

Using Dynamic Field Theory to Rethink Infant Habituation

Gregor Schöner
Ruhr-Universität

Esther Thelen
Indiana University

Much of what psychologists know about infant perception and cognition is based on habituation, but the process itself is still poorly understood. Here the authors offer a dynamic field model of infant visual habituation, which simulates the known features of habituation, including familiarity and novelty effects, stimulus intensity effects, and age and individual differences. The model is based on a general class of dynamic (time-based) models that integrate environmental input in varying metric dimensions to reach a single decision. Here the authors provide simulated visual input of varying strengths, distances, and durations to 2 coupled and interacting fields. The 1st represents the activation that drives “looking,” and the 2nd, the inhibition that leads to “looking away,” or habituation. By varying the parameters of the field, the authors simulate the time course of habituation trials and show how these dynamics can lead to different depths of habituation, which then determine how the system dishabituates. The authors use the model to simulate a set of influential experiments by R. Baillargeon (1986, 1987a, 1987b) using the well-known “drawbridge” paradigm. The dynamic field model provides a coherent explanation without invoking infant object knowledge. The authors show that small changes in model parameters can lead to qualitatively different outcomes. Because in typical infant cognition experiments, critical parameters are unknown, effects attributed to conceptual knowledge may be explained by the dynamics of habituation.

Keywords: infant habituation, dynamic systems theory, infant perception, infant cognition

What do infants know? When do they know it? Are some concepts inherent from the beginning, or does knowledge emerge only through experience in the world? There is a lot at stake in the answers to these questions. For millennia, philosophers and scientists, from Aristotle and Kant to Piaget and Chomsky, have tried to understand the human mind from speculating about how cognition begins. Grand epistemological edifices are built on assumptions about the initial state. Human nature may be very different if babies are born as blank slates, absorbing their environments like a sponge, than if they have kernels or essences of knowledge already built in and need only to elaborate them. The questions are old ones, but they are still hotly debated by philosophers, psychologists, and linguists (e.g., Haith, 1998; Pinker, 1994; Spelke & Newport, 1998; Thelen & Smith, 1994; and many others).

Unfortunately, it is very difficult to ask babies these important questions directly. Infants do not talk and thus have limited ways to express the workings and contents of their minds. For many months, they have little or no control of their limbs, cannot follow verbal directions, and are remarkably indifferent to their privileged status in the philosophy of mind. Until very recently, speculations about what infants know remained only speculations.

In the last 2 decades or so, a new experimental procedure appeared to finally open this window on infant cognition. The procedure capitalizes on what young infants can do—to look at visual displays and look away—and the pervasive tendency of all animals to decrease reactivity to a repeated stimulus—to habituate—and to reactivate interest when the stimulus changes. In infant habituation studies, babies are shown repeated displays of items or events until they indicate by looking away that they are no longer interested. Then infants are tested with new items or events. If the infants’ looking times recover to the test items or events, experimenters infer that the infants detect some salient difference between the first, habituating displays and the second, dishabituating stimuli.

To use habituation to ask what infants know, experimenters design the habituating displays to embody some aspect of the concepts in question. They then construct the dishabituating events to offer a violation of that concept. If infants dishabituate to the “wrong” display, researchers believe that the babies notice that the world is behaving badly. For example, in well-known experiments, Spelke, Breinlinger, Macomber, and Jacobson (1992) asked whether 3- to 4-month-old infants know that an object that falls from view follows a continuous path that cannot pass through solid surfaces. They habituated infants to a display of a ball falling behind a screen. They then lifted the screen to show the infant the ball resting on the floor of the display. For the test trials, they introduced a second horizontal surface above the floor. In the “consistent,” or “possible,” test condition, infants saw the ball resting on the top of the second surface. In the second test condition, the ball rested under the second surface, “inconsistent,” or “impossible,” because falling balls cannot pass through or jump over solid surfaces. Infants looked longer at the inconsistent event, because, the experimenters wrote, they

Gregor Schöner, Institut für Neuroinformatik, Ruhr-Universität, Bochum, Germany. Esther Thelen (deceased December, 29, 2004), Department of Psychology, Indiana University.

We thank Rowan Candy, Melissa Clearfield, Kelly Mix, and Virgil Whitmyer for their very helpful comments. This work was supported by National Institutes of Health Grant R01 HD2283013 to Esther Thelen.

