Developmental Psychology

Computational Modeling Approaches to Emotional Development

Andrea G. Stein and Seth D. Pollak

Online First Publication, September 26, 2024. https://dx.doi.org/10.1037/dev0001830

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

https://doi.org/10.1037/dev0001830

Computational Modeling Approaches to Emotional Development

Andrea G. Stein and Seth D. Pollak
Department of Psychology and Waisman Center, University of Wisconsin–Madison

Computational models of development have the potential to address a key challenge in emotional development research: investigating not only what changes across development but also how these changes come about. Drawing on connectionist and Bayesian methods, this review considers how computational modeling could augment the processes of theorizing and behavioral research to investigate causal processes underlying emotional development. As an illustrative example, we consider how different modeling approaches could help researchers evaluate different ideas about how children come to reason about others' emotions in increasingly sophisticated ways across development. This example is just a starting point; we propose that computational modeling could be an invaluable tool for exploring a variety of yet unresolved "how" questions in emotional development.

Public Significance Statement
This article demonstrates how computational modeling may be used to shed new light on the processes underlying changes in children's emotional development. It provides a brief introduction to computational models of development for researchers who are unfamiliar with this approach but interested in its potential to help us understand how changes in emotion come about.

Keywords: emotion, emotional development, computational modeling

Across decades of investigations into emotional development, researchers have documented the many changes that occur on the path from birth to the complexity observed later in development (e.g., Lapate & Shackman, 2018; Ruba & Pollak, 2020). For example, we know that children's emotion reasoning—the process by which they make inferences and predictions about others' emotions (Ruba & Pollak, 2020)—becomes increasingly sophisticated across development, as they come to take into account new factors such as individuals' desires (Repacholi & Gopnik, 1997; Ruffman et al., 2018), past experiences (Lagattuta et al., 1997; Lagattuta & Wellman, 2001), and expectations (Asaba et al., 2019; Lara et al., 2019) in their reasoning. However, we know little about how these new emotion reasoning tendencies emerge (see Buss et al., 2019; Pollak et al., 2019). Here, we propose that computational models designed to simulate how children develop offer a valuable and currently underutilized tool for beginning to answer these sorts of "how" questions in emotional development.

In particular, we consider how computational modeling might play a critical role in answering "how" questions in emotional development by acting as a bridge between researchers' ideas about the origins of developmental change and empirical research related to these ideas. (For similar ideas about how modeling may fit into psychological theory building across research areas, see Borsboom et al., 2021; Guest & Martin, 2021; van Rooij & Baggio, 2021.) Broadly speaking, we propose that computational modeling can connect theory and behavioral research by fulfilling three key roles: helping researchers elaborate on their ideas about the origins of developmental change, allowing them to test the plausibility of these ideas, and providing a tool for them to generate new, falsifiable hypotheses consistent with these ideas.

First, the exercise of formalizing a model often enables the researcher to elaborate upon their initial proposed explanation for a given developmental change. Consider two researchers seeking to explain how children come to take into account individuals' prior expectations when reasoning about their emotional responses to events. Researcher A might posit that the emergence of such expectations-based reasoning is strictly learning-based—perhaps the result of inductive learning. Researcher B, meanwhile, might posit that this same developmental change is instead driven largely by cognitive maturation—such as maturation of the child's long-term

Amanda E. Guyer served as action editor.
Andrea G. Stein https://orcid.org/0000-0002-1792-6606
Seth D. Pollak https://orcid.org/0000-0001-5184-9846
Andrea G. Stein was supported by an Emotion Research Training Grant T32MH018931 and Seth D. Pollak was supported by Grant R01MH61285, both from the National Institute of Mental Health. The authors appreciate the helpful feedback that Karen Smith, Ashley Ruba, and Joe Austerweil provided on an earlier draft of this article.

Andrea G. Stein played a lead role in writing–original draft and an equal role in conceptualization and writing–review and editing. Seth D. Pollak played a supporting role in writing–original draft and an equal role in conceptualization and writing–review and editing.
Correspondence concerning this article should be addressed to Andrea G. Stein, Department of Psychology and Waisman Center, University of Wisconsin–Madison, 1500 Highland Avenue, Room 399, Madison, WI 53705, United States. Email: agstein@wisc.edu

memory. Developmental explanations such as these often suffer from underspecification. They may propose a broad mechanism that accounts for the change of interest (e.g., a particular variety of learning or the development of a particular cognitive skill), but they rarely elucidate how, in a step-by-step manner, that mechanism might produce the given change (Benton, 2023). As we will illustrate, computational modeling gives the researcher a tool to do just that.

Second, computational modeling allows researchers to test the plausibility of their proposed explanations for developmental change. Such ideas are often difficult to evaluate directly via behavioral studies because they implicate factors that the researcher cannot manipulate. Researchers cannot control the maturation of a child's memory or all the instances of emotion to which a child is exposed in their day-to-day life. Computational modeling, on the other hand, allows the researcher to manipulate these variables. We will discuss how more precisely researchers can go about this, through their specification of a model and a learning environment, in the next section. For now, suffice it to say that computational modeling allows researchers to test whether—and under what conditions—they can reproduce the developmental change of interest. For example, when Researcher B specifies a model whose cognitive capacity increases across development, do these increases contribute to the emergence of expectations-based emotion reasoning? Alternatively, when Researcher A specifies a model that learns via inductive inference, does this same expectations-based emotion reasoning emerge with accumulating experience? Through such simulations, computational modeling can provide an initial test of the plausibility of one's ideas about the origins of developmental change.

Third, once a researcher has developed a plausible model, this model can be used to generate new hypotheses that are more directly falsifiable via empirical studies. For instance, as we will discuss, having established a plausible model of the emergence of expectations-based emotion reasoning, a researcher might then test the effects of different learning environments on the model's learning and use the results as the basis for new hypotheses about the impact of individual differences in children's environments on the pattern or timing of developmental change. These hypotheses could then be tested via behavioral studies, the results of which could inform modifications (or falsification) of the original model and idea (see Stojnić et al., 2023). Under this approach, computational modeling accordingly becomes a key component in a cycle of theorizing, modeling, and empirical behavioral research, all targeted at furthering our understanding of the origins of emotional development.

In the sections that follow, we first provide a brief primer on the key undertaking of computational modeling: specifying a learner and a learning environment consistent with one's ideas about development. As an illustrative example, we then explore how researchers holding two very different ideas about the origins of developmental changes in children's emotion reasoning might use computational modeling to evaluate their theories. We conclude with a discussion of what we see as the most important ongoing challenges facing researchers seeking to apply computational modeling to their investigations of emotional development.

Building Blocks of Computational Modeling

The most basic building blocks of computational modeling mirror the building blocks of development: There must be both a learner and experiences from which that learner can learn. In computational modeling, these conditions are met by specifying a model and a training environment. There are a number of excellent reviews and tutorials covering the nuts and bolts of this specification process in greater detail (e.g., Nobandegani & Shultz, 2022; Perfors et al., 2011; Yermolayeva & Rakison, 2014). Here, we focus on broader theoretical considerations relevant to researchers starting to consider how they might usefully apply computational modeling methods to questions in emotional development.

The Learner

To formalize a computational model, the researcher must first choose what general form their model will take. This article focuses on the two types of computational models that are currently most widely used to explain developmental change: Bayesian models and connectionist models. Each type of model can and has been applied across diverse domains of development, to investigate the potential contributions of both learning and maturation to development. However, as we detail here, the two types of models differ in that they rely on different representations of human thought and learning.

