aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Perspectives on Psychological Science
2021, Vol. 16(4) 803–815

DOI: 10.1177/1745691620966794
www.psychologicalscience.org/PPS
SAGE

Anatomy of a Psychological Theory: Integrating Construct-Validation and Computational-Modeling Methods to Advance Theorizing

Ivan Grahek[1], Mark Schaller[2], and Jennifer L. Tackett[3]
[1]Department of Cognitive, Linguistic, and Psychological Sciences, Brown University; [2]Department of Psychology, The University of British Columbia; and [3]Department of Psychology, Northwestern University

Abstract
Discussions about the replicability of psychological studies have primarily focused on improving research methods and practices, with less attention paid to the role of well-specified theories in facilitating the production of reliable empirical results. The field is currently in need of clearly articulated steps to theory specification and development, particularly regarding frameworks that may generalize across different fields of psychology. Here we focus on two approaches to theory specification and development that are typically associated with distinct research traditions: computational modeling and construct validation. We outline the points of convergence and divergence between them to illuminate the anatomy of a scientific theory in psychology—what a well-specified theory should contain and how it should be interrogated and revised through iterative theory-development processes. We propose how these two approaches can be used in complementary ways to increase the quality of explanations and the precision of predictions offered by psychological theories.

Keywords
theory, computational modeling, construct validation

Psychology is a hyperempirical science. Psychological scientists receive substantial training in data analysis and research design and use those skills to discover and publish novel findings. It is not surprising that, in the wake of prominent failures to replicate prior results (Open Science Collaboration, 2015), the ways in which psychological scientists collect, analyze, and report data have been scrutinized (Benjamin et al., 2018; Nosek et al., 2018; Simmons et al., 2011). But empirical practices are not the only ones being more closely examined. Building on older arguments about issues related to weak theories in psychology (Cummins, 2000; Gigerenzer, 1998; Kruglanski, 2001; Meehl, 1978, 1990), many have pointed toward the poor state of psychological theories as one of the causes of the current replication crisis (Fiedler, 2017; Gray, 2017; Morey et al., 2018; Muthukrishna & Henrich, 2019; Oberauer & Lewandowsky, 2019; Schaller, 2016; Smaldino, 2017).

Although there is a wealth of articles identifying the benefits that can accrue from better theories, there is a lack of instruction on how to accomplish this goal. It is not clear what constitutes a strong theory in psychology, how such theories should be built, and how the strength of the existing theories can be evaluated. With this deficit in mind, here we focus on the methodology of theory building. We highlight two existing methods: construct validation and computational modeling. These methods can be taught, learned, and used to produce better theories and hypotheses.

Construct validation and computational modeling have been a part of the psychological landscape for decades, and each is commonly used to accomplish specific research objectives. However, their broader utility in the service of theory development remains underappreciated. In the sections below, we outline the

Corresponding Author:
Ivan Grahek, Department of Cognitive, Linguistic, and Psychological Science, Brown University
E-mail: ivan_grahek@brown.edu

points of convergence and divergence among them to illuminate the anatomy of a theory in psychology—what a well-specified theory should contain and how it should be continually revised through an iterative theory-development process. We propose that these two methods can be combined to achieve two characteristics of strong theory: mechanistic explanations of psychological phenomena and predictions with a narrow range. We argue throughout that theoretical and methodological rigor are interdependent and that both are necessary for progress within the psychological sciences.

Theory in Psychology: Weak Explanations and Wide Predictions

A scientific theory consists of a set of connected statements that work together to explain and predict observable phenomena. Explanations offered by a theory lead to testable predictions, often resulting in theory revision. This process leads to cumulative knowledge as theories are iteratively modified, improved, or eventually abandoned (Meehl, 1990). This may seem like a straightforward recipe, but it is not. Few psychological phenomena are directly observable, and psychologists theorize about latent constructs such as neuroticism, working memory, or altruism. Psychological theories establish the existence of such constructs and postulate how they interact to explain and predict psychological phenomena (Fidler et al., 2018; Gawronski & Bodenhausen, 2015). The fundamental challenge facing theorists in psychology is to create testable theories that provide increasingly better explanations that produce ever more precise predictions.[1]

Explanations in psychology

What does it mean to explain a psychological phenomenon? In physics, explanation commonly requires subsuming a phenomenon under a universal law. In contrast, psychology, and many other life sciences, rarely postulate such laws. Rather, explanations in these sciences specify parts and relationships of a system to explain how they give rise to phenomena. For example, the increased probability of performing a previously rewarded action can be explained by a learning mechanism that tracks the expected value of actions and updates those values on the basis of difference between expected and observed rewards following different actions (i.e., reinforcement learning models; Dolan & Dayan, 2013). To offer another such example from another domain of psychology, the likelihood of a child demonstrating a relationally aggressive behavior (e.g., social exclusion) may be predicted by examining other aspects of the child's personality, the presence of other forms of psychopathology in the child, and the nature of the child's social relationships and functioning (Brandes et al., 2021).

Such explanations aim to be mechanistic and are commonly found in psychology and biology. These types of explanations rely on the precise specification of the component parts, the operations performed by these parts, and dynamic patterns of change in the properties of the components and their operations (Bechtel & Abrahamsen, 2005, 2010; Craver, 2006). A good mechanistic understanding provides an explanation by detailing the system that gives rise to the phenomenon. In psychology, many theoretical explanations attempt to do this but fall short because of imprecisely specified components and imprecisely defined relationships between components (Cummins, 2000; Hommel, 2020). Psychological theories often specify weak relationships (e.g., A decreases B) between poorly understood latent constructs (e.g., anxiety and inhibition).