Correspondence concerning this article should be addressed to Gregor Schöner, Institut für Neuroinformatik, Ruhr-Universität Bochum 44780, Bochum, Germany. E-mail: Gregor.Schoener@rub.de

“inferred that the ball would move on a connected, unobstructed path” (p. 613).

It is not an understatement to say that much of what we know about infant cognition is based on experiments of this general class. Infants are habituated to a display that is consistent with the real attributes of the world, such as number of things or events or the solidity or continuity of objects, or notions of support, causality, inertia, or gravity. The dishabituating event is similar to the habituating event, except for a slight change that makes the event “impossible” in the real world. If infants’ looking time to this “impossible” event increases, the reasoning goes, it is because it violates what they already know, and they are surprised by this violation. Infants thus demonstrate by the simple act of looking that they understand how the world works. They need not wait for language or even more complex manual responses to show the hidden contents of their minds.

It is also not an understatement to say that experiments of this type have revolutionized the field of infant cognition. First, on the basis of such procedures, experimenters have challenged the very foundations of Piagetian theory. Piaget (1952, 1937/1954) proposed that infants construct concepts gradually through their continued activities in the world. For example, infants learn about object properties through looking at them, handling them, watching others with them, and so on. But the looking experiments suggested that infants know about the existence of objects and their properties at much younger ages than predicted by Piaget’s epigenetic account—at ages when they cannot even manipulate objects—and perhaps already at birth. If true, Piaget seriously underestimated both the initial state and the processes by which knowledge is acquired. Second, the habituation experiments have led to claims of remarkable intellectual precocity in young infants. For instance, from her experiments, Wynn (1995) has claimed that infants have “impressive early competence” in the domain of numbers and that “a system of numerical knowledge may be part of the inherent structure of the human mind” (p. 172).

Recently, this class of “possible–impossible” experiments have themselves come under a storm of criticism. The heart of the critique is that researchers are making unwarranted inferences about complex cognitive abilities from experiments that reveal only simpler perceptual processing effects (Haith, 1998). In all uses of the habituation paradigm, the most conservative interpretation of dishabituation after habituation is that infants notice a difference between the stimuli. Indeed, the procedure has been used extensively since the 1960s to learn about infant perception, with the assumption that if infants dishabituate to a new stimulus, we can be confident that they can perceive whatever changed (see Kellman & Arterberry, 1998, for review). For example, Slater, Morison, and Somers (1988) tested newborn infants’ sensitivity to visual orientation. They habituated infants to a striped pattern oriented at 45° from the vertical and tested them with the identical display and one shifted 90°. Newborns showed by recovered looking times that they indeed were sensitive to this difference in orientation. Here the question asked of the infants was a simple perceptual one, and the infants’ answers seem straightforward. More skepticism is introduced when the conceptual content of the displays are at issue. Are infants processing the higher level concepts (such as knowledge of the paths of object when they fall) or simply noticing that something is different between two events?

Moreover, critics have argued that experiments do not control for the dynamics of the habituation process itself. For example, infants appear to prefer familiar stimuli after a few habituation trials and only switch to a novelty preference after many trials. Thus, in some experiments, increased looking at “impossible” events may be a reflection of preference for the familiar, rather than detection of physical law violation (Schilling, 2000). Likewise, critics say that experimenters may have failed to consider the relative novelty and salience of their events, and thus attribute complex cognitive significance to events processed at a more perceptual level (Rivera, Wakeley, & Langer, 1999). There are also troubling individual difference effects, such that infants who habituate quickly demonstrate different conceptual knowledge than those who do so slowly. The implication is that all of the effects attributed to “knowledge” must be reevaluated in light of these basic processes.

Given the central role of habituation to the entire field of infant cognition, it is critical to deeply understand the process itself. Whether looking experiments are called “violation of expectancy” studies (Baillargeon, 2000) or “preferential looking” (Spelke & Newport, 1998), the structure is identical: Infants are shown one stimulus event until they reach some level of disinterest and then tested with a different event. Thus, regardless of the putative cognitive content of the displays, or the intent of the experimenters, processes of habituation and dishabituation must be involved. Many questions remain about the nature of these processes in infants, and inferences about cognition are unresolved until we know more about how cognition is measured.

