is that collectively, psychological researchers lack a *lingua franca* that enables us to communicate about the complexities of our field. While numerous succinct psychological constructs exist to describe complicated ideas (e.g., "stereotype threat" (Steele & Aronson, 1995), "affective forecasting" (Gilbert & Wilson, 2000), this terminology is fundamentally ad hoc, making it challenging to develop a shared and extensible framework for the communication and development of new ideas. In contrast to social science fields that rely on verbal descriptions of phenomena (e.g., psychology) (Watts, 2017), scientific disciplines that have methods of communicating about high-dimensional problems using mathematics (e.g., physics, computer science, economics), concrete physical models (e.g., biology, astronomy), or their own notation (e.g., chemistry) may be less susceptible to the Flatland fallacy. We argue that in the absence of a foundational dialect, psychology has suffered from an inability to communicate complexity, instead generating theories which make
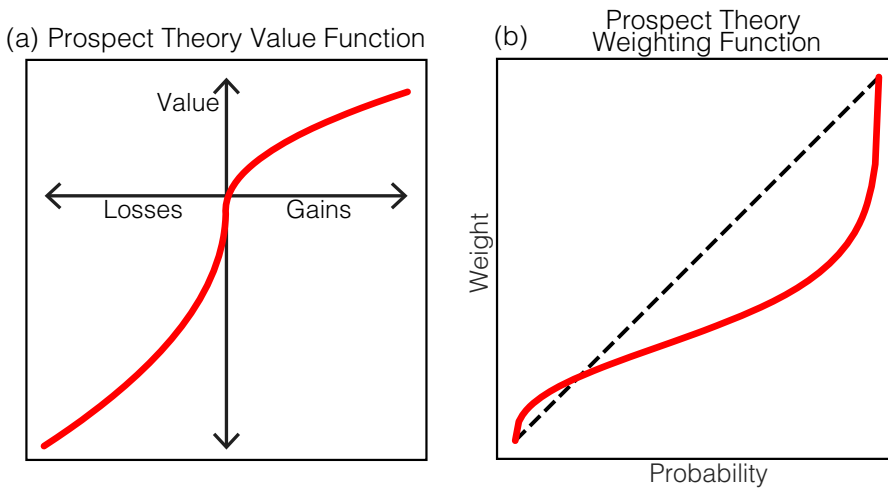
Fig. 2. (a) Prospect theory describes a mathematical *function* that maps between financial gains and losses and the subjective value that individuals experience. This function can be used to predict how an individual will make decisions and explains their tendency to exhibit risk-seeking or risk-averse behavior. (b) A central part of this theory is a functional account of how individuals treat probability values when making decisions, illustrating how small probability events receive more consideration than large probability events when making choices.

crude and imprecise predictions that may be difficult to falsify (Meehl, 1990). As a result, psychological discourse is more like a pidgin language than a *lingua franca*: simplified mixtures and non-specific generalizations that make it challenging to build a cumulative science.

This communication problem is particularly exacerbated by the limited number of dimensions available to psychologists to visualize their findings. In theory, data visualizations provide a useful tool for detecting patterns and interpreting research findings. Several recent technical and software advances in this area can provide aid in overcoming this limitation by learning "low-dimensional embeddings" of high-dimensional data that attempt to preserve distance (Zhang, Huang, & Wang, 2010), such as t-SNE (t-distributed stochastic neighbor embedding) (van der Maaten & Hinton, 2008) and UMAP (uniform manifold approximation and projection) (McInnes & Healy, 2018). Both of these algorithms are implemented in easy-to-use open source software packages such as hypertools (Heusser, Ziman, Owen, & Manning, 2018) However, even with such advanced techniques, graphical representations of research findings are typically limited to about three dimensions.[1] This is particularly problematic in the absence of a formal framework to build a cumulative science (e.g., mathematics), as psychologists are only able to interpret experiments that independently manipulate a few dimensions (e.g., a three-way ANOVA). This creates a tension between theoretical interpretability and theoretical extensibility. Because psychological science lacks a formal discourse, researchers are motivated to design experiments around low-dimensional theory testing and simple visualizations, but because experiments designed to test low-dimensional theories are not by themselves

extensible, psychological science exists as a patchwork of disparate, loosely connected ideas. Even recent advances in "theory mapping" (Gray, 2017), which provide organizational instructions as to how to connect psychological theories to each other, are underspecified and underconstrained (Newell, 1973) because they fail to provide comprehensive parameterized *models* of psychology that can be used to make useful quantitative predictions (Yarkoni & Westfall, 2017) about behavior or cognition.