In Bayesian models, thought is represented via sets of hypotheses that the learner considers about a set of phenomena in the world. This hypothesis space may be relatively constrained or, particularly in the case of Bayesian nonparametric models, relatively flexible (for a review, see Austerweil et al., 2015). For example, a child might consider different hypotheses about the factors that determine someone's emotional responses to an event. Each hypothesis is assigned a probability indexing its perceived probability. Bayesian models capture the relations involved in such hypotheses via a variety of structures, such as graphical networks (essentially, flowcharts) or probabilistic programs (for more on probabilistic programming models, see Ghahramani, 2015; Goodman, 2013).

What distinguishes Bayesian modeling from other approaches is that learning consists of the learner updating the probability they assign to each hypothesis after new observations in accordance with Bayes' Rule. Updated probabilities, called posterior probabilities, reflect two factors. First, a posterior probability reflects a hypothesis's prior probability—that is, how much the learner believed the hypothesis before observing any new information. Second, it reflects the hypothesis's likelihood, which captures how well the hypothesis fits the newly observed data. Specifically, it represents the probability of observing the new data if the hypothesis in question were true. This updating process occurs repeatedly, meaning that as experience accumulates, more weight is given to the fit of each hypothesis to this experience. In this way, learning within the Bayesian framework thus balances both the learner's prior knowledge and the learner's new experiences when evaluating potential hypotheses about how the world works.

In connectionist models, thought is represented via patterns of activation on a collection of information processors, referred to as units, designed to roughly parallel the function of neurons in the brain. In the feedforward connectionist models discussed here, the broad role of the model is to receive a set of observations on a set of input units. Then, via weighted connections among units, the model transmits and transforms these activations to ultimately produce a prediction of some kind. For example, a model might be designed to receive observations related to a particular event and produce a prediction regarding the emotional response likely to follow. To make such predictions, connections between input and output units in the model are often mediated by layers of "hidden" units that allow for internal representations to be formed and refined over the course of learning. The simplest connectionist models may have one or even no hidden layers, while so-called deep neural network models may have hundreds of hidden layers, allowing for highly complex internal representations (for more on deep neural networks, see LeCun et al., 2015; Schmidhuber, 2015).

In these models, learning is implemented by an algorithm that updates connection weights among units in the model, paralleling changes in the learner's representations. In the most commonly used algorithms for feedforward connectionist models, backpropagation-based algorithms, these adjustments are prediction error-based. That is, adjustments are made based on the discrepancies between a model's output and a researcher-provided "correct" response. This process parallels, for example, a learner predicting an individual's emotional response to an event, observing their actual response, and comparing the two. These discrepancies are then propagated backward through the model to adjust connection weights, effectively changing the learner's internal representations to be more in line with what was observed. In this way, backpropagation-based connectionist models allow the learner to use prediction error to gradually adjust their understanding of how the world works.

There has been extensive debate regarding the advantages and disadvantages of Bayesian and connectionist modeling and about which approach most veridically captures human cognition in learning (e.g., Griffiths et al., 2010; McClelland et al., 2010). In our view, it would be shortsighted to claim that one framework or the other is preferable for excavating the causal processes underlying emotional development. Rather, if both approaches can plausibly simulate developmental phenomena of interest, both may provide important insights into how developmental change occurs and allow for the refinement of future hypotheses and models. Accordingly, we do not advocate for any one approach. Rather, we recommend researchers carefully consider how each approach implements cognition and learning as described above and choose the method most consistent with their own perspective on these topics.

Finally, while so far we have discussed how Bayesian and connectionist models implement learning, it is important to note that both these approaches may also be used to investigate the contributions of maturation to developmental change. Often, this simulation is achieved through a coarse model comparison approach, in which researchers compare the performance of models that vary on a set of parameters proposed to correspond to a particular capacity that matures across development. Researchers have used this approach to examine the influence of developmental differences in working memory on infants' changing representations of animates and inanimates (Rakison & Lupyan, 2008) and the influence of shrinking neural receptive fields on infants' category learning abilities (Westermann & Mareschal, 2012). Some approaches incorporate such changes continuously in a single model. For example, constructive neural network models deploy algorithms that allow change to take place not only through changes to the connection weights between units but also through the recruitment of additional hidden units, in a process that has been broadly understood as a proxy for synaptogenesis and/or neurogenesis in the human brain (Shultz, 2017; Westermann et al., 2006). Through such methods, computational modeling may be used to examine not only how learning contributes to development but also how development is shaped by brain maturation.

The Learning Environment

Like a child, in order to learn, a model must be exposed to new experiences. For example, in the case of a model designed to learn to predict individuals' emotional responses to events, these learning experiences might be a series of pairings between life events and individuals' emotional responses. Here, the researcher will have to make a series of important decisions about how they break these observations down into their more primitive features. In our example, for instance, the researcher must decide what aspects or dimensions of events will be included as input, as well as how they will represent individuals' emotional responses (e.g., will they represent them in terms of discrete categories, dimensions, raw perceptual features, or something else?). Plainly, this process will draw on researchers' ideas about the nature of emotions and other theoretical commitments.

Having determined these primitive features, the researcher must also develop the specific set of instances the model will learn from. In the modeling applications we focus on here, the goal is to develop a set of training exemplars that is as consistent as possible with what we know about children's day-to-day experiences with emotions—for instance, how commonly they encounter different emotions, how perceptually variable these instances of emotion are, how consistent event–emotion pairings are, and so on. Of course, to anticipate one major challenge that we will return to later, our knowledge in these areas is often quite limited! To mitigate this, and to gain a better understanding of environmental constraints on the behavioral change of interest, researchers will often develop multiple training sets and examine for which training set(s) the change of interest is reproduced, effectively establishing some boundary conditions for their results.

In the remainder of this article, we explore ways in which computational modeling might help address unanswered "how" questions in emotional development. As an illustrative example, we explore how computational models might be used to evaluate the plausibility of two different ideas about how children's emotion reasoning becomes more sophisticated across development. Our intent with this example is not to advocate for any one idea about the origins of emotional development over any other, nor to endorse the use of any one modeling approach. Rather, our goal is to provide a broader illustration of the potential of these approaches to help researchers elaborate on their ideas about the causes of developmental change, evaluate the initial plausibility of these ideas, and generate testable hypotheses.

An Illustrative Example: How Does Children's Emotion Reasoning Become More Sophisticated?

As adults, we consider many different factors when reasoning about others' emotions. For example, imagine predicting another person's reaction to losing a board game. When predicting their reaction, we are likely to take into account their desires. In doing so, we would understand that a competitive person might be especially upset to lose, whereas a person with little interest in games might not react to the loss at all. Meanwhile, a person who was trying to let someone else win might derive pleasure from losing. We might also consider the person's past experiences. A younger sibling who nearly always loses to an older sibling might react more strongly to another loss, as might someone who has already had a rough day. We would also think about our own expectations of that person—is this someone who tends to respond strongly to negative events, or someone who tends to be very even-keeled or upbeat? We would likely factor in the individual's prior expectations, with the understanding that someone with high expectations about their performance would likely react differently from someone who never expected to win in the first place, who would be less likely to be disappointed. And of course, there would be the social context: Is the game or relationship between the players more important than it might seem at first glance? Are all the players adults, family members, friends, or strangers?