Predictions in psychology

Strong explanations give rise to precise predictions that can be empirically verified. In this way, theories are corroborated or gradually abandoned. However, relative to other sciences, psychology progresses slowly toward cumulative knowledge, and this slow scientific progress is due in part to the difficulty of actually falsifying predictions derived from most psychological theories (Meehl, 1978, 1990). This difficulty follows from two sources of imprecision. First, many theories yield predictions that cover a wide range of possible empirical outcomes (i.e., *spielraum* of a theory; Meehl, 1990). For example, a theory might simply predict a nonzero positive correlation between constructs without offering any more precise prediction about the magnitude of that nonzero effect. This would be the case if we predict that neuroticism and depressive symptoms are positively related. Second, many theories are characterized by vaguely defined constructs that cover a broad range of possible empirical operationalizations (Oberauer & Lewandowsky, 2019). This makes falsifying predictions difficult: When we fail to observe the predicted relationship between neuroticism and depression, we can try to use a different self-report measure for each of them or assess them through task-based measures. To be truly falsifiable and to produce the most confident predictions, theories in psychology should be articulated in a manner that tightly constrains the search space of operational variables that might plausibly correspond to hypothetical constructs. Theories should also predict a tightly constrained range of results.[2]

Building Better Theory in Psychology: Two Methods

Here we introduce two existing sets of methodological tools and practices that can be used to construct theories that offer better explanations and more precise predictions. They come from two distinct research traditions: individual-differences and experimental psychology (Cronbach, 1957). However, both methods share the same goal of increasing precision in the specification of psychological constructs and their relationships. Both strive to achieve this goal through an iterative process of theory building, testing, and revision.

Construct validation

Construct validation (Clark & Watson, 2019; Cronbach & Meehl, 1955; Meehl, 1986) is an iterative theory-development process based on the idea that construct operationalization and measurement should reflect an intertwined iterative process directly informing—and informed by—overall theory development in psychological science. Despite its widely relevant and generalizable implications across psychological science, construct-validation approaches have been primarily used in individual-differences research areas such as clinical and personality psychology.

The construct-validation process can (and should) consist of many components (Clark & Watson, 2019; Tay & Jebb, 2018), including construct operationalization, psychometric properties of a given scale presumed to measure the construct (e.g., reliability, structural validity), and adherence to a theoretically specified nomological network of other measures in association with the target construct (e.g., convergent/discriminant validity and predictive validity). All aspects of the construct-validation process rely on extant theory. Each aspect of the construct-validation process requires formulating scientific hypotheses (e.g., to what other psychological variables your construct of interest should or should not be related) and testing those hypotheses with data (Cronbach & Meehl, 1955). In this way, all scientific investigations are limited by the measures researchers use, and the measures used are similarly limited by the theoretical integrity underpinning them.

Despite the widespread citation of the classic 1955 article by Cronbach and Meehl (Google scholar currently reports more than 12,500 citations), the process of construct validation is often misunderstood in modern psychological science. This misunderstanding leads to the creation and use of measures with little, if any, established evidence of construct validity. Common misconceptions of the construct-validation process include focusing on one aspect of construct validity to the exclusion of others (e.g., equating a satisfactory Cronbach's alpha or confirmatory factor structure with "achieving" construct validity), misunderstanding the integral and foundational role of theory in the construct-validation process, or attributing construct validity to a property of an instrument rather than an ongoing and iterative theory-development process via interpretations of an instrument's measurements (Clark & Watson, 2019; Cronbach & Meehl, 1955; Tay & Jebb, 2018). These misunderstandings likely result in researchers in some areas of psychology viewing construct validation as irrelevant for them and researchers in other areas of psychology who identify more with construct validation often approaching it incompletely and, thus, inaccurately.

Examples of construct-validation efforts do exist, although the limited scope of the current article does not allow an exhaustive review of these efforts. Efforts focused on comprehensive and thorough construct-validation efforts in their own right are less common, although it is clear that such efforts have far-reaching implications for measurement and theory (for an example of a comprehensive construct-validation effort, see Brandes et al., 2021). A comprehensive approach to construct validation would also elucidate the extensive overlap with the steps involved in the computational-modeling process. The common practice of ignoring construct validation can result in a host of problems, including improper construct specification, poor psychometric properties of instruments, and a proliferation of redundant and amalgam constructs in the literature, all of which are likely contributors to the current crisis of replication in psychological science (Maul, 2017; Meehl, 1986). Narrative justification for scale creation or use is no substitute: "Rationalization is not construct validation" (Cronbach & Meehl, 1955, p. 291). The importance of construct-validation efforts is clear, but the field remains in dire need of more empirical demonstrations of this process in the literature.

Computational modeling

Whereas construct-validation methods focus on specific constructs and their operationalization, computational-modeling methods are typically used to inform researchers about specific processes such as decision-making or the evolution of altruism. Computational models formalize psychological processes through mathematical models. They specify the algorithms through which inputs to the model (e.g., rewards following actions) are transformed into outputs (e.g., selection of actions). In this way, computational models offer a formal account of the parts of psychological processes and their functions, as well as a set of testable predictions (Farell & Lewandowsky, 2018; Forstmann & Wagenmakers, 2015; Smaldino, 2017).

Cognitive science has a long tradition of developing formal models of processes such as decision-making (Ratcliff, 1978, 1981), learning (Rescorla & Wagner, 1972; Sutton & Barto, 1981), and vision (Marr & Poggio, 1979). Such models can simulate trial-by-trial behavior (e.g., decisions or reaction times) and thus offer strong predictions that can be compared with empirical results (for introductions to modeling in cognitive science, see Farell & Lewandowsky, 2018; Heathcote et al., 2015; Lee et al., 2019; Wilson & Collins, 2019). In the next step, computational models can be fit to data to extract latent variables (e.g., the expected value of different actions in the environment). As a result of these benefits, computational models are becoming increasingly popular in fields outside of the strict scope of cognitive science. Formal models of cognitive processes are used to understand psychopathology (Grahek et al., 2019), emotion (Eldar et al., 2016), the relationship between emotion and cognition (Grahek et al., 2020), and the neural implementation of cognitive processes (Forstmann & Wagenmakers, 2015).