3. What are some solutions?

We believe that overcoming the Flatland fallacy requires advances in the methodological approach that psychological scientists take toward their own work and also pedagogical changes that train future generations of researchers to build upon and extend extant work. First, we highlight the role that *computational models* can play in enabling researchers to overcome the biases and limitations in their own cognition, as well as enabling multiple researchers to work together to build a more cumulative science. Second, we outline some suggestions for the improvement of *psychological training*, stressing the importance of teaching students the fundamentals of mathematics and computer programming. We believe that both of these approaches are required to mature the field of psychological science.

3.1. Computational models

Formalizing psychological theories using computational models provides a way to overcome the Flatland fallacy through the consideration of *high dimensional* explanations of psychological phenomena. Indeed, recent methodological advances in neuroscience have demonstrated how information in the brain is encoded with incredibly high dimensionality with respect to both space and time (Haxby, Connolly, & Guntupalli, 2014). We believe the use of computational models will likewise better enable researchers to capture this complexity within psychological theories. Most psychological researchers are already familiar with regression as an instance of a *statistical model*, specifically a linear one that combines features according to a set of weights, in order to predict the value of a dependent variable. We encourage researchers to think of models in more general terms: a general mathematical function that transforms inputs into specific outputs.[2] In this way, models can be likened to cooking recipes that describe how to best combine ingredients into prepared foods (Crockett, 2016). We believe this analogy can help elucidate how models can play a critical role in overcoming the Flatland fallacy.

Imagine eating a piece of cake: observing the colors and designs that draw your eye, tasting the different flavors and textures as you take a bite, and experiencing the way numerous ingredients come together to create a delightful sensory experience. What ingredients determine the colors you see and the flavors you taste? Was that a hint of vanilla? Does the frosting contain cardamom? You might attempt to recreate this sensory experience in your own kitchen, combining numerous ingredients in different proportions,

transforming those ingredients through baking at different temperatures, until you can reliably reproduce a tasty baked good. Through a painstaking trial-and-error process you might converge on a *recipe* for *how* to combine and transform a set of raw ingredients into a finished product. In much the same way that recipes serve as a set of instructions for combining specific ingredients together that produce a cake, models provide mathematical formulations for combining different input *features* together that produce an outcome. These features can span a space that comprises only a few dimensions just as it takes only takes a handful of ingredients to make a pound cake. They can also, however, be incredibly enumerate and involve complex nonlinear interactions, just as a dobos torte requires intricately interleaving many layers of a thin sponge cake with buttercream. Indeed, with recent advances in deep learning, some models can comprise millions of distinct features organized into hidden layers that learn many different ways to filter and transform input data (e.g., Huang, Sun, Liu, Sedra, & Weinberger, 2016). Yet despite this wide range of complexity, researchers need not manipulate so many inputs at once single-handedly. Instead, they can rely on a model which serves as powerful and reliable assistive tool.

At their core, recipes are comprised of ingredients, proportions of ingredients, and instructions for how to combine ingredients to produce a finished product. Like recipes, models are comprised of *features* that are scaled by *weights* and combined in a specific *formula* to produce a *prediction* (Table 1). Central to our argument is that models serve as tools to both reason and communicate about *high–dimensional* spaces. Models allow researchers to consider what dimensions of a problem are most relevant and predict outcomes based on complex sets of interactions. Models also allow researchers to build intuitions about their components through simulation and application to novel datasets (Yarkoni & Westfall, 2017), akin to children taking devices apart to figure out how they work. Moreover, models can be shared between researchers, permitting the collective development of a cumulative science whereby weak or redundant theories are pruned and robust, predictive theories are retained.