Here we present a field model of infant habituation grounded in dynamic systems principles of behavior and development. We show that this model can accommodate the known experimental effects that are now the topic of intense debate in the infant literature. Furthermore, we offer new predictions based on the model which can be tested empirically. If this model fulfills its promise as a general-purpose explanation of habituation in infancy, then indeed, it can be used to assess whether inferences from habituation experiments are warranted.

The model is based on two dynamically coupled processes that drive looking and looking away, an activation process and an inhibitory one. The dynamic field equations allow metric specification of stimulus strength, distance, stimulus intervals, and coupling between activation and inhibition. Because these variables can assume continuous parameters, we can precisely simulate small and gradual changes in the habituation and test events.

There is an important take-home message from the dynamic model and the simulations that accompany it: Seemingly very small changes in the stimuli, in the timing of presentations, in the metric differences of the test events, in the number of trials and habituation criteria, and in individual infants can radically change the outcome measure of whether a baby dishabituates to a test stimulus. In many (or even most) infant habituation studies, these factors are not well specified and often are unknown and unreported. When infants are shown complex events that occur over time, and when test stimuli differ from each other and from the habituating stimuli in subtle but unknown ways, the reasons for infants’ behavior at the test remain obscure. The benefits of a model, therefore, are to make these interacting processes explicit and to make principled predictions using parameters that may interact in complex and nonlinear ways.

But there is a second reason to understand the dynamics of infant habituation. Habituation and dishabituation are fundamental properties of animal nervous systems that allow decrements of response to repeated stimuli and renewed responses to novel stimuli. They provide mechanisms to ignore the environment when it presents no immediate threats or rewards and to focus attention on potentially important new input, and even to seek novelty. The processes must be working whether the stimuli have rich “cognitive” content or much simpler aversive consequences (Thompson & Spencer, 1966). We see habituation–dishabituation, therefore, as the counterpart to processes by which habits are formed. Organisms need to form memories of successful encounters with the world to build a repertoire of adaptive responses and provide stability for the self and others (James, 1899). But too deep a habit means getting stuck, limiting flexibility as the situation changes. In the discussion, we address the implications of our model both for the specific issue of infant cognition and for the more general problem of how animals develop both stability and flexibility.

Habituation: What Needs to Be Explained

Hundreds of experiments use the habituation method to test what infants perceive or what they know, but there are a number of unresolved issues surrounding the procedure. For example: (a) The habituation procedure leads to preferences for familiarity rather than novelty under some conditions; does such familiarity preference lead to spurious evidence for infant knowledge? (b) Dishabituation depends on the order of test stimuli; why is looking in response to violation of expectancy weaker when the surprising stimulus comes second on test? (c) Individual differences lead to different rates of habituation; Why does most evidence for infant knowledge come from fast habituators? The dynamic field model provides explanations for these questions in the context of understanding the basic mechanisms of habituation itself.

Knowledge or Novelty Preference?

All habituation researchers share the assumption that when infants’ looking times increase with the presentation of a test stimulus, the infants must have noticed a difference between the habituating stimulus and the test. What is unclear is the meaning of this difference to the infant.

A widely accepted assumption is that after viewing a scene consistent with their knowledge of the world, infants express surprise when the dishabituating event is inconsistent with that knowledge. Critics, on the other hand, contend that infants are responding on the basis of a more basic and content-free preference for novel or for familiar stimuli (see *Infancy* volume 1, No. 4, 2000). These critics suggest that the class of possible–impossible experiments have not adequately considered and controlled for novelty versus familiarity preferences. Because test stimuli are usually perceptually similar to the habituating events, increased looking to impossible events may simply reflect a tendency to look at something familiar. From this point of view, then, the response to the test event is determined by the dynamics of the habituating stimuli and does not reflect any core “knowledge.”

In fact, there has been considerable research on infants’ preferences for familiarity versus novelty in both looking and manipulation. This question was quite thoroughly studied several decades

ago (Colombo, Mitchell, O’Brien, & Horowitz, 1987; Hunter, & Ames, 1988; Hunter, Ames, & Koopman, 1983; Hunter, Ross, & Ames, 1982; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982). The data from these and a spate of more recent studies (Bogartz, Shinskey, & Schilling, 2000; Cashon & Cohen, 2000; Cohen, 2002; Roder, Bushnell, & Sasseville, 2000; Schilling, 2000) converge on a set of well-documented principles:

1. Infants attend longer to a familiar stimulus if their habituation time has been relatively brief.
2. Infants display novelty preferences after longer habituation times.
3. Novelty preferences emerge after fewer habituation trials if the stimulus is simple rather than complex.
4. Novelty preferences emerge after fewer habituation trials if the infants are older.