Outside of the domain of cognitive science, computational models are commonly used to identify plausible outcomes of evolutionary processes (e.g., the evolution of altruism; Fletcher & Doebeli, 2009; Hammond & Axelrod, 2006). Such models, often used in biology, are increasingly implemented in social and cultural psychology to model long-term population-level processes. These agent-based models are used to generate testable hypotheses in domains that defy easy logical analysis (e.g., models of complex social-influence dynamics and their implications for cultural evolution; Muthukrishna & Schaller, 2020; for additional social-psychological applications, see Smith & Conrey, 2007; Jackson et al., 2017).

Although computational models have recently been applied to a wide range of psychological research topics, modeling methods remain an underused tool for theory development in psychological sciences (e.g., Jackson et al., 2017; Muthukrishna & Schaller, 2020; Smaldino et al., 2015). As theory-building tools, they offer two benefits: First, they force the theorist to be explicit about all of the assumptions that go into the model (e.g., in construct validation, assumptions guiding choices for operationalization and relationships specified in the nomological network); second, they offer the possibility to derive precise predictions from the model (in a construct-validation framework, guiding predictions around convergent/divergent and criterion validity). When computational-modeling approaches are overlaid on construct-validation approaches, as we detail more extensively below, resultant theories are substantially strengthened by leveraging the benefits of both.

Theory-Building Steps in Construct Validation and Modeling

Although construct-validation and computational-modeling methods are typically applied to different kinds of research questions, these methods share some fundamental similarities. One important similarity is that both methods truly are *methods*—practical procedures that can be taught and learned and thus incorporated into textbooks and courses on research methods, which can help to overcome the lamentable deficit in theory-oriented methodological training (Gray, 2017; Kruglanski, 2001). These methods also share broad procedural similarities. First, both methods require researchers to conceptually articulate a theoretical model and specify how the theoretical model can be empirically assessed. Second, both methods call for iterative revision of the model on the basis of empirical results. However, there are also important differences between these methods, which we review below.

Articulation of a theoretical model

Theoretical models developed through construct validation articulate the latent construct of interest by specifying its measurement and positioning it within a wider network of related constructs. Constructs are defined in a way that enables the specification of the hierarchical structure of the construct (Clark & Watson, 2019) and the continuum of measurable responses that characterize the construct (Tay & Jebb, 2018). Alongside the conceptualization of the focal construct, researchers also identify an entire universe of other constructs that are related to it (as causes, consequences, or correlates). The theoretical core of construct validation is the nomological network (Cronbach & Meehl, 1955). The network specifies how the construct of interest is related to other latent constructs, relying on existing scientific knowledge to create testable hypotheses about every component of the network (and thus about the construct itself).

The nomological network offers a set of hypotheses specifying which constructs should be positively related to the focal construct (convergent validity) and which should be negatively related to the main construct (discriminant validity). In doing so, the types of psychological processes and mechanisms frequently evoked in computational modeling are embedded in the construct-validation process. A construct-validation study focusing on a specific operationalization of depression may specify a nomological network hypothesizing a strong, positive association with a measure of trait neuroticism and a strong negative association with trait extraversion. Such hypotheses allow for the explication

of underlying mechanisms: For example, the experience of depression may reflect a more severe behavioral manifestation of underlying tendencies to experience high levels of negative affect as well as tendencies to experience low levels of reward motivation or behavioral activation. In this way, each aspect of the nomological network embeds the type of mechanistic features of a computational model, although this is not typically made explicit in many construct-validation efforts. Likewise, articulating these hypotheses in a construct-validation effort produces directly testable empirical hypotheses that can be tested, resulting in information entered back into the model, revisions made, and new tests performed—another process that is very similar to a standard computational-modeling design.

Computational-modeling projects are typically motivated by a conceptual analysis of a specific psychological process. This analysis determines the problem that a process is trying to solve and how to solve it (Farell & Lewandowsky, 2018; Marr, 1982; Wilson & Collins, 2019). For example, assuming that attention is a limited resource, a process can solve the problem of how and when to allocate this resource. After determining the problem, the analysis proceeds to propose a good algorithm that can effectively solve this problem (e.g., weigh the costs and benefits of allocating attention in a certain way). The specification of the goal of a process is a crucial step in selecting an algorithm through which the process can achieve that goal.

In computational modeling a theoretical model is implemented by a set of algorithms that specify the components of the model (which would align with "constructs" in a construct-validation sense) and the interactions between them. These algorithms transform inputs to the model (e.g., rewards and costs associated with different ways of allocating attention) into outputs (e.g., the allocation of attention in a specific way). In this way, the model produces behavior that can then be compared with the empirically observed behavior in experiments (Heathcote et al., 2015; Lee et al., 2019; Wilson & Collins, 2019). This step commonly includes simulating the results from the model by varying inputs (i.e., experimental conditions) and parameter values of the model (i.e., components of the model) to simulate experimental conditions. These procedures also allow researchers to systematically vary the model inputs and parameter values to simulate outputs that are predicted to occur under different experimental conditions. In this way, the model produces precise quantitative results that can then be compared with the empirical results generated by conceptually analogous experiments.

Although both approaches begin with the conceptual articulation of a construct under investigation, they do so in different ways. Construct validation defines the construct of interest (e.g., neuroticism) and positions it within a wider network of related constructs. In this approach the construct under investigation is of primary interest, whereas the specification of the relationships between latent constructs is secondary. In contrast, computational modeling focuses primarily on specifying the formal relationships through which constructs interact to solve a problem (e.g., how much cognitive resources to invest). In this approach the relationships between constructs are of primary interest, whereas the specification and measurement of the included constructs often come second.

Appraisal of evidence and consequent conceptual revision

Both of these methods involve testing and revising theory on the basis of empirical data. The construct-validation approach tests aspects of the theoretically specified operationalized construct and the hypothesized nomological network. This involves testing the measurement of constructs, the relationships between them, and the operationalization of the main construct of interest. All of these aspects of the theory are further revised on the basis of the data, and the nomological network is elaborated and the theoretical understanding of the main construct of interest is increased.