This sort of sophisticated, multivariate emotion reasoning appears to emerge gradually over the course of development. Behavioral evidence suggests that infants begin to demonstrate an understanding of the role of desires in others' emotions somewhere around one and a half (Repacholi & Gopnik, 1997) to two (Ruffman et al., 2018) years old. It is not until around age 5 (Asaba et al., 2019) to 6 (Lara et al., 2019) that children demonstrate an emergent understanding that people's emotional reactions will depend on their prior expectations surrounding an event. And the understanding that people's present emotions depend on their past experiences and life histories appears to emerge between ages 3 and 7 (Lagattuta, 2014).

These behavioral studies can tell us a great deal about the time course and pattern by which children's emotion reasoning changes, as they come to incorporate new factors into their reasoning. An open question remains, however, regarding what drives these changes. That is, what leads infants to start to consider people's desires, or children to begin to consider people's expectations and past experiences, when reasoning about others' emotions?

Sample Testable Hypothesis 1: Experience Drives Changes in Emotion Reasoning

One potential explanation for changes in infants' and children's emotion reasoning is that these changes are largely the product of their accumulating experiences with the world. Specifically, it is possible that very early in development, when infants have limited life experience, it is rational for them to have simpler ideas about how people's emotions work. But as they gain experience, perhaps these ideas are disconfirmed, and they move toward more sophisticated reasoning about others' emotions.

This proposal about the role of experience in transitions from simpler ideas to more sophisticated ones draws from research in other domains on children's intuitive theory development (e.g., Gerstenberg & Tenenbaum, 2017; Gopnik & Wellman, 2012). Researchers who study intuitive theory development posit that infants and children have implicit theories about how various phenomena in the world work. For example, infants and children might have different implicit theories about how others' emotional responses work, which might vary with regard to the role that variables like people's desires, past experiences, and expectations play in generating their emotional responses (e.g., Wu et al., 2021). The question then becomes: Why do children appear to shift from favoring simpler models of people's emotions to favoring increasingly more complex ones (e.g., Asaba et al., 2019; Lara et al., 2019; Repacholi & Gopnik, 1997; Ruffman et al., 2018)?

One way researchers have investigated similar questions is through a Bayesian model comparison approach. This approach involves setting up competing Bayesian cognitive models, typically in the form of graphical networks, each of which represents a different intuitive theory a child might have of how the phenomenon of interest works. Researchers then compare the posterior probability of these models over time, as the models are exposed to new observations—just as children are exposed to new experiences across development. A shift in posterior probability from favoring one model to another suggests a shift from favoring one theory to another.

This approach has been fruitful in explaining developmental transitions in reasoning in a number of nonemotional domains (for a review, see Gopnik & Bonawitz, 2015). Goodman et al. (2006) used this approach to demonstrate how, with accumulating experience, Bayesian learning could account for children's shift from an omniscient theory of mind (i.e., a model of the world in which individuals' beliefs are always set to the true state of the world) to a representational theory of mind (i.e., a model of the world in which individuals' beliefs may differ from the true state of the world). Similarly, Lucas et al. (2014) used this approach to demonstrate how Bayesian learning could account for infants' shift from an understanding of preferences as universal (i.e., a model of the world in which preferences are fixed to be constant across individuals) to an understanding of preferences as subjective (i.e., a model of the world in which preferences are free to vary across individuals). As we will discuss later, this issue of variability across individuals represents the major learning problem in understanding emotion.

In both of the studies described above, the developmental transition of interest was explained by the same phenomenon: the changing trade-off between the simplicity of a model and the goodness of fit of that model to new observations of the world. This changing trade-off led to a pattern in which, early on, the simpler model (i.e., the model with fewer free parameters) had the higher posterior probability. However, with accumulating experience—particularly experience with observations that disconfirmed this simpler model—favor shifted to the more complex model.

In human terms, these studies suggest that very early in development, we may tend to hold simpler internal models of how the world works; this simplicity is rational when we have limited life experience. When it comes to emotion reasoning, perhaps very early in development, infants start out predicting others' emotional reactions to events purely based on their own appraisal of those events, without taking into account the other individual's desires, expectations, or past experiences. This simple model of others' emotions might produce good predictions in many situations. Furthermore, with limited experience, any exceptions might initially be hard to distinguish from random noise.

However, with increasing life experience, these exceptions may accumulate, making it easier for the developing child to pick up on the systematic patterns among them. For example, by 18–24 months (Repacholi & Gopnik, 1997; Ruffman et al., 2018), an infant might have noticed that when their sibling loses a game of Candyland, they tend to be quite unhappy, but when their parent loses a game of Candyland, they seem perfectly content. In other words, from these and other similar patterns, they might begin to reason that someone else's emotional reactions may depend on their particular desires. Likewise, perhaps by age 5 to 6 (Asaba et al., 2019; Lara et al., 2019), a child might have noticed that when their sibling who considers themselves a checkers whiz loses a game, they tend to be particularly disappointed, while when their friend who just learned to play loses a game, they do not take it as hard. Again, from there, and from other similar patterns, they might begin to reason that others' emotional reactions to events may depend on their expectations regarding that event. These are trivial examples, but speak to a broader idea: As children gain life experience, they likely gain more and more experience with systematic patterns of behavior that disconfirm their simpler theories of others' emotions; the accumulation of such patterns of experience may lead to an eventual rejection of simpler theories of emotion in favor of more complex, but well-fitting, theories of others' emotions.

Understanding How Experience Might Drive Change

Computational modeling also provides us with a framework for understanding why, more precisely, such a process might account for transitions in children's emotion reasoning. For instance, why might infants initially prefer simpler models of others' emotions? In a Bayesian framework, this preference is rational because, all else equal, simpler models have higher posterior probabilities. Specifically, in a phenomenon that is often referred to as the Bayesian Occam's razor (Jefferys & Berger, 1992; MacKay, 2003), a preference for simplicity will be automatically encoded in the likelihood of each model. That is, if a learner encounters a set of observations that are equally compatible with multiple models of the world, these observations will have a higher likelihood under the simplest models. Why? Recall that the likelihood of a given model specifically encodes the probability of encountering a particular set of observations if the model in question were true. In the simplest model, the set of observations the learner actually observed was one of fewer sets of observations that were ever possible; thus, this particular set of observations had a higher likelihood of being observed. In intuitive terms, the Bayesian Occam's razor captures our sense of coincidence. For example, it would be quite a coincidence to observe a number of people who are all unhappy after losing a game if the desire to win were entirely subjective, but less of a coincidence to observe the same pattern if the desire to win were universal.

Critically, however, this rational bias in favor of simplicity is not the only factor determining a model's likelihood. As we have experiences with the world that conflict with a model, this conflict will also be reflected in the likelihood of the model. For example, as we start to observe that our sibling is consistently unhappy when losing a game, but our parents often do not seem to mind when they lose, our model that represents the desire to win as universal starts to seem quite unlikely.

In sum, researchers have demonstrated that developmental shifts such as the shift from an understanding of preferences as universal to an understanding of preferences as subjective (Lucas et al., 2014), or the shift from an omniscient theory of mind to a representational theory of mind (Goodman et al., 2006), may be the product of a rational bias in favor of simpler models of the world that is eventually outweighed, over the course of learning, by the systematic accumulation of experiences that conflict with these simpler models. The same phenomena may underlie transitions in children's emotion reasoning—a possibility that may be evaluated using computational modeling.