In our model, we show that familiarity and novelty preferences can be subsumed under a single dynamic regime, and that indeed, they can emerge in any habituation paradigm regardless of content.

The Order of Test Stimuli

Familiarity or novelty preference is thus a central issue in the current debate, and it is incumbent upon a model of habituation to explain these results. But there is another, less discussed, finding that also warrants attention. In her highly cited and now-contentious “drawbridge” study of early infant knowledge, Baillargeon (1987a) reported an important asymmetry. Three-and-a-half and 4.5-month-old infants were habituated to a screen rotating back and forth 180°. The two test events were the possible event, which included a box which halted the rotation of the screen and the impossible event in which the box did not interrupt the full motion of the screen. Infants looked longer at the impossible event, but only if they saw this event first in the test sequence. Infants who saw the possible event first did not reliably dishabituate to either the possible or the impossible event when it was shown. This order effect was also found in Baillargeon (1987b) and in Baillargeon, Spelke, and Wasserman (1985) and was replicated in Rivera et al. (1999). Although this effect was dismissed as “of little theoretical interest” by Baillargeon (1987a, p. 659), it is intriguing and indeed important. We show how this order effect falls out of our habituation dynamics along with familiarity and novelty preferences. We also show that this effect is fragile, explaining how, under some regimes, no order effect emerges (see, e.g., Clearfield & Mix, 1999).

“Fast” and “Slow” Habitulators

Some babies take many trials to reach habituation criteria, whereas others habituate quickly. Researchers have studied these individual differences in infants’ abilities to engage and disengage from visual stimuli extensively over the last 25 years, primarily as indicators of information processing rates and as predictors of later intellectual functioning (see Bornstein, 1989; DeLoache, 1976; Frick, Colombo, & Saxon, 1999). These individual differences are also important for interpreting habituation experiments because

there is considerable evidence that fast and slow habituators also show different responses to the test stimuli. For instance, in Bailargeon (1987a), for 3.5-month-olds, only the fast habituators showed dishabituation to the impossible event. The slow habituators did not differentiate between the possible and impossible test events. Cashon and Cohen (2000) also found differences between the infants who did not habituate within the 20 trials and those who did, with the nonhabituaors apparently preferring the familiar over the novel test event. We show how such individual differences can be understood with the dynamic field model.

Theoretical Accounts of Infant Habituation

Completing the Representation

The currently accepted explanation of infant habituation centers on the familiarity versus novelty effects. Scholars reason that infants require a certain amount of exposure to the stimulus in order to completely encode it and form a complete representation. If the number of habituation trials is small, this encoding is incomplete. Infants will thus prefer to look at the familiar pattern to complete the encoding. When, however, infants are fully habituated and they have a well-coalesced memory representation, then they are attracted to the novel stimulus. More simple stimuli are encoded faster, leading to a novelty preference with fewer trials. Likewise, older infants also encode more quickly and are also more likely to show a novelty preference with fewer trials. Roder et al. (2000) posit three stages in the habituation process. First, infants display no preferences for familiar or novel. As habituation trials continue, they prefer familiarity when tested, and finally, with additional trials, a novelty preference emerges. They believe that such a process is adaptive, facilitating attention to stimuli long enough to learn and remember the features, but not so long as to capture attention from potentially new and more informative stimuli.

This information-processing account of infant habituation is a version of Sokolov's (1963) model-stimulus comparator theory. Sokolov proposed that organisms construct internal models or mental representations of the external world. According to this theory, when an organism first encounters a novel stimulus, it tries to match it with a stored representation. If no match is found, the organism attends to the stimulus. As the stimulus is repeatedly encountered, a full representation is built. When new stimuli match the representation, attention is inhibited. Response decrements therefore provide behavioral evidence of the construction of the representation in the nervous system.