Many—although not all—computational-modeling projects also involve the collection and critical appraisal of empirical data. This step involves comparing the simulated and observed data as well as the fitting of the model to the empirical data. Further, in this step the parameters of the different latent constructs are extracted after the fitting to the empirical data and then validated against other relevant constructs such as self-report measures of psychopathology (Rouhani & Niv, 2019) or neural data (Daw et al., 2011; Fischer & Ullsperger, 2013). The model is further refined on the basis of the results of this step; unnecessary components can be removed from the model, and new components can be added (a parallel step to refining the nomological network). Finally, in both approaches different competing models can be compared in their goodness of fit to the data.

Both construct-validation and computational-modeling methods have an iterative quality. Neither methodology is static, and rarely will the end product of either method be a simple falsification and abandonment of a theory. Rather, both methodologies are dynamic and highly constructive sets of tools designed for theory building. When considered together, these tools can lead to more highly elaborated research programs (Lakatos, 1976) and can guide psychological scientists toward more refined understandings of latent constructs

and their relationships, as well as better ways of measuring them.

Toward an Integrated Method: How to Develop a Psychological Theory

Building a theory starts with a question. To build a good theory, we need to ask good questions and imagine what satisfying answers would look like. Thinking about questions and answers allows us to set the appropriate level of abstraction of our theory. For example, if our goal is to understand the reasons behind individual differences in job performance, we will probably want to generate a theory at the level of personality traits, motivation, and social interactions. That would be the correct level of abstraction that will likely generate good and useful predictions. The process of accumulating visual evidence toward a decision bound is definitely a process involved in job performance; however, this level of abstraction is unlikely to be useful for our theory. Of course, the reverse is also true: If we want to understand the properties of stimulus-driven attention, including the construct of work motivation in our explanation is unlikely to be helpful. Both construct validation and computational modeling can be helpful tools regardless of our chosen level of abstraction. What is important is that our theory can offer satisfying explanations and useful predictions at the level of abstraction at which it operates.

Once we have determined the appropriate level of abstraction of our theory, we can start thinking about how to create good explanations and generate good predictions. This is where construct validation and computational modeling have unique strengths and weaknesses, and this is precisely where an integrated approach to theory building, leveraging the strengths of both methods, can be useful. The quality of explanations and predictions offered by our psychological theory depends on its anatomy: the clarity of its parts (latent constructs) and the precision of relationships between the parts (interactions between latent constructs). Construct validation can help us better define our constructs, and computational modeling is a great tool for developing a precise understanding of the relationships between constructs. Here we detail these advantages and how they can be integrated.

Better explanations

A complete explanation of a psychological phenomenon requires an understanding of underlying mechanisms. To provide a good description of such a mechanism, we need to precisely specify its component parts and stipulate how they interact to produce the phenomenon of interest (Bechtel & Abrahamsen, 2010; Craver, 2006). Theories developed through construct validation and computational modeling both attempt to do so but have complementary strengths and weaknesses in their ability to provide mechanistic explanations. In practice, construct validation is often more focused on determining and defining the components, whereas computational modeling is more focused on how the components interact to produce phenomena of interest.

Detecting the parts: lessons from construct validation. Computational models are characterized by the structure that parallels mechanistic explanations: They specify components of a process and the nature of relations between components. However, although all computational models have this kind of structure, models differ in the quality of their explanations. An important part of these differences comes from the extent to which the components of a model are well defined and can be reliably measured.

For instance, in some areas of biology, computational modeling commences after researchers have already accumulated substantial knowledge about the components, their operations, and the organization of a mechanism (see Bechtel & Abrahamsen, 2010; Servedio et al., 2014). Further, in these biological disciplines, there often exist well-established, highly reliable, and demonstrably valid means of measuring the essential components (e.g., sex ratio) or mechanistic processes (e.g., protein synthesis). Although this is sometimes the case in psychological science, it very often is not. Psychological scientists often create models of processes that are difficult to measure and include components that have only face validity. Many modeling projects begin with assuming the components and their interactions, and models are usually validated by demonstrating that the model can simulate the relevant phenomena exhibited by humans or other animals. This is especially true in areas outside of the strict scope of cognitive science, such as clinical psychology, affective science, or social psychology. For example, let us go back to our task of creating a theory to explain individual differences in job performance. One can easily create a model with many interacting parameters (e.g., Job Performance = Ability + Level of Education + Motivation / (Anxiety × Hostility)). This model will be able to generate a number of predictions, some of which will likely resemble empirical data (e.g., job performance will improve with an increase in work motivation). However, this model is useless because it operates on ill-defined constructs that have undefined measurement properties and an unclear hierarchical structure.

A part of this problem can be addressed by applying construct-validation methods along with computational modeling during theory development and revision. Construct validation offers a systematic methodology that can be used to define a construct at a conceptual level of analysis—including its hierarchical structure and the continuum of stimuli or responses that might characterize it (Clark & Watson, 2019; Tay & Jebb, 2018). In this way, the components included in the computational model can be validated. Before starting to simulate findings from our model of job performance, we will need to conduct a construct-validation study aimed at creating a refined nomological network of this construct. This network can include all of the abovementioned components, but it will iteratively revise the network until a satisfying set of connections, hierarchical structures, and measurement properties are defined. For example, this means spending time on testing whether anxiety and hostility can be explained by one factor in the network or not. This will also include developing reliable measures for each of these constructs.

The focus of construct validation on quantifying and reducing measurement error can be very useful in developing and testing theories. For example, if our model of stimulus-driven attention predicts a 20-ms difference between experimental conditions but the measurement error on a behavioral task assessing the relevant construct is 100 ms, a researcher might be wise to delay model testing until a more sensitive and reliable measure can be developed. This can be particularly useful in those domains of experimental psychology in which researchers have historically paid relatively little attention to measurement issues, especially when applied to individual-differences research with behavioral tasks (Hedge et al., 2018; Rouder & Haaf, 2019).