A Potential Model of Experience-Driven Changes in Emotion Reasoning

Researchers investigating transitions in children's emotion reasoning using Bayesian modeling might draw inspiration from researchers proposing models of emotion reasoning in adults (de Melo et al., 2014; Houlihan et al., 2023; Ong et al., 2019; Saxe & Houlihan, 2017; Wu et al., 2018). These models provide a formal Bayesian framework for understanding how we might make a variety of emotion-related inferences, such as predicting someone's behavior based on cues to their emotions, inferring what caused someone's emotions, or, most critically to the present example, predicting someone's emotional responses. Researchers have also started testing these models against behavioral data relevant to emotion. For example, Ong et al. (2015) used Bayesian modeling to capture the process by which participants integrated event outcomes and people's facial configurations to infer others' emotions in a simple game setup, while Wu et al. (2018) modeled the process by which participants integrated individuals' facial configurations and their actions to infer their beliefs and desires in an emotion reasoning study. Likewise, Houlihan et al. (2022) modeled how participants reasoned backward from individuals' expressions to the outcome that may have preceded these expressions, and Teo et al. (2022) modeled how participants used individuals' expressions in the context of their exploration of a new technology to reason about both the individuals' mental states and how the technology worked. Investigations of this sort into adults' emotion reasoning represent a rapidly growing area of research.

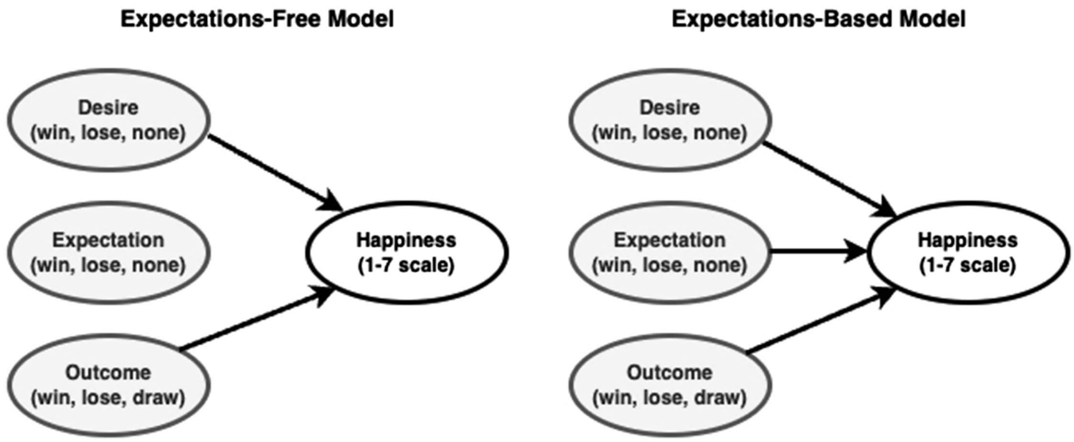
In line with this tradition, models like Goodman et al.'s (2006) model could easily be adapted to investigate the origins of developmental change in emotion reasoning. For example, imagine we sought to model one of the transitions in children's emotion reasoning described earlier: how, around the age of 5 or 6 (Asaba et al., 2019; Lara et al., 2019), children start to take into account individuals' prior expectations when predicting their emotional responses to events. We will refer to this transition as the emergence of expectations-based emotion reasoning.

Specifically, imagine we sought to model the same kind of prediction that children made in Lara et al.'s (2019) study of the emergence of expectations-based emotion reasoning. In our adaptation of the task, the agent is responsible for predicting how happy someone else will be (on a 1–7 scale) following an event (e.g., the outcome of a game). To make this prediction, the agent is provided with information about the target individual's desires surrounding the event (e.g., desiring to win the game, desiring to lose the game, or having no desire regarding the game), their expectations about the outcome (e.g., expecting to win, expecting to lose, or having no expectation), and the actual outcome (e.g., a win, a loss, or a draw).

To model how children might learn to take into account individuals' prior expectations when predicting their emotions, we could set up two competing Bayesian network models (Figure 1). In the simpler, expectations-free model, an individual's emotional response to an event could be modeled as depending on only whether the individual wanted to win the game and whether they did. The more complex, expectations-based model, meanwhile, could be similar, except that an individual's emotional response would also be allowed to depend on whether the individual expected to win the game beforehand.

We could then expose these models to a set of game scenarios, varying in terms of the individual's desires and expectations, the actual outcome of the game, and the individual's ensuing emotional response. In developing these scenarios, we would want to pay particular attention to including a proportion of scenarios violating the expectations-free model that we believed was consistent with

Figure 1
Architecture of Potential Bayesian Models of Emotion Reasoning



Note. In the simpler, expectations-free model of others' emotional responses (left), an individual's happiness in response to an event depends only on the outcome of that event and their desires surrounding that event. In the more complex, expectations-based model (right), that emotional response can also depend on their expectations about the event. Over the course of learning, parameters for each of the causal relationships in each model can be learned, and the posterior probabilities of the two models can be compared.

the proportion of exceptions to this model children are likely to encounter in their day-to-day lives—a key challenge we will return to later.

Once training sets were formed, and the corresponding observations were presented to the model, we could then use Bayes' Rule to compute and compare the posterior probabilities of the two models, to determine which model was favored at different points in development. Specifically, we could examine whether, initially, the expectations-free model of emotion reasoning was favored due to its relative simplicity and whether, with accumulating experience, favor shifted to the expectations-based model, paralleling the emergence of expectations-based emotion reasoning in children. This result would help us understand whether, as children gain life experience—particularly with situations in which individuals' expectations play a key role in their emotional responses, such as in cases of disappointment and relief—a changing trade-off between simplicity and goodness of fit might account for an emerging tendency to consider others' expectations when reasoning about their emotions.

Sample Testable Hypothesis 2: Cognitive Maturation Drives Changes in Emotion Reasoning

Another possibility is that developments in infants' and children's emotion reasoning are the product not only of their accumulating experience with the world but also of the ongoing maturation of their cognitive abilities. This possibility is consistent with recent calls from several researchers (Hoemann et al., 2020; Ruba & Pollak, 2020; Walle et al., 2022) to consider the role that cognitive development may play in emotional development.

Reasoning about another individual's emotions requires a child to direct their attention to relevant cues, draw on their contextually relevant emotion knowledge, and integrate all of this information to make inferences. However, cognitive capacities such as attention, long-term memory, and working memory that underlie these processes undergo protracted development across childhood as the brain continues to mature (Bauer, 2013; Rueda & Posner, 2013; Spencer, 2020). Given the cognitive requisites of emotion reasoning, it is plausible to consider that constraints provided by cognitive development play a role in shaping children's emotional development.

As noted earlier, constructive neural networks are one way of modeling the contributions of cognitive maturation to developmental change in other domains. Constructive neural network models are similar to other connectionist models in that they can learn by adjusting connection weights among units in the model based on prediction error. Unlike the archetypal connectionist model, however, the architecture of units in a constructive neural network model is not fixed. Instead, constructive neural network models are designed to increase in complexity over time (i.e., across development). Specifically, constructive neural networks can incorporate additional hidden units—the units responsible for the learner's internal representations—whenever they have exhausted the reduction in prediction error achievable through learning (i.e., connection weight adjustments). From a neuro-constructivist perspective, this process is generally understood as implementing learning-dependent brain plasticity (Westermann et al., 2006). More broadly, however, constructive neural network models can also instantiate the principle that for certain developmental changes, progress may not be achievable through learning alone, but instead may also rely on the development of infants' and children's cognitive resources.