A key property of formalizing psychological theories as computational models is that it enables researchers to share their ideas in *extensible* ways. In contrast, the significance result of an ANOVA is fundamentally useless to other researchers trying to extend prior theoretical work (Schmidt, 1996): *p*-values and effect sizes speak to the likelihood of the data under a null-hypothesis but are unable to provide precise predictions (point-estimates) about new data in new contexts (Meehl, 1967) or evaluate the predictive capabilities of competing ideas (i.e., model selection/comparison). At best, researchers will only be able to devise novel experiments to test contextual moderators on the coarse treatment effects that a theory predicts (Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016). On the other hand, models are simply recipes for combining inputs to generate predictions, and they can therefore be easily applied to novel contexts using new data. For example, we have developed models of how emotions (Chang & Smith, 2015; Chang et al., 2011), social norms (Chang & Sanfey, 2013; Sanfey, Stallen, & Chang, 2014), and inferences about others (Chang, Doll, van't Wout, Frank, & Sanfey, 2010; Fareri, Chang, & Delgado, 2015; Sul, Güroğlu, Crone, & Chang, 2017) can predict decisions to cooperate. In

Table 1
Relationship between models and recipes

Recipe	Model	Purpose
Ingredients	Inputs/features	Characterize the possible building blocks (dimension) necessary to create an output
Proportions	Weights/parameters	Characterize the relative importance of each input
Instructions	Formula/model specification	A describe a set of mathematical rules for ways in which inputs should be combined
Product	Prediction	The output consequent of combining inputs according to a set of rules with a set of fixed levels of importance

addition, we and others have developed models for how activity in different brain regions might be combined to produce an affective response (Chang, Gianaros, Manuck, Krishnan, & Wager, 2015; Eisenbarth, Chang, & Wager, 2016; Krishnan et al., 2016; Wager et al., 2013). Perhaps one of the best examples of how models can lead to a cumulative and extensible study of cognition can be found in neurally inspired connectionist models. This work strives to synthesize neuroscience findings into high-dimensional models of how the brain implements specific cognitive functions such as learning (O'Reilly & Frank, 2006; O'Reilly, Frank, Hazy, & Watz, 2007), memory (Norman & O'Reilly, 2003), and decision-making (Frank & Claus, 2006). These models have been integrated into a programming framework (e.g., Leabra) that provides a holistic architecture capable of making precise predictions of a wide array of cognitive processes (O'Reilly, Hazy, & Herd, 2016). The advantage of these types of quantitative models is that they permit precise evaluation of how *sensitive* a model is for capturing a psychological construct, as well as how *specific* a model is to a given construct as compared to other psychological states and processes. Further, these models are shareable and extensible by other researchers, allowing them to directly build upon previous work to test how a given model, such as a marker of negative affect, responds in a new experimental condition and generalizes to a novel context (e.g., Gilead et al., 2016; Krishnan et al., 2016). This process of iterative construct validation through model sharing and testing on many types of data is critical for developing a cumulative science of what comprises psychological states and how they are encoded and represented in the brain (Woo, Chang, Lindquist, & Wager, 2017).

3.2. Improved quantitative training

Although computational modeling offers an approach for uncovering psychological phenomena in higher dimensions, it requires a dramatic reform in the way the discipline of psychology carries out quantitative training. Instead of providing a limited introduction to inference using statistics and research design and separately prioritizing the memorization of psychological effects in different domains of psychology, we believe there should be increased emphasis on teaching technical skills and a better education of data-driven inferential reasoning within all psychology courses. Moreover, psychology's curriculum could be expanded to adapt to the recent technological advances that have resulted in the

exponential growth in the collection of data via the Internet, online commerce, mobile sensing, and so on (Griffiths, 2015; Lazer et al., 2009; Yarkoni, 2012). Beyond the narrow domain of academic psychology, there already exists intense demand for skilled workers who can gain insights about human behavior from data in almost every industry, including government, journalism, business, and healthcare. We believe that with improved training, psychologists could become increasingly involved in these efforts, ultimately providing an opportunity to inform an array of diverse and important industries and issues.