Detecting the relationships: lessons from computational modeling. Construct-validation methods, especially in the domains of personality and clinical psychology, can be used to develop theoretical models—nomological networks—that establish the existence and measurement of constructs, as well as directional and hierarchical relationships between constructs (e.g., DeYoung, 2015; Kotov et al., 2017). Such networks systematically specify the universe of other constructs—causes, consequences, correlates—that might plausibly be related to a construct of focal interest. In our example this would correspond to a well-developed nomological network of job performance. However, in practice, nomological networks commonly lack the formal specification of the exact causal mechanisms, such as explicit statements about the quantitative aspects of the association (e.g., a positive correlation between .40 and .60) or full elucidation of the dynamical operations and interactions between constructs in the nomological network.

Computational modeling begins with the question about the problem that a process is trying to solve. From that starting point, a modeler thinks about the potential sets of components and their relationships that can solve the posed problem. For example, we can pose the question of what determines how well people perform on their job. One proposal could be that people are trying to balance the costs (e.g., effort, work hours, stress), benefits (e.g., great colleagues, a sense of purpose, salary), and their own abilities (e.g., level of education) to determine their level of performance. Modeling can then proceed to search for different algorithms that could plausibly be useful in modeling this complex decision-making process. Once such algorithms are found, computational models can rely on construct-validation efforts for the specification and testing of the included components. However, modeling methods will push construct-validation efforts to move beyond vague unspecified hypotheses (e.g., external motivation is positively related to job performance).

Computational models generate a clear articulation of the relationships through which components of a model interact. In this way, computational models provide very specific predictions at both the conceptual and numerical level. For example, a computational model can suggest that the relationship between external motivation and job performance depends on previous learning history related to received rewards on other jobs. In this way, the developed theory can be conceptually strengthened, which will translate into a constrained empirical space for hypothesis testing (e.g., the theory will need to be revised if it predicts a correlation between .40 and .60 but the observed correlation is .15). It is this kind of additional conceptual development—the transparent and precise mechanistic articulation of each individual strand within a nomological network—for which computational-modeling methods are ideally suited (Oberauer & Lewandowsky, 2019; Smaldino, 2017). In this way, the theories developed through construct validation could capture the more dynamic picture of the underlying psychological processes.

Better predictions

Good theories produce increasingly narrow predictions that can be tested against empirical data. This is the main route to cumulative knowledge in science: Theories are corroborated, modified, or abandoned on the basis of empirical evidence. If our theory predicts that job performance will be positively related to the motivation of a worker, our theory is practically worthless:

This is an extremely wide prediction that is difficult to reliably test. To improve our theory, we need to narrow the prediction range of our model and specify the ways of measuring the components of the model. Computational modeling and construct validation play complementary roles in achieving these goals and thus improving the precision of predictions derived from psychological theories.

Narrowing the prediction range: benefits of computational modeling. A crucial characteristic of a good theory is that it produces predictions that can be corroborated or falsified. This is the case only when the predictions of a theory have a narrow range relative to all possible empirical outcomes (i.e., spielraum; Meehl, 1990). When constructing a theory, or analyzing one, we should detect the entire range of possible outcomes of the crucial variables. Then we should analyze how much of that space is restricted by the prediction made by the theory. For example, our weak theory can predict that job performance will improve as a function of motivation. The entire range of possible empirical outcomes ranges from the maximal negative correlation to maximal positive correlation. The directional prediction of our theory covers 50% of that space. Thus, our theory is not doing better than a coin flip in predicting psychological phenomena.

The main advantage of computational-modeling methods is that they push researchers to formalize their models. Formal models constrain the prediction range in at least three useful ways. First, they specify interactions between multiple parameters that are commonly beyond simple linear relationships between constructs. In this way, predictions from a theory can be more specific (e.g., an inverted U-shape relationship between job difficulty and performance), and thus the prediction range is narrowed. Second, formal models can simulate psychological phenomena (Wilson & Collins, 2019), allowing for the generation of numerical predictions from the model. Indeed, this is one of the primary benefits of computational-modeling methods. They enable researchers to produce numerical predictions that are far superior to the directional predictions that characterize many verbal theories. Third, these methods compel researchers to transparently specify the logical components of a hypothesis—including underlying assumptions, logical derivations from those assumptions, and the sequence of psychological events through which one construct is hypothesized to influence another construct. When these logical components and mediating mechanisms are made transparent, researchers are more likely to consider them critically and to appraise their generalizability. A key assumption may hold in some cultural contexts, for instance, but not others; and a presumptive sequence of psychological events may readily occur under some conditions but be disrupted under others (Klatzky & Creswell, 2014; Schaller, 2016). By compelling researchers to transparently specify the mechanistic logic underlying a hypothesis, computational modeling provides a systematic means through which researchers can more readily discover—and therefore make more precise predictions about—specific conditions under which a hypothesized effect is likely, or unlikely, to occur.

Better specification of predictions: benefits of construct validation. The narrow prediction interval of a theory is not useful unless the ways of testing that prediction are precisely specified. For example, a narrow prediction interval about the level of correlation between two constructs is useless if those two constructs can be measured with several different measures that will produce different results. Our computational model specifying the process through which people set their level of performance at work will be of little use if it includes components that have large measurement error relative to the prediction interval.

A necessary step in improving the predictions arising from a theory is to precisely specify the ways in which the relevant constructs can be measured (Oberauer & Lewandowsky, 2019). Construct-validation methods provide a systematic method through which the measurement of latent constructs can be specified. A construct-validation study would aim to directly specify the measurement of each of the constructs within the nomological network of job performance. This limits researchers' degrees of freedom when testing the underlying theoretical model, thus allowing that theory to be more truly falsifiable.

Putting These Principles Into Practice

This article has attempted to explain why construct-validation and computational-modeling methods can be an asset to theory development in psychology, but we have not directly addressed the question of how these principles can be put into practice. We do not have space here to provide a fulsome and detailed answer to that question. Instead, we briefly identify some potentially useful points of departure.