Constructive neural networks have been used to model a number of transitions in infants' and children's reasoning in nonemotional domains (for a review, see Shultz, 2017). For example, Berthiaume et al. (2013) used constructive neural networks to model transitions in children's theory of mind, including the transition from an omniscient theory of mind to a representational theory of mind. Berthiaume et al. (2013) found that early in training, before incorporating any hidden units, the model made predictions consistent with an omniscient theory of mind. However, once the model increased its representational capacity by recruiting a hidden unit, it learned to make predictions consistent with a representational theory of mind. This result suggested a critical role for changes in children's cognitive resources in their development of a theory of mind.

Understanding How Cognitive Maturation Might Drive Change

As in the previous example, computational modeling allows us not only to establish that the development of children's cognitive resources might account for transitions in their reasoning but also to understand how, more precisely, it might do so. Berthiaume et al.'s (2013) modeling suggested that cognitive development might drive children's transition on the false belief task because said development enables them to consider more complex relations among features of their environment than they were previously capable of considering. One of the researchers' key insights was that the traditional false belief task is what is referred to as a nonlinearly separable classification problem: a problem in which it is impossible to consistently make the correct prediction by only considering each variable in isolation (McClelland & Rumelhart, 1989). (In more precise geometric terms, a nonlinearly separable classification problem is a problem that cannot be perfectly solved by drawing a single line or plane through the n-dimensional space of input value combinations.) To consistently succeed at the false belief task, a child (or a model) must be able to take into account the starting location of the object, the end location of the object, and whether the agent was able to observe any movement from starting to end location. However, it is not sufficient to take into account any (or even all) of these features independently. Instead, for a model or a child to make consistently correct predictions, it must learn an interaction among these features: When the agent is able to observe any movements, the correct prediction depends on the end location of the object, and when the agent does not observe the movements, the correct prediction depends on the starting location of the object.

Importantly, connectionist models without hidden units—that is, models that are in a state of relative cognitive immaturity—cannot solve nonlinearly separable problems like the false belief task. Without hidden units mediating connections between input and output, they have no way of encoding interactions (Minsky & Papert, 1969; Rumelhart et al., 1985). Early in development, when Berthiaume et al.'s (2013) model had no hidden units, the best predictions it could make were predictions consistent with an omniscient theory of mind (i.e., predictions that depended on the end location of the object, in isolation). However, once the model recruited its first hidden unit, this increased representational capacity enabled it to learn the interaction among agent observation, starting location, and ending location described earlier. It acquired a representational theory of mind.

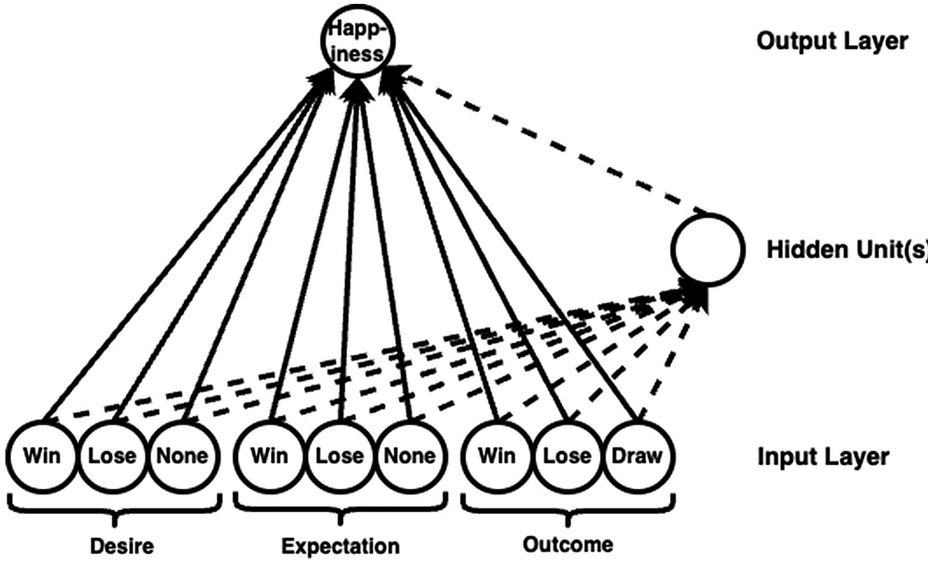
Put in human terms, then, perhaps children's reasoning becomes more sophisticated across development in part because improvements in their cognitive resources allow them to learn new, more complex relations among features of their environments. In the case of emotion reasoning, perhaps maturation only gradually provides infants and children with the capacity to represent the ways in which people's emotional responses are the interactive product of not only what happens in the world but also, among other factors, their desires and expectations surrounding these events. Thus, perhaps it is not just learning, but also cognitive development, that drives the development of children's emotion reasoning.

A Potential Model of Cognitive Maturation-Driven Changes in Emotion Reasoning

Imagine that we wanted to model the same transition in children's emotion reasoning as in the previous example—the emergence of expectations-based reasoning—using the same game scenario and the same learning environment/training set. But this time, we wanted to model how cognitive maturation might contribute to this developmental change. To do so, we could use a constructive neural model similar to those developed by Berthiaume et al. (2013) and others. Specifically, recall that in the aforementioned game scenario, the learner has information about an individual's desires (e.g., desiring to win the game, desiring to lose, or having no particular desire) and expectations (e.g., expecting to win, expecting to lose, or having no expectation) surrounding the game, as well as the actual outcome of the game (win, lose, or draw), and is tasked with predicting the individual's ensuing emotional response (e.g., their level of happiness on a 1–7 scale). In this case, we could specify a constructive neural network model with input units corresponding to each of these potential desires, expectations, and outcomes (Figure 2). For example, one input unit would activate only in scenarios in which the individual desired to win the game, while another would activate only in scenarios in which the individual desired to lose, and so on, for each potential desire, expectation, and outcome. On the output layer, a single unit with a continuous activation function could capture the model's prediction regarding the individual's happiness level, to be compared to their observed level of happiness. In the beginning, the model would have no hidden units to mediate connections between input units and the output unit, mirroring immature representational capacity, but, characteristic of a constructive neural network model, its algorithm would allow it to recruit additional units anytime it could no longer improve its predictions solely via learning from prediction error.

We could then train this constructive neural network model on the same set of game scenarios described in the previous section. Over the course of training (i.e., development), we could then examine how the model's predictions for each scenario changed. Specifically, this examination often involves a technique called output contribution analysis, which allows the researcher to quantify, for each pattern, how much each input and hidden unit contributed to the model's output activation(s)/prediction (Sanger, 1989; Shultz & Elman, 1993). Accordingly, in this case, we would be interested in whether, early in training, the three input units capturing the three different event outcomes had high output contributions (suggesting the model was making predictions based largely on outcomes). We would also be interested in whether, after the model came to recruit additional units, the output contributions of first the input units capturing the individual's desires and then, most critically, the input units capturing their expectations started to increase. This result would help us understand whether increased representational capacity resulting from children's ongoing cognitive development might account for the emergence of a tendency to consider others' expectations when reasoning about their emotions.

Figure 2
Architecture of Potential Constructive Neural Network Model of Emotion Reasoning Development



Note. The model has nine input units: three encoding the individual's desire in a given game scenario, three encoding their expectation, and three encoding the outcome. It has one output unit responsible for producing a prediction regarding the individual's happiness level. Because the model is a constructive neural network, it starts with no hidden units. For each game scenario, one desire unit, one expectation unit, and one outcome unit are activated. That activation feeds forward into the output unit, which is responsible for producing a prediction regarding the individual's happiness level. The activation of this output unit depends on the weights that the model has learned to place on the connections linking it to each activated input unit (solid lines). Per its learning algorithm, the model then compares its prediction to the "correct" output—in this case, the individual's observed level of happiness—and uses this prediction error to adjust its connection weights. But each time the model has exhausted the reductions in prediction error that may be made via learning (i.e., via adjustments to connection weights), it is permitted to recruit an additional hidden unit, which may mediate a connection between any given input unit(s) and the output unit (dotted lines), allowing for increasingly complex internal representations.