To do so will require reimagining training in psychology. Working with large, complicated datasets requires basic training in areas traditionally associated with computer science and informatics, including programming, algorithms, databases, and computing. This requires extending basic education of statistical training to include skills such as data manipulation, generating predictive models, machine-learning, natural language processing, graph theory, and visualization (Montag, Duke, & Markowetz, 2016; Yarkoni, 2012; Yarkoni & Westfall, 2017). These types of competencies comprise an emerging growth of applied statistics or "data science" programs (Anderson, Bowring, McCauley, Pothering, & Starr, 2014). We believe that increased emphasis on training basic technical and quantitative skills will improve the ability of psychology majors to participate in the enormous endeavor of understanding human behavior from data. However, we are certainly not advocating that psychology majors should additionally pursue an accompanying degree in statistics or computer science. Instead, we recommend that training programs in psychology consider adding additional requirements to curricula (e.g., programming and computing for psychologists & advanced statistics), providing better integration of data-driven inferential reasoning skills into existing curricula (e.g., requiring data analysis projects), and offering more courses on advanced research methods (e.g., natural language processing, mobile sensing, and/or social network analysis). One practical recommendation akin to training in other STEM disciplines (e.g., biology, chemistry, physics) would be to add accompanying laboratories to the core psychology topic classes and provide hands on training for making inferences using these types of methods. We believe that this is an exciting opportunity to advance our field to new dimensions.

3.3. Toward computational thinking

We hesitate to leave interested readers with the intuition that improved quantitative skills and the application of computational modeling are simply additional "tools" that researchers should strive to acquire in service of conducting high–dimensional psychological research. Rather, through the act of engaging in *computational thinking*, psychological researchers themselves can fundamentally change the way they approach psychological problems (Anderson, 2016; Newell, 1973).

For example, in addition to the psychological explanations enumerated above, there is also a statistical explanation for the Flatland fallacy. It is well established in the field of machine learning that making predictions entails a trade-off between minimizing bias and variance (Hastie, Tibshirani, & Friedman, 2009). Specifically, errors in predictions can be

decomposed into three separate components: (a) *irreducible error* is the level of noise intrinsic to the problem, (b) *bias error* describes errors between the true data-generating model and the learned model averaged across many data samples, characterizing the degree to which a learned model "underfits" a data sample, and (c) *variance error* reflects the learned model's sensitivity to the idiosyncrasies present within individual data samples, characterizing the degree to which a learned model "overfits" a data sample (Geman, Bienenstock, & Doursat, 1992). It has been demonstrated that in the context of small sample sizes and high-dimensional signals, lower dimensional models associated with greater bias error can counterintuitively make *more* accurate out-of-sample predictions than the true high-dimensional model of an underlying signal (Friedman, 1997). This means that there can be a computational benefit to prioritizing parsimony and bias (Gigerenzer & Brighton, 2009) when predicting complex psychological phenomenon from small datasets. It also offers an additional explanation for why so many researchers have converged on low-dimensional accounts of psychological phenomena: Lower dimensional theories better explain small or inadequately sampled datasets. We provide a simulation illustrating this point by demonstrating that a principal components analysis of high-dimensional data is biased to find a low dimensional solution when undersampled
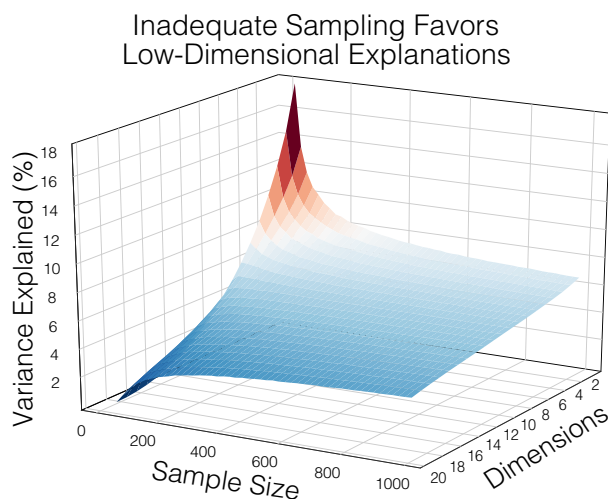