One step in the right direction requires no technical expertise whatsoever. It is simply to adopt the habit of asking conceptual questions of the sort that are associated with construct-validation and computational-modeling methods. Consider first the kinds of questions associated with construct validation. For any psychological construct of interest (e.g., embodied cognition, relational aggression, parental caregiving motivation), it is valuable to ask questions that help to define the conceptual space occupied by that construct, such as:

What exactly is it? What isn't it (i.e., in what ways is it conceptually distinct from other similar-seeming constructs)? Is it truly a single coherent psychological construct, or might multiple meaningful constructs be lurking here that might best be conceptually defined, and operationalized, as separate psychological entities? These questions are complemented by additional questions—about causes, correlates, and consequences—that are necessary for constructing a nomological network. As with all scientific endeavors, these questions are best addressed in a systematic way. For instance, deeper insights into the causes of many psychological constructs may be achieved through deliberate consideration of distinct questions about underlying mechanisms, development, adaptive function, and phylogenetic history (i.e., "Tinbergen's four questions"; Nesse, 2013). Questions about correlates and consequences can also be systematically subdivided into separate questions pitched at different levels of psychological analysis (e.g., behavioral manifestations, cognitive operations, neural processes). These are just examples. The broader point is that when researchers make an effort to systematically consider questions that characterize the "conceptualization" phase of construct validation (Clark & Watson, 2019), the effort can pay off in the form of more precise conceptual definitions of constructs (and, consequently, better measures of those constructs), more carefully articulated theories about those constructs, and more nuanced hypotheses that make accurate predictions.

A complementary set of questions are associated with computational-modeling methods. Because these methods require that constructs be implemented in terms of algorithmic operations, researchers are faced with questions about how best to do so (e.g., "Which of these possible algorithmic operations corresponds most closely to the conceptual definition of this construct?"). Researchers are thus compelled to more precisely define the conceptual space occupied by a construct. In addition, because computational models require researchers to specify computational steps through which each construct exerts a hypothesized effect on another construct, researchers are confronted with questions about the mechanistic logic of the hypothesized relation (e.g., "What precise sequence of intervening psychological events might plausibly account for the hypothesized effect?"). Consequently, after specifying these sequences of operations, researchers are in a better position to ask additional questions that probe the robustness of each hypothesized relation (e.g., "What assumptions must one make to be confident about the plausibility of this particular sequence of psychological events? Under what conditions might this particular sequence of events be less plausible?").

To more fully avail oneself of the benefits of construct-validation and computational-modeling methods, one needs to acquire some expertise in these methods. It would be great if construct-validation and computational-modeling methods were part of every psychological scientist's formal education, but the absence of formal instruction need not be a barrier. A little reading goes a long way. Clark and Watson (2019) provide an excellent introduction to construct-validation methods, and Brandes et al. (2021) offer an illustrative example showing how construct-validation methods can be systematically used to achieve a better understanding of one particular construct (relational aggression). As for computational modeling, Smaldino (2020) provides guidance on specific steps that researchers can follow to translate an imprecise verbal theory into a more precise formal model, and Jackson et al. (2017) provide an excellent introduction—including examples, tips, and a step-by-step guide for beginners—to one particular modeling methodology (agent-based models) and its applications within social psychology. For the future modelers interested in cognitive psychology, Wilson and Collins (2019) provide an excellent step-by-step introduction. And although some computational-modeling articles are highly technical and difficult for beginners to decode, there also exist more approachable examples that, with a minimum of technical detail, illustrate how computational-modeling methods can aid theory building in domains of "soft" psychology—such as personality, social, and cultural psychology (e.g., Muthukrishna & Schaller, 2020; Smaldino et al., 2015; Smaldino, Lukaszewski, von Rueden, & Gurven et al., 2019).

Conclusion

Psychology is a hyperempirical science, but this was not always the case. Early textbooks in the field, such as James (1892/1961) and McDougall (1908), were strong on ideas but weak on evidentiary support for them. The early history of psychology is full of bold theoretical models proposed to explain perception (Fechner, 1987; Helmholtz, 1866/1962), learning (Skinner, 1965; Thorndike, 1933; Tolman, 1925), and personality (Freud, 1920/1959). In the intervening decades, the field was much more focused on creating a large body of empirical results. However, the increased empirical focus of psychology was rarely paralleled with theoretical progress. Rather, many fields of psychology experienced a proliferation of weak and increasingly narrow theories that are difficult to refute or corroborate. This is especially true for theories in "soft psychology." Decades ago, Meehl (1978) observed that theories in clinical, social, and personality psychology were "scientifically unimpressive and technologically worthless" (p. 806). However, these are not the only areas of psychology that are

experiencing such issues. Similar concerns about the weakness of theories have been raised in cognitive psychology (Gigerenzer, 1998; Hommel, 2020).

Meehl (1978) suggested that weak theories lead to a myopic overreliance on the empirical tools that psychological scientists use to analyze and draw inferences from empirical data. Instead of engaging in theory testing, psychologists seek to establish the statistical significance of effects, and this practice does not lead to cumulative knowledge. These concerns have been echoed many times since (e.g., Cohen, 1994; Gigerenzer, 1998; Greenwald, 2012; Higgins, 1992), most recently in the wake of the "replication crisis" (e.g., Fiedler, 2017; Gray, 2017; Klein, 2014; Morey et al., 2018; Muthukrishna & Henrich, 2019; Oberauer & Lewandowsky, 2019; Schaller, 2016). Despite the recognition of the poor state of theory in many fields of psychology, and the problems associated with such a state, it remains unclear what steps should be taken to improve psychological theories.

Here we have argued that there is a need for more clearly articulated methods that can be used to build better theories with increased explanatory and predictive power. Despite the poor overall state of theory in psychology, certain subfields have a long tradition of research focused on theory development and testing. We have described computational modeling and construct validation—two systematic methods for theory development. Both methods include a close relationship between theory development and theory testing and represent dynamic workflows through which a theory is iteratively refined on the basis of empirical data. We have shown how they can complement each other and be combined to develop strong psychological theories. Only when combined can these methods produce process models that include precisely specified psychological constructs and clear algorithms through which they interact to give rise to psychological phenomena. These methods represent a set of teachable tools that can produce theories that offer better explanations, make more precise predictions, and more fully satisfy the standards of scientific practice.