Model Evaluation and Hypothesis Generation

Once a researcher has models that appear plausible, the next step is to evaluate said models more thoroughly using behavioral evidence. For example, imagine that both of the aforementioned modeling approaches successfully reproduced, under some set of learning conditions and model specifications, the emergence of children's expectations-based emotion reasoning. Such a result suggests that the model—and by extension, the researcher's idea—provides an initially plausible explanation for how this change in children's emotion reasoning comes about. However, more rigorous testing is needed.

In many computational modeling contexts, researchers can use established quantitative methods to evaluate how well their models fit behavioral phenomena. These methods are generally applicable when researchers are directly modeling behavioral data capturing a phenomenon of interest (e.g., a researcher might seek to model participant behavior in a specific experiment). In these cases, the best practice is typically to first ensure the model is ready for fitting to behavioral data, by using simulated data (with known parameter values) to test that model parameters are recoverable and by confirming that the model is identified. (See Wilson & Collins, 2019, for details on and a workflow for these procedures.) From there, there are generally at least two broad, complementary classes of analyses available to researchers (Palminteri et al., 2017; Wilson & Collins, 2019). The first class consists of simulation methods. The researcher uses the best-fitting version of their model to generate synthetic data, which can then be compared to the observed behavioral data both qualitatively (did the model generally reproduce all of the effects in the observed data?) and quantitatively. The second class of analyses consists of model comparison methods. The researcher fits a set of competing models to the behavioral data and compares their predictive performance using their preferred model comparison criterion. Based on this criterion, the researcher can select a "winning" model.

What these methods share is that they involve direct assessment of the correspondence between one's model(s) and observed behavioral data, generally from experiments. Such direct quantitative assessment is, however, not the only way to evaluate the viability of a model (Kording et al., 2020; Wilson & Collins, 2019). Furthermore, in many cases, such direct quantitative assessment is not immediately available for the kinds of developmental computational models described here, as there is often not the same one-to-one correspondence between model and behavioral data. Developmental computational models may seek to explain phenomena that have been documented in behavioral experiments. For example, here, we sought to explain a developmental difference in expectations-based emotion reasoning documented behaviorally by Lara et al.

(2019) and Asaba et al. (2019). Likewise, developmental computational models may have architectures inspired by the setup of these experiments, as when we specified architectures for both our Bayesian models and our constructive neural network model that were based on Lara et al.'s (2019) game paradigm. However, critically, developmental computational models like the ones we discussed are not aimed at merely capturing participants' behavior in the context of such experiments. Instead, they aim to capture all of the learning and maturational processes that led up to this behavior. In other words, they are modeling more than what the behavioral data is designed to capture.

This distinction, however, does not mean that researchers interested in using computational modeling to understand the processes underlying emotional development have no recourse, beyond initial plausibility testing, when it comes to evaluating their models and the ideas about development they instantiate. On the broadest theoretical level, there are many standards against which these models and ideas may be evaluated such as their parsimony, breadth, and analogy to more established models and theories in other domains (Borsboom et al., 2021; Kording et al., 2020). Ultimately, however, what we advocate for is a process of model evaluation that fits into the cycle of theorizing, modeling, and behavioral research.

That is, one of the key advantages of having a plausible computational model of developmental change is that this model can be used to generate new, empirically testable hypotheses. One of the most straightforward ways to go about this is to manipulate the model's specification and/or the learning environment. The results can then be used to generate new hypotheses about the implications of individual differences in children's cognition or their life experiences. For example, we might explore whether there are learning environments in which one of our proposed models fails to produce the emergence of expectations-based emotion reasoning or produces this transition much earlier or later in training. We suspect (and could simulate) that in both of our proposed models, the time frame of the emergence of expectations-based emotion reasoning depends greatly on the proportion of patterns in the training set involving a violation of expectations. Applying this simulation result to a case of individual differences in child development, we might accordingly hypothesize that unpredictability in a child's environment—and specifically their exposure to others' responses to this unpredictability—might be associated with earlier emergence of expectations-based emotion reasoning. This is an empirically testable hypothesis; we could design a study that investigated this association. The results of this study could then help falsify, or inspire modifications to, one's model and underlying theory. This approach is particularly fruitful, furthermore, in situations where competing models make qualitatively different predictions, as this difference allows researchers to adjudicate between models. For example, if one of our models ended up predicting a positive association between unpredictability and the emergence of expectations-based emotion reasoning, and the other predicted the opposite, a descriptive study of the observed association between these factors could provide valuable evidence in favor of one model over the other. However, regardless of whether one's focus is adjudicating among models or validating a single model, the principle is the same: It is through a cycle of theorizing, model-building, and empirical research that we believe computational modeling can help us investigate the processes underlying emotional developmental change.

Ongoing Challenges

Computational modeling, like any research method, has limitations. First, any model of emotional development will be founded on certain assumptions about the nature of emotion and the nature of knowledge representation and learning. Modeling the development of emotion reasoning, for example, requires commitments about the factors that we might consider when thinking about others' emotions and which factors can be disregarded. Likewise, when we choose a particular type of model, we necessarily make certain commitments about the nature of human learning and knowledge more generally. For example, connectionist models assume our representations are graded and lack explicit structure. However, this may not be the only way that we represent knowledge in the brain, thus limiting the types of learning that can be captured via connectionist models (Griffiths et al., 2010). Bayesian models, meanwhile, make assumptions about how we update our beliefs. In particular, Bayesian modeling depends on our reasoning being "approximately rational" (Griffiths et al., 2008), specifically in a way that approximates the process of Bayesian inference. (Though we need not actually perform Bayesian inference, which would quickly become computationally intensive.) On the one hand, one potential benefit of computational modeling is that, by requiring precise model specifications, rather than just verbal theorizing, the modeling process can help make otherwise implicit assumptions like these explicit (Guest & Martin, 2021). On the other hand, to the extent that the researcher's assumptions do not in fact hold in the real world, modeling may be of limited use or even lead us astray. To evaluate the practicability of such assumptions, we thus view ongoing research on both the nature of human emotion and the nature of knowledge representation and learning as essential.

Second, it is important to keep in mind that just like a child learning about emotions, a computational model's learning will be highly dependent on the experiences to which it is exposed. This dependency realistically reflects the inherent variability (across individuals, contexts, groups, and cultures) in how people think about and experience emotions. Furthermore, it presents a tremendous opportunity to explore the impact of different inputs on the developing child. Indeed, researchers in other fields of development have begun leveraging computational models in such ways. For example, language development researchers have explored how the quality and quantity of caregivers' verbal input to children may affect their developing word learning abilities (Borovsky & Elman, 2006; Jones & Rowland, 2017). Computational modeling could be used to explore similar input-related questions about emotion learning, such as how having harsh, hostile, or unpredictable caregivers could affect a child's learning.

However, this dependency can also be a challenge when our knowledge of what constitutes the range of typical input to the developing child is limited (Ruba et al., 2022). Such is arguably the case in emotional development, which as a field has yet to see the explosion in descriptive research on natural input that has characterized fields such as language development in recent years (e.g., Bergelson et al., 2019). We currently know relatively little about children's exposure to the variety of different emotion cues they encounter across ages and cultures. Likewise, researchers are just starting to quantify children's exposure to different emotion words across development (e.g., Ogren & Sandhofer, 2021). Particularly when we seek to model longer term developmental phenomena, this leaves us without a firm empirical foundation for deciding what data should be used for model training to most veridically simulate the input that children receive. We thus see descriptive research on children's everyday emotional experiences as essential to improving the potential of computational models of emotional development going forward.