Fig. 3. This simulation illustrates how small sample sizes are biased to favor low-dimensional explanations even if the true underlying dimensionality of the data is high. We first generate a multivariate gaussian cloud using 10,000 observations comprised of 20 orthogonal dimensions and repeatedly draw random samples of increasing size from this space (100 repetitions per sample size). We then attempt to recover the dimensionality of the simulated data using principal components analysis and plot the distribution of variance explained across the computed components. When small samples are collected from a high–dimensional space, the majority of variance explained comes from the first few (2–3) components, which may lead researchers to mistakenly believe that the population itself is low dimensional. However, when larger data samples are collected, the amount of variance recovered across the dimensions becomes more uniform and in line with the data-generating process, which would lead to the correct conclusion that the sampled data and therefore the population from which they came are high dimensional.

(Fig. 3). While psychologists might be naturally inclined to further simplify their experimental manipulations and increase their sample size to improve power, computational thinking predicts adopting an alternative strategy. Rather than reducing high bias error via collecting larger sample sizes alone, computational thinking highlights the importance of measuring psychological phenomena with greater sampling *diversity*. For example, naturalistic experiments, comprised of free-viewing/listening to dynamic movies and unconstrained social interactions, elicit a greater range of psychological experiences (e.g., Chen et al., 2017; Haxby et al., 2011; Huth et al., 2016; Zadbood, Chen, Leong, Norman, & Hasson, 2017). By measuring and eliciting psychological phenomena in numerous ways, researchers can more richly sample high–dimensional effects of interest. It is with this data diversity that high-dimensional models can outperform biased lower dimensional alternatives. To this end, to combat the Flatland fallacy, we believe that psychological scientists should additionally strive for large sample sizes as well as large data *diversity*, richly sampling from a larger spectrum of human experience.

4. Summary

In this paper, we have outlined how subjective biases, limitations of human cognition, social and cultural norms surrounding experimental design, analysis and pedagogy, and communicative shortcomings in discussing complex ideas can lead psychological researchers to consistently converge on low-dimensional explanations of human psychology. To avoid committing this "Flatland fallacy," we have proposed a reimagining of the field of psychology, which emphasizes a culture of developing, testing, and sharing computational models, accompanied by improved quantitative and technical training. Like cooking recipes, models provide a formal framework for taking inputs and transforming them into products (predictions). More specifically, models offer researchers a tool to assist reasoning in higher dimensional ways, approaching psychological science as an explanatory and predictive discipline (Yarkoni & Westfall, 2017), and most important, facilitating the development of a cumulative science rather than one characterized by a patchwork of disparate low-dimensional theories (i.e., "you can't play 20 questions with nature and win" [Newell, 1973]).

We note that this is not a trivial proposal. Redesigning the entire curriculum for undergraduate and graduate training in psychology will take many years. The adoption of a computational framework will inevitably generate a host of additional complications and difficulties. For example, increasingly complex computational methods will necessarily increase the difficulty in communicating research findings. However, our methods of inquiry as scientists should not be determined by the ease with which we can communicate our work. Rather, our methods of inquiry should strive to directly tackle the immense complexity of our discipline while making it easier to collaborate, share, and develop our cumulative understanding. We believe these initial obfuscations provide a net benefit to both the general public and other scientists: They remind us that understanding how the human mind works is an incredibly challenging endeavor easily rivaling the

difficulty of problems studied for decades in disciplines such as physics, astronomy, chemistry, biology, and computer science and should be accordingly approached with commensurate awe, rigor, and humility.

Acknowledgments

Author contributions

E.J. and L.J.C. conceived of and wrote the manuscript.

Competing financial interests

Notes

1. With the clever manipulation of visual properties (e.g., color, texture, animation), it may be possible to graphically represent multiple interacting factors, but as the number of factors increases, difficulties in interpretation increase dramatically.
2. This definition captures a variety of statistical and machine learning approaches, including linear and nonlinear supervised, unsupervised, and reinforcement learning models.