Transparency

Action Editors: Travis Proulx and Richard Morey
Advisory Editor: Richard Lucas
Editor: Laura A. King
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
Funding
This work was supported by the Special Research Fund (BOF) of Ghent University (Grant 01D02415 to I. Grahek) and by the Social Sciences and Humanities Research Council of Canada (Insight Development Grant 430-2018-00218 to M. Schaller).

ORCID iDs

Ivan Grahek https://orcid.org/0000-0002-9247-5956
Mark Schaller https://orcid.org/0000-0003-1243-4650

Notes

1. Evaluation of theory strength is a matter of debate among philosophers of science. Although there is no agreed-on set of criteria, there is broad agreement that the quality of explanations and predictions offered by a theory are crucial for evaluating the strength of a theory (Fidler et al., 2018; Gawronski & Bodenhausen, 2015).
2. This narrow prediction range does not necessarily have to be numerical. Predicting a rank order, a shape of a function, or a nuanced set of multiple orthogonal directional predictions are all ways to constrain the range of predicted results (Meehl, 1990).

References

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Biological and Biomedical Sciences*, *36*, 421–441. https://doi.org/10.1016/j.shpsc.2005.03.010

Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science*, *41*, 321–333. https://doi.org/10.1016/j.shpsa.2010.07.003

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L., Camerer, C., Cesarini, D., Chambers, C. D., Clyde, M., Cook, T. D., De Boeck, P., Dienes, Z., Dreber, A., Easwaran, K., Efferson, C., . . . Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*, 6–10. https://doi.org/10.1038/s41562-017-0189-z

Brandes, C. M., Reardon, K. W., Shields, A. N., & Tackett, J. L. (2021). Towards construct validity of relational aggression: An examination of the Children's Social Behavior Scale. *Psychological Assessment*. Advance online publication. https://doi.org/10.1037/pas0001005

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, *31*, 1412–1427. https://doi.org/10.1037/pas0000626

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*, 997–1003. https://doi.org/10.1037/0003-066X.49.12.997

Craver, C. F. (2006). When mechanistic models explain. *Synthese*, *153*, 355–376. https://doi.org/10.1007/s11229-006-9097-x

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302. https://doi.org/10.4324/9781315128498

Cummins, R. (2000). "How does it work?" versus "What are the laws?": Two conceptions of psychological explanation. In R. A. Wilson & F. C. Keil (Eds.), *Explanation and Cognition* (pp. 117–144). MIT Press.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*, 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

DeYoung, C. G. (2015). Cybernetic big five theory. *Journal of Research in Personality*, *56*, 33–58. https://doi.org/10.1016/j.jrp.2014.07.004

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*, 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as representation of momentum. *Trends in Cognitive Sciences*, *20*, 15–24. https://doi.org/10.1016/j.tics.2015.07.010

Farell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge University Press.

Fechner, G. T. (1987). My own viewpoint on mental measurement (1887). *Psychological Research*, *49*, 213–219. https://doi.org/10.1007/BF00309029

Fidler, F., Thorn, F. S., Barnett, A., Kambouris, S., & Kruger, A. (2018). The epistemic importance of establishing the absence of an effect. *Advances in Methods and Practices in Psychological Science*, *1*, 237–244. https://doi.org/10.1177/2515245918770407

Fiedler, K. (2017). What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspectives on Psychological Science*, *12*, 46–61. https://doi.org/10.1177/1745691616654458

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*, 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Fletcher, J. A., & Doebeli, M. (2009). A simple and general explanation for the evolution of altruism. *Proceedings of the Royal Society B: Biological Sciences*, *276*, 13–19. https://doi.org/10.1098/rspb.2008.0829

Forstmann, B. U., & Wagenmakers, E. J. (2015). Model-based cognitive neuroscience: A conceptual introduction. In B. U. Forstmann & E. J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 139–156). Springer.

Freud, S. (1959). *Beyond the pleasure principle* (J. Strachey, Trans.). Bantam. (Original work published 1920)

Gawronski, B., & Bodenhausen, G. V. (2015). Theory evaluation. In B. Gawronski & G. V. Bodenhausen (Eds.), *Theory and explanation in social psychology* (pp. 3–23). Guilford Press.

Gigerenzer, G. (1998). Surrogates for theories. *Theory & Psychology*, *8*, 183–194.

Grahek, I., Musslick, S., & Shenhav, A. (2020). A computational perspective on the roles of affect in cognitive control. *International Journal of Psychophysiology*, *151*, 25–34. https://doi.org/10.1016/j.ijpsycho.2020.02.001

Grahek, I., Shenhav, A., Musslick, S., Krebs, R. M., & Koster, E. H. W. (2019). Motivation and cognitive control in depression. *Neuroscience and Biobehavioral Reviews*, *102*, 371–381. https://doi.org/10.1016/j.neubiorev.2019.04.011

Gray, K. (2017). How to map theory: Reliable methods are fruitless without rigorous theory. *Perspectives on Psychological Science*, *12*, 731–741. https://doi.org/10.1177/1745691617691949

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science*, *7*, 99–108. https://doi.org/10.1177/1745691611434210

Hammond, R. A., & Axelrod, R. (2006). Evolution of contingent altruism when cooperation is expensive. *Theoretical Population Biology*, *69*, 333–338. https://doi.org/10.1016/j.tpb.2005.12.002

Heathcote, A., Brown, S. D., & Wagenmakers, E. J. (2015). An introduction to good practices in cognitive modeling. In E. J. Wagenmakers & B. U. Forstmann (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 25–48). Springer.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*, 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Helmholtz, H. (1962). *Concerning the perceptions in general, treatise on physiological optics* (J. P. C. Southall, Trans.). Dover. (Original work published 1866)

Higgins, E. T. (1992). Increasingly complex but less interesting articles: Scientific progress or regulatory problem? *Personality and Social Psychology Bulletin*, *18*, 489–492.