Third, any model of children's emotional development will necessarily zoom in on certain developmental processes, at the expense of simplifying out others. In recent years, researchers have made progress when it comes to the breadth of behaviors they are able to capture in a single model. For example, researchers have developed computational models that can switch among different types of learning tasks and transfer knowledge across these tasks, as humans do (e.g., Lampinen & McClelland, 2020; Mnih et al., 2015). Likewise, outside of the field of development, many of the Bayesian models of emotion inference discussed earlier (de Melo et al., 2014; Houlihan et al., 2023; Ong et al., 2019; Saxe & Houlihan, 2017; Wu et al., 2018) are designed as unifying frameworks meant to capture, in a single model, how adults might make a variety of forward and backward inferences surrounding others' emotions. However, particularly when we begin to take development into account, it is important to keep in mind that no model is likely to capture all of the myriad cognitive tasks and learning opportunities the developing child faces at the same time. Accordingly, anyone seeking to build a computational model of emotional development must make careful, empirically informed decisions about what to include in their model, and what to leave out; consumers of these models should likewise maintain a critical eye toward these omissions. However, even with these omissions, computational modeling may provide insight into the potential causes of the developmental phenomena we observe. In this sense, we concur with Box's (1976) idea that "all models are wrong, but some are useful."

Conclusion

So far, emotional development researchers have not tapped the potential of computational models of development, but we believe these methods are ripe with opportunity. Applications of these models in other domains suggest they might provide us with generative insights about the origins of developmental change, connecting theory to empirical research and furthering our understanding of the mechanisms underlying emotional development. These models have the potential to serve as invaluable tools in the next phase of emotional development research, as we seek to understand not only what changes across emotional development but also how these changes come about.

References

Asaba, M., Ong, D. C., & Gweon, H. (2019). Integrating expectations and outcomes: Preschoolers' developing ability to reason about others' emotions. *Developmental Psychology*, *55*(8), 1680–1693. https://doi.org/10.1037/dev0000749

Austerweil, J. L., Gershman, S. J., Tenenbaum, J. B., & Griffiths, T. L. (2015). Structure and flexibility in Bayesian models of cognition. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford handbook of computational and mathematical psychology* (pp. 187–208). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199957996.013.9

Bauer, P. J. (2013). Memory. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology* (Vol. 1, pp. 505–541). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199958450.013.0018

Benton, D. T. (2023). 'But what is the mechanism?' Demystifying the ever elusive 'developmental mechanism'. *Infant and Child Development*, *32*(6), Article e2355. https://doi.org/10.1002/icd.2355

Bergelson, E., Casillas, M., Soderstrom, M., Seidl, A., Warlaumont, A. S., & Amatuni, A. (2019). What do North American babies hear? A large-scale cross-corpus analysis. *Developmental Science*, *22*(1), Article e12724. https://doi.org/10.1111/desc.12724

Berthiaume, V. G., Shultz, T. R., & Onishi, K. H. (2013). A constructivist connectionist model of transitions on false-belief tasks. *Cognition*, *126*(3), 441–458. https://doi.org/10.1016/j.cognition.2012.11.005

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language*, *33*(4), 759–790. https://doi.org/10.1017/S0305000906007574

Borsboom, D., van der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, *16*(4), 756–766. https://doi.org/10.1177/1745691620969647

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799. https://doi.org/10.1080/01621459.1976.10480949

Buss, K. A., Cole, P. M., & Zhou, A. M. (2019). Theories of emotional development: Where have we been and where are we now? In V. LoBue, K. Pérez-Edgar, & K. A. Buss (Eds.), *Handbook of emotional development* (pp. 7–25). Springer International Publishing. https://doi.org/10.1007/978-3-030-17332-6_2

de Melo, C. M., Carnevale, P. J., Read, S. J., & Gratch, J. (2014). Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology*, *106*(1), 73–88. https://doi.org/10.1037/a0034251

Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 515–548). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199399550.013.28

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, *521*(7553), 452–459. https://doi.org/10.1038/nature14541

Goodman, N. D. (2013). The principles and practice of probabilistic programming. *SIGPLAN Notices*, *48*(1), 399–402. https://doi.org/10.1145/2480359.2429117

Goodman, N. D., Baker, C. L., Bonawitz, E. B., Mansinghka, V. K., Gopnik, A., Wellman, H., Schulz, L., & Tenenbaum, J. B. (2006). *Intuitive theories of mind: A rational approach to false belief* [Conference session]. Annual Meeting of the Cognitive Science Society, Vancouver, BC, Canada. https://escholarship.org/uc/item/69k8c7v6

Gopnik, A., & Bonawitz, E. (2015). Bayesian models of child development. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 75–86. https://doi.org/10.1002/wcs.1330

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory. *Psychological Bulletin*, *138*(6), 1085–1108. https://doi.org/10.1037/a0028044

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364. https://doi.org/10.1016/j.tics.2010.05.004

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (pp. 59–100). Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.006

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, *16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Hoemann, K., Wu, R., LoBue, V., Oakes, L. M., Xu, F., & Barrett, L. F. (2020). Developing an understanding of emotion categories: Lessons from objects. *Trends in Cognitive Sciences*, *24*(1), 39–51. https://doi.org/10.1016/j.tics.2019.10.010

Houlihan, S. D., Kleiman-Weiner, M., Hewitt, L. B., Tenenbaum, J. B., & Saxe, R. (2023). Emotion prediction as computation over a generative theory of mind. *Philosophical Transactions of the Royal Society A: Mathematical, Physical, and Engineering Sciences*, *381*(2251), Article 20220047. https://doi.org/10.1098/rsta.2022.0047

Houlihan, S. D., Ong, D., Cusimano, M., & Saxe, R. (2022). *Reasoning about the antecedents of emotions: Bayesian causal inference over an intuitive theory of mind* [Conference session]. Annual Meeting of the Cognitive Science Society, Toronto, ON, Canada. https://escholarship.org/uc/item/7sn3w3n2

Jefferys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist*, *80*(1), 64–72.

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive Psychology*, *98*, 1–21. https://doi.org/10.1016/j.cogpsych.2017.07.002

Kording, K. P., Blohm, G., Schrater, P., & Kay, K. (2020). Appreciating the variety of goals in computational neuroscience. *Neurons, Behavior, Data Analysis, and Theory*, *3*(6), 1–12. https://doi.org/10.48550/arXiv.2002.03211

Lagattuta, K. H. (2014). Linking past, present, and future: Children's ability to connect mental states and emotions across time. *Child Development Perspectives*, *8*(2), 90–95. https://doi.org/10.1111/cdep.12065

Lagattuta, K. H., & Wellman, H. M. (2001). Thinking about the past: Early knowledge about links between prior experience, thinking, and emotion. *Child Development*, *72*(1), 82–102. https://doi.org/10.1111/1467-8624.00267

Lagattuta, K. H., Wellman, H. M., & Flavell, J. H. (1997). Preschoolers' understanding of the link between thinking and feeling: Cognitive cuing and emotional change. *Child Development*, *68*(6), 1081–1104. https://doi.org/10.1111/j.1467-8624.1997.tb01986.x

Lampinen, A. K., & McClelland, J. L. (2020). Transforming task representations to perform novel tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(52), 32970–32981. https://doi.org/10.1073/pnas.2008852117

Lapate, R. C., & Shackman, A. J. (2018). Afterword: What develops in emotional development? In A. S. Fox, R. C. Lapate, A. J. Shackman, & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (2nd ed., pp. 399–402). Oxford University Press.