References

Abbott, E. A. (1884). *Flatland: A romance of many dimensions, by A. Square*. London: Seeley & Co.
Aiken, L. S., West, S. G., & Millsap, R. E. (2008). Doctoral training in statistics, measurement, and methodology in psychology: Replication and extension of Aiken, West, Sechrest, and Reno's (1990) survey of PhD programs in North America. *The American Psychologist*, *63*(1), 32.
Aiken, L. S., West, S. G., Sechrest, L., Reno, R. R., Roediger, H. L., III, Scarr, S., Kazdin, A. E., & Sherman, S. J. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of PhD programs in North America. *The American Psychologist*, *45*(6), 721–734.
Anderson, N. D. (2016). A call for computational thinking in undergraduate psychology. *Psychology Learning & Teaching*, *15*(3), 226–234.

Anderson, P., Bowring, J., McCauley, R., Pothering, G., & Starr, C. (2014). An undergraduate degree in data science: Curriculum and a decade of implementation experience. In *Proceedings of the 45th ACM technical symposium on computer science education* (pp. 145–150). New York: ACM.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481.

Bolles, R. C. (1962). The difference between statistical hypotheses and scientific hypotheses. *Pharmacological Reports: PR*, *11*(7), 639.

Brunswick, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861–898.

Camerer, C. F., Ho, T.-H., & Chong, J. K. (2015). A psychological approach to strategic thinking in games. *Current Opinion in Behavioral Sciences*, *3*, 157–162.

Camerer, C. F., Johnson, E., Rymon, T., & Sen, S. (1993). Cognition and framing in sequential bargaining for gains and losses. *Frontiers of Game Theory*, *1*, 27–47.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105.

Chang, L. J., Gianaros, P. J., Manuck, S. B., Krishnan, A., & Wager, T. D. (2015). A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biology*, *13*(6), e1002180.

Chang, L. J., & Sanfey, A. G. (2008). Emotion, decision-making and the brain. In D. Houser & K. McCabe (Eds.), *Neuroeconomics* (Vol. 20, pp. 31–53). Bingley, West Yorkshire, UK: Emerald.

Chang, L. J., & Sanfey, A. G. (2013). Great expectations: Neural computations underlying the use of social norms in decision-making. *Social Cognitive and Affective Neuroscience*, *8*(3), 277–284.

Chang, L. J., & Smith, A. (2015). Social emotions and psychological games. *Current Opinion in Behavioral Sciences*, *5*, 133–140.

Chang, L. J., Smith, A., Dufwenberg, M., & Sanfey, A. G. (2011). Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*, *70*(3), 560–572.

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125.

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *Journal of Abnormal Psychology*, *51*(2), 291–294.

Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica: Journal of the Econometric Society*, *69*(5), 1193–1235.

Crockett, M. J. (2016). How formal models can illuminate mechanisms of moral judgment and decision making. *Current Directions in Psychological Science*, *25*(2), 85–90.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, *40*, 61–149.

Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York: Harper Collins.

Davidson, R. J. (1993). Cerebral asymmetry and emotion: Conceptual and methodological conundrums. *Cognition and Emotion*, *7*(1), 115–138.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.

Dawes, R. M. (1971). A case study of graduate admissions: Application of three principles of human decision making. *The American Psychologist*, *26*(2), 180–188.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*(4899), 1668–1674.

Deane, D. H., Hammond, K. R., & Summers, D. A. (1972). Acquisition and application of knowledge in complex inference tasks. *Journal of Experimental Psychology*, *92*(1), 20–26.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, *311*(5763), 1005–1007.

Dudycha, L. W., & Naylor, J. C. (1966). Characteristics of the human inference process in complex choice behavior situations. *Organizational Behavior and Human Performance*, *1*(1), 110–128.

Eisenbarth, H., Chang, L. J., & Wager, T. D. (2016). Multivariate brain prediction of heart rate and skin conductance responses to social threat. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(47), 11987–11998.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *The American Psychologist*, *49*(8), 709–724.

Evans, J. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459.

Evans, J., & Frankish, K. (2009). *In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.

Fareri, D. S., Chang, L. J., & Delgado, M. R. (2015). Computational substrates of social value in interpersonal collaboration. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(21), 8170–8180.

Fisher, R. A. (1925). Statistical methods for research workers. In F. A. E. Crew & D. Ward Cutler (Eds.), *Breakthroughs in statistics* (pp. 66–70). New York: Springer.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review*, *113*(2), 300–326.