It is unclear, however, from the comparator accounts what "encoding" actually means or how, or at what level, infants "build a representation" of the stimulus. Nor is it clear why infants need to complete a representation and what constitutes complete. Moreover, comparator theory has no principled way to explain dishabituation. How much, and in what ways, is the display or event which elicits renewed interest different from the old, boring one? Is it a difference in meaning or content, or just a difference? And if so, what constitutes a meaningful difference (Bornstein, 1989)?

For instance, it is possible that the dishabituating stimulus is just sufficiently different from the habituating stimulus to activate different neural pathways and to thus evoke a renewed amount of orientation and looking. In this case, the infant need not assign

either stimuli any meaning at all but may simply be responding on the basis of new perceptual information. It is also possible that dishabituation is not sparked by differences in kind but simply by differences in intensity. That is, the dishabituating stimulus may excite more pathways or raise the habituated pathways to new levels of activation. In either case, however, the process may be purely metric, that is, free of stored knowledge or meaning.

Dual-Process Theory

Indeed, both of these processes—difference and intensity—are invoked in a second, well-accepted general theory of habituation, the dual-process model (Thompson & Spencer, 1966). According to the dual-process model, habituation is driven by two independent and additive processes. When the organism receives repeated stimulation, one process tends to suppress responsiveness. This habituation occurs in pathways specific to the stimulus and leads to a decrement in responding. At the same time, the stimulus provokes a more general level of activation or arousal. This sensitization amplifies the activity of the stimulus-response pathway. If the habituating stimulus is sufficiently strong, the sensitization process causes an initial increase in responding. (This has been demonstrated in both infant and other animal experiments but is sometimes obscured by the manner in which trials are averaged.) As stimulus events are repeated, however, the habituation processes overwhelm the initial arousal. The rate of habituation is thus a function of the strength, or activating power, of the stimulus. More salient or arousing stimuli activate the system more strongly, leading to slower habituation. Dishabituation, in this view, may result from either, or both, processes: new-stimulus response pathways or increased sensitization. For example, dual-process theory predicts that after habituation, the response to a novel stimulus should be stronger when the novel stimulus is also more intense, because intensity increases overall sensitization.

There are many experimental results supporting dual-process theory. Most important is what Kaplan, Werner, and Rudy (1990) called "Thompson-Spencer dishabituation," after the authors of the theory. Thompson and Spencer (1966) describe an experiment in which rats are given skin shocks at a certain rate until they habituated. These stimuli were discontinued, and a single tail pinch was delivered. The rats responded to the tail pinch, but when given a test of the skin shock again, their responses to that old stimulus recovered, and even exceeded the original rate. A similar increased responsiveness was observed even when a third stimulus was introduced after the tail pinch. According to the authors, these experiments demonstrate the independence of the habituating and activating events. The tail pinch did not simply interrupt the habituation process but superimposed a response-incremental process, so that either familiar or novel stimuli received increased responding.

Indeed, Kaplan and Werner (1986) found some evidence of Thompson-Spencer dishabituation in infants. They habituated infants to checkerboard patterns that were of either low (4×4), medium (12×12), or high (20×20) densities. Infants initially looked longer at the high density patterns, as expected. After habituation, they responded most to the novel patterns that initially captured the longest looks. After the novel stimulus, infants were then retested with the original patterns. Only those infants dishabituated to the high-density novel stimulus showed recovery of the

original looking responses. The arousing high-density patterns, the authors reasoned, provided a nonspecific boost that increased interest in the original, habituated pattern. Dual-process theory thus predicts asymmetries in the dishabituation process.

Because dual-process theory involves the addition of two simultaneous processes, the timing of the habituating and dishabituating stimuli are especially critical. For example, in this theory, the length of the intervals between habituating stimuli should matter profoundly because the activating effect of the specific stimulus has a time-related delay. Kaplan and Werner (1986) showed this to be true: 4-month-old infants were shown eight presentations of the 4×4 checkerboard patterns, followed by two presentations of the highly activating 20×20 grids. They were then retested with the original 4×4 pattern either 10, 20, or 30 s after the termination of the novel pattern. Their amount of dishabituation to the original stimulus (Thompson–Spencer dishabituation) declined dramatically as a function of these delays and the presumed decay of the nonspecific arousing feature of the 20×20 display. Our model retains some aspects of dual-process theory in proposing an activation metric, but we incorporate this in a single-process mechanism.