Lara, K. H., Lagattuta, K. H., & Kramer, H. J. (2019). Is there a downside to anticipating the upside? Children's and adults' reasoning about how prior expectations shape future emotions. *Child Development*, *90*(4), 1170–1184. https://doi.org/10.1111/cdev.12994

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., Markson, L., & Hu, J. (2014). The child as econometrician: A rational model of preference understanding in children. *PLOS ONE*, *9*(3), Article e92160. https://doi.org/10.1371/journal.pone.0092160

MacKay, D. J. (2003). *Information theory, inference, and learning algorithms*. Cambridge University Press.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*(8), 348–356. https://doi.org/10.1016/j.tics.2010.06.002

McClelland, J. L., & Rumelhart, D. E. (1989). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. MIT Press.

Minsky, M., & Papert, S. (1969). *Perceptrons*. MIT Press.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Nobandegani, A. S., & Shultz, T. R. (2022). Computational approaches to cognitive development. In O. Houdé & G. Borst (Eds.), *The Cambridge handbook of cognitive development* (pp. 318–338). Cambridge University Press. https://doi.org/10.1017/9781108399838.018

Ogren, M., & Sandhofer, C. M. (2021). Emotion words in early childhood: A language transcript analysis. *Cognitive Development*, *60*, Article 101122. https://doi.org/10.1016/j.cogdev.2021.101122

Ong, D. C., Zaki, J., & Goodman, N. D. (2015). Affective cognition: Exploring lay theories of emotion. *Cognition*, *143*, 141–162. https://doi.org/10.1016/j.cognition.2015.06.010

Ong, D. C., Zaki, J., & Goodman, N. D. (2019). Computational models of emotion inference in theory of mind: A review and roadmap. *Topics in Cognitive Science*, *11*(2), 338–357. https://doi.org/10.1111/tops.12371

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, *21*(6), 425–433. https://doi.org/10.1016/j.tics.2017.03.011

Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to Bayesian models of cognitive development. *Cognition*, *120*(3), 302–321. https://doi.org/10.1016/j.cognition.2010.11.015

Pollak, S. D., Camras, L. A., & Cole, P. M. (2019). Progress in understanding the emergence of human emotion. *Developmental Psychology*, *55*(9), 1801–1811. https://doi.org/10.1037/dev0000789

Rakison, D. H., & Lupyan, G. (2008). Developing object concepts in infancy: An associative learning perspective. *Monographs of the Society for Research in Child Development*, *73*(1), 1–110. https://doi.org/10.1111/j.1540-5834.2008.00454.x

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21. https://doi.org/10.1037/0012-1649.33.1.12

Ruba, A. L., & Pollak, S. D. (2020). The development of emotion reasoning in infancy and early childhood. *Annual Review of Developmental Psychology*, *2*(1), 503–531. https://doi.org/10.1146/annurev-devpsych-060320-102556

Ruba, A. L., Pollak, S. D., & Saffran, J. R. (2022). Acquiring complex communicative systems: Statistical learning of language and emotion. *Topics in Cognitive Science*, *14*(3), 432–450. https://doi.org/10.1111/tops.12612

Rueda, M. R., & Posner, M. I. (2013). Development of attention networks. In P. D. Zelazo (Ed.), *The oxford handbook of developmental psychology* (Vol. 1, pp. 683–705). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199958450.013.0024

Ruffman, T., Aitken, J., Wilson, A., Puri, A., & Taumoepeau, M. (2018). A re-examination of the broccoli task: Implications for children's understanding of subjective desire. *Cognitive Development*, *46*, 79–85. https://doi.org/10.1016/j.cogdev.2017.08.001

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation*. https://apps.dtic.mil/sti/citations/ADA164453

Sanger, D. (1989). Contribution analysis: A technique for assigning responsibilities to hidden units in connectionist networks. *Connection Science*, *1*(2), 115–138. https://doi.org/10.1080/09540098908915632

Saxe, R., & Houlihan, S. D. (2017). Formalizing emotion concepts within a Bayesian model of theory of mind. *Current Opinion in Psychology*, *17*, 15–21. https://doi.org/10.1016/j.copsyc.2017.04.019

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Shultz, T. R. (2017). Constructive artificial neural-network models for cognitive development. In N. Budwig, E. Turiel, & P. D. Zelazo (Eds.), *New perspectives on human development* (pp. 15–26). Cambridge University Press. https://doi.org/10.1017/CBO9781316282755.003

Shultz, T. R., & Elman, J. (1993). Analyzing cross-connected net works. In M. I. Jordan, Y. LeCun, & S. A. Solla (Eds.), *Advances in neural information processing systems* (Vol. 6, pp. 1117–1124). Morgan-Kaufmann. https://proceedings.neurips.cc/paper_files/paper/1993/file/81e74d678581a3bb7a720b019f4f1a93-Paper.pdf

Spencer, J. P. (2020). The development of working memory. *Current Directions in Psychological Science*, *29*(6), 545–553. https://doi.org/10.1177/0963721420959835

Stojnić, G., Gandhi, K., Yasuda, S., Lake, B. M., & Dillon, M. R. (2023). Commonsense psychology in human infants and machines. *Cognition*, *235*, Article 105406. https://doi.org/10.1016/j.cognition.2023.105406

Teo, D. W. H., Ang, Z. Y., & Ong, D. (2022). *Modeling causal inference from emotional displays* [Conference session]. Annual Meeting of the Cognitive Science Society, Toronto, ON, Canada. https://escholarship.org/uc/item/3w06g892

van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science*, *16*(4), 682–697. https://doi.org/10.1177/1745691620970604

Walle, E. A., Lopez, L. D., & Castillo, A. (2022). Emotional development: A field in need of a (cognitive) revolution. In D. Dukes, A. C. Samson, & E. A. Walle (Eds.), *The Oxford handbook of emotional development* (pp. 61–77). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198855903.013.16

Westermann, G., & Mareschal, D. (2012). Mechanisms of developmental change in infant categorization. *Cognitive Development*, *27*(4), 367–382. https://doi.org/10.1016/j.cogdev.2012.08.004

Westermann, G., Sirois, S., Shultz, T. R., & Mareschal, D. (2006). Modeling developmental cognitive neuroscience. *Trends in Cognitive Sciences*, *10*(5), 227–232. https://doi.org/10.1016/j.tics.2006.03.009

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

Wu, Y., Baker, C. L., Tenenbaum, J. B., & Schulz, L. E. (2018). Rational inference of beliefs and desires from emotional expressions. *Cognitive Science*, *42*(3), 850–884. https://doi.org/10.1111/cogs.12548

Wu, Y., Schulz, L. E., Frank, M. C., & Gweon, H. (2021). Emotion as information in early social learning. *Current Directions in Psychological Science*, *30*(6), 468–475. https://doi.org/10.1177/09637214211040779

Yermolayeva, Y., & Rakison, D. H. (2014). Connectionist modeling of developmental changes in infancy: Approaches, challenges, and contributions. *Psychological Bulletin*, *140*(1), 224–255. https://doi.org/10.1037/a0032150

Received October 17, 2023
Revision received April 3, 2024
Accepted June 29, 2024 ■