Frank, M. J., Cohen, M. X., & Sanfey, A. G. (2009). Multiple systems in decision making. *Current Directions in Psychological Science*, *18*(2), 73–77.

Friedman, J. H. (1997). On bias, variance, 0/1—loss, and the curse-of-dimensionality. *Data Mining and Knowledge Discovery*, *1*, 55–77.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, *4*, 1–58.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*(1), 107–143.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–669.

Gilbert, D. T., & Wilson, T. D. (2000). Miswanting: Some problems in the forecasting of future affective states. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (vols. 1–8, pp. 178–197). Cambridge, UK: Cambridge University Press.

Gilead, M., Boccagno, C., Silverman, M., Hassin, R. R., Weber, J., & Ochsner, K. N. (2016). Self-regulation via neural simulation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(36), 10037–10042.

Gray, K. (2017). How to map theory: Reliable methods are fruitless without rigorous theory. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *12*(5), 737–741.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619.

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, *96*(3), 505–520.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105–2108.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4–27.

Griessinger, T., & Coricelli, G. (2015). The neuroeconomics of strategic interaction. *Current Opinion in Behavioral Sciences*, *3*(Suppl. C), 73–79.

Griffiths, T. L. (2015). Manifesto for a new (computational) cognitive revolution. *Cognition*, *135*, 21–23.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*(2), 129–154.

Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *10*(3), 252–264.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference and prediction*. New York: Springer.

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, *37*, 435–456.

Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., Hanke, M., & Ramadge, P. J. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, *72*(2), 404–416.

Heatherton, T. F., & Wagner, D. D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in Cognitive Sciences*, *15*(3), 132–139.

Heusser, A. C., Ziman, K., Owen, L. L. W., & Manning, J. R. (2018). HyperTools: A Python toolbox for gaining geometric insights into high-dimensional data. *Journal of Machine Learning Research*, *18*(152), 1–6.

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *4*(2), 162–176.

Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K. (2016, March 30). Deep networks with stochastic depth. *arXiv [cs.LG]* [accessed on October 1, 2017]. Available at http://arxiv.org/abs/1603. 09382

Huth, A. G., Lee, T., Nishimoto, S., Bilenko, N. Y., Vu, A. T., & Gallant, J. L. (2016). Decoding the semantic content of natural movies from human brain activity. *Frontiers in Systems Neuroscience*, *10*, 81.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *The American Psychologist*, *58*(9), 697–720.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*(3), 404–426.

Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. *Journal of Linguistics*, *43*(02), 365.

Karlsson, F. (2010). Syntactic recursion and iteration. In H. van der Hulst & H. van der Hulst (Eds.), *Recursion and human language* (pp. 43–68). Berlin: De Gruyter Mouton.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609–639.

Krishnan, A., Woo, C.-W., Chang, L. J., Ruzic, L., Gu, X., López-Solà, M., Jackson, P. L., Pujol, J., Fan, J., & Wager, T. D. (2016). Somatic and vicarious pain are represented by dissociable multivariate brain patterns. *eLife*, *5*. https://doi.org/10.7554/elife.15166

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Social science. Computational social science. *Science*, *323*(5915), 721–723.

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503–507.

McInnes, L., & Healy, J. (2018, February 9). UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv [stat.ML]* [accessed on June 28, 2018]. Available at http://arxiv.org/abs/1802. 03426

Medin, D. L. (1989). Concepts and conceptual structure. *The American Psychologist*, *44*(12), 1469–1481.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science*, *34*(2), 103–115.

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, *66*(1), 195.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89–115.

Miller, G. A. (1956). The magic number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 91–97.

Mitchell, J. P. (2005). The false dichotomy between simulation and theory-theory: The argument's error. *Trends in Cognitive Sciences*, *9*(8), 363–364.

Montag, C., Duke, É., & Markowetz, A. (2016). Toward psychoinformatics: Computer science meets psychology. *Computational and Mathematical Methods in Medicine*, *2016*, 2983685.