The Dynamic Field Model

A general model of infant habituation is important because it provides a framework by which to assess the meaning of hundreds of experiments believed to tap into the infant mind. If habituation processes are governed by systematic dynamics, then those dynamics must apply to all experiments which have a similar structure, whatever the specific nature of the stimuli or their putative cognitive content or whether the process is labeled “violation of expectancy” or “preferential looking.” Thus, we begin with what is known about such general mechanisms.

A useful model of infant habituation must account for the following known experimental results¹ (see also Sirois & Marschal, 2002a):

1. Habituation is a process of declining interest in a stimulus display or event with repeated trials.
2. Initially, there is an increase of response, but then there is a decline.
3. The more arousing or interesting or complex the habituating stimulus, the more infants look at it and the longer it takes to reach a habituation criterion.
4. The strength of the dishabituation is a function both of the difference of the dishabituating stimulus from the original and its intensity or complexity.
5. When the dishabituating stimulus is presented after a few trials, infants appear to prefer the old stimulus.
6. In contrast, when the dishabituating stimulus is presented after many trials, infants appear to prefer the novel stimulus.
7. After habituation, there can be an asymmetric response to two test stimuli, based on which test (familiar or novel) is given first.

8. Dishabituation to a novel stimulus can reinvigorate interest in the familiar stimulus.
9. A nonspecific activation “boost” can reinvigorate interest in the familiar stimulus.
10. There are individual differences in the rates of habituation, which in turn, determine dishabituation.

Properties of Field Theories

Our model is a version of a general class of models that express behavioral decisions in the dynamics of continuously evolving activation fields, or distributions of activation over metric spaces. Field theories are based on the assumption that actions are the dynamic (and often nonlinear) function of both the immediate stimuli in the environment and both the recent and longer term history of the system in similar situations. Thus, such fields have the ability to integrate environmental information in the form of inputs of varying metric dimensions that may variously reinforce or compete and to reach a single decision to act. The critical dimensions of such theories, then, are their specification of both the inputs to the field and the decision field in continuous metric parameters, and their realization of an action decision from these multiple inputs.

The “dynamic field” concept has been used to model other behavioral choices that evolve with multiple perceptual inputs, including motor planning (Erlhagen & Schöner, 2002; Schöner, Kopecz, & Erlhagen, 1997), saccadic eye movements (Kopecz & Schöner, 1995), robot behavior (Schöner, Dose, & Engels, 1995), location memory (Schutte, Spencer, & Schöner, 2003), and infant perseverative reaching (Thelen, Schöner, Scheier, & Smith, 2001). Common to all these applications are (a) the representation of variables in a continuous metric space and (b) the evolution of the decision process gradually over time. In the first case, this means that all stimuli can be assigned continuous, metric variables that indicate their relative distance from one another (these assumptions are commonly made in theories of perceptual similarity). Additionally, each variable can assume different activation strengths, with an understanding that only variables that reach a particular threshold enter into the behavioral decision. Moreover, within a field, interaction between different activation variables can also be specified. In the second case, the field models assume that perceptual variables have dynamics, that is, they evolve gradually over time according to some dynamic regime, which may be

¹ For a critical discussion of Point 2, see Colombo, Frick, and Gorman, 1997. Points 3 and 4 are classical; see, for example, Fantz and Fagan, 1975; Karmel, 1969; McCall and Kagan, 1967. Points 5 and 6 are also well known; see, for example, Hunter et al., 1982, and Rose et al., 1982; Points 5 and 6 have recently been critically examined; for example, see Roder et al. (2000, and commentary in that issue); Clearfield and Westfahl (in press) demonstrates dramatically the crucial role of familiarization. Point 7 has been reported repeatedly in Baillargeon’s work (Baillargeon, 1986, 1987a, 1987b; Baillargeon et al., 1985); later studies did not analyze the role of order. For studies that support Point 8, see, for example, Kaplan et al. (1990) and critical assessment in Colombo et al. (1997). Support for Point 9 can be seen in, for example, Bogartz, Shinsky, and Speaker (1997). Point 10 is broadly supported; see, for example, Schilling (2000), for a recent study.

nonlinear. This assumption of continuous time means that the history of the system is critical, both within the experimental task but also from previous activities in similar tasks (see Erlhagen & Schöner, 2002). For example, in the field