Hommel, B. (2020). Pseudo-mechanistic explanations in psychology and cognitive neuroscience. *Topics in Cognitive Science*, *12*, 1294–1305. https://doi.org/10.1111/tops.12448

Jackson, J. C., Rand, D., Lewis, K., Norton, M. I., & Gray, K. (2017). Agent-based modeling: A guide for social psychologists. *Social Psychological and Personality Science*, *8*, 387–395.

James, W. (1961). *Psychology: The briefer course*. Harper. (Original work published 1892)

Klatzky, R. L., & Creswell, J. D. (2014). An intersensory interaction account of priming effects—and their absence. *Perspectives on Psychological Science*, *9*, 49–58. https://doi.org/10.1177/1745691613513468

Klein, S. B. (2014). What can recent replication failures tell us about the theoretical commitments of psychology? *Theory & Psychology*, *24*, 326–338.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., . . . Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*, 454–477. https://doi.org/10.1037/abn0000258

Kruglanski, A. W. (2001). That "vision thing": The state of theory in social and personality psychology at the edge of the new millennium. *Journal of Personality and Social Psychology*, *80*, 871–875. https://doi.org/10.1037/0022-3514.80.6.871

Lakatos, I. (1976). Falsification and the methodology of scientific research programmes. In S. G. Harding (Ed.), *Can theories be refuted* (pp. 205–259). Springer.

Lee, M. D., Criss, A. H., Devezer, B., Donkin, C., Etz, A., Leite, F. P., Matzke, D., Rouder, J. N., Trueblood, J. S., White, C. N., & Vandekerckhove, J. (2019). Robust modeling in cognitive science. *Computational Brain & Behavior*, *2*, 141–153. https://doi.org/10.1007/s42113-019-00029-y

Marr, D. (1982). The philosophy and the approach. In *Vision: A computational investigation into the human representation and processing of visual information* (pp. 8–38). Henry Holt and Co.

Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London B: Biological Sciences*, *204*, 301–328. https://doi.org/10.1016/B978-1-4832-1446-7.50046-7

Maul, A. (2017). Rethinking traditional methods of survey validation. *Measurement*, *15*, 51–69. https://doi.org/10.1080/15366367.2017.1348108

McDougall, W. (1908). *An introduction to social psychology*. Methuen.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, *46*, 806–834.

Meehl, P. E. (1986). What social scientists don't understand. In D. W. Fiske & R. A. Shweder (Eds.), *Metatheory in social science: Pluralisms and subjectivities* (pp. 315–338). University of Chicago Press.

Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, *1*, 108–141.

Morey, R. D., Homer, S., & Proulx, T. (2018). Beyond statistics: Accepting the null hypothesis in mature sciences. *Advances in Methods and Practices in Psychological Science*, *1*, 245–258. https://doi.org/10.1177/2515245918776023

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behavior*, *3*(3), 221–229. https://doi.org/10.1038/s41562-018-0522-1

Muthukrishna, M., & Schaller, M. (2020). Are collectivistic cultures more prone to rapid transformation? Computational models of cross-cultural differences, social network structure, dynamic social influence, and cultural change. *Personality and Social Psychology Review*, *24*, 103–120.

Nesse, R. M. (2013). Tinbergen's four questions, organized: A response to Bateson and Laland. *Trends in Ecology & Evolution*, *28*, 681–682.

Nosek, B. A., Ebersole, C. R., Dehaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences, USA*, *115*, 2600–2606. https://doi.org/10.1073/pnas.1708274114

Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, *26*, 1596–1618.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *349*(6251), Article aac4716. https://doi.org/10.1126/science.aac4716

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, *88*, 552–572. https://doi.org/10.1037/0033-295X.88.6.552

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). Appleton-Century-Crofts.

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin and Review*, *26*, 452–467. https://doi.org/10.3758/s13423-018-1558-y

Rouhani, N., & Niv, Y. (2019). Depressive symptoms bias the prediction-error enhancement of memory towards negative events in reinforcement learning. *Psychopharmacology*, *236*, 2425–2435. https://doi.org/10.1007/s00213-019-05322-z

Schaller, M. (2016). The empirical benefits of conceptual rigor: Systematic articulation of conceptual hypotheses can reduce the risk of non-replicable results (and facilitate novel discoveries too). *Journal of Experimental Social Psychology*, *66*, 107–115. https://doi.org/10.1016/j.jesp.2015.09.006

Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., Van Cleve, J., & Yeh, D. J. (2014). Not just a theory—The utility of mathematical models in evolutionary biology. *PLOS Biology*, *12*(12), Article e1002017. https://doi.org/10.1371/journal.pbio.1002017

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366. https://doi.org/10.1177/0956797611417632

Skinner, B. F. (1965). *Science and human behavior*. Simon and Schuster.

Smaldino, P. E. (2017). Models are stupid, and we need more of them. In R. Vallacher, S. Read, & A. Nowak (Eds.), *Computational social psychology* (pp. 311–331). Routledge. https://doi.org/10.4324/9781315173726-14

Smaldino, P. E. (2020). How to translate a verbal theory into a formal model. *Social Psychology*, *51*, 207–218.

Smaldino, P. E., Calanchini, J., & Pickett, C. L. (2015). Theory development with agent-based models. *Organizational Psychology Review*, *5*, 300–317.

Smaldino, P. E., Lukaszewski, A., von Rueden, C., & Gurven, M. (2019). Niche diversity can explain cross-cultural differences in personality structure. *Nature Human Behaviour*, *3*(12), 1276–1283.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, *11*(1), 87–104.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*, 135–170. https://doi.org/10.1037/0033-295X.88.2.135

Tay, L., & Jebb, A. T. (2018). Establishing construct continua in construct validation: The process of continuum specification. *Advances in Methods and Practices in Psychological Science*, *1*, 375–388.https://doi.org/10.1177/2515245918775707

Thorndike, E. (1933). A proof of the law of effect. *Science*, *77*, 173–175.

Tolman, E. C. (1925). Purpose and cognition: The determiners of animal learning. *Psychological Review*, *32*, 285–297. https://doi.org/10.1037/h0072784

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, *8*, Article e49547. https://doi.org/10.7554/eLife.49547