Montague, R. (2007). *Your brain is (almost) perfect: How we make decisions*. New York: Penguin.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium [accessed on June 28, 2018]. Available at http://repository.cmu.edu/cgi/viewcontent.cgi?article=3032&context=compsci

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*(4), 611–646.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. E. (2002). Rethinking feelings: An FMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, *14*(8), 1215–1229.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328.

O'Reilly, R. C., Frank, M. J., Hazy, T. E., & Watz, B. (2007). PVLV: The primary value and learned value Pavlovian learning algorithm. *Behavioral Neuroscience*, *121*(1), 31–49.

O'Reilly, R. C., Hazy, T. E., & Herd, S. A. (2016). The leabra cognitive architecture: How to play 20 principles with nature. In S. E. F. Chipman (Ed.), *The Oxford handbook of cognitive science* (vol. 91, pp. 91–116). New York: Oxford University Press.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, *16*(2), 366–387.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.

Pinker, S. (1999). How the mind works. *Annals of the New York Academy of Sciences*, *882*, 119–127; discussion 128–134.

Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734.

Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, *489* (7416), 427–430.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*(5), 521–562.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Sanfey, A. G., & Chang, L. J. (2008). Multiple systems in decision making. *Annals of the New York Academy of Sciences*, *1128*, 53–62.

Sanfey, A. G., Stallen, M., & Chang, L. J. (2014). Norms and expectations in social decision-making. *Trends in Cognitive Sciences*, *18*(4), 172–174.

Saxe, R. (2005). Against simulation: The argument from error. *Trends in Cognitive Sciences*, *9*(4), 174–179.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*, 379–399.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods*, *1*(2), 115.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

Shiffrin, R. M., & Schneider, W. (1984). Automatic and controlled processing revisited. *Psychological Review*, *91*(2), 269–276.

Simon, H. A. (1957). *Models of man: Social and rational*. New York, NY: Wiley.

Simon, H. A. (1987). Satisficing. In P. Macmillan (Ed.), *The new Palgrave dictionary of economics* (pp. 1–3). Basingstoke, UK: Palgrave Macmillan.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*(1), 3–22.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*(1), 3.

Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*(1), 218–254.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *The Behavioral and Brain Sciences*, *23*(5), 645–665; discussion 665–726.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811.

Stoloff, M., McCarthy, M., Keller, L., Varfolomeeva, V., Lynch, J., Makara, K., Simmons, S., & Smiley, W. (2009). The undergraduate psychology major: An examination of structure and sequence. *Teaching of Psychology*, *37*(1), 4–15.

Strevens, M. (2000). The essentialist aspect of naive theories. *Cognition*, *74*(2), 149–175.

Sul, S., Güroğlu, B., Crone, E. A., & Chang, L. J. (2017). Medial prefrontal cortical thinning mediates shifts in other-regarding preferences during adolescence. *Scientific Reports*, *7*(1), 8510.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(23), 6454–6459.

van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research: JMLR*, *9*, 2579–2605.

Van Horn, J. D., & Toga, A. W. (2014). Human neuroimaging as a"Big Data" science. *Brain Imaging and Behavior*, *8*(2), 323–331.

Wager, T. D., Atlas, L. Y., Lindquist, M. A., Roy, M., Woo, C.-W., & Kross, E. (2013). An fMRI-based neurologic signature of physical pain. *The New England Journal of Medicine*, *368*(15), 1388–1397.

Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour*, *1*(1), s41562–016–0015.

Waytz, A., & Mitchell, J. P. (2011). Two mechanisms for simulating other minds: Dissociations between mirroring and self-projection. *Current Directions in Psychological Science*, *20*(3), 197–200.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*(6), 1049–1062.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, *60*(2), 181–192.

Woo, C.-W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, *20*(3), 365–377.

Yarkoni, T. (2012). Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science*, *21*(6), 391–397.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *12*(6), 1100–1122.

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How we transmit memories to other brains: Constructing shared neural representations via communication. *Cerebral Cortex*, *27*(10), 4988–5000.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *The American Psychologist*, *35*(2), 151–175.

Zhang, J., Huang, H., & Wang, J. (2010). Manifold learning for visualizing and analyzing high-dimensional data. *IEEE Intelligent Systems*, *25*(4), 54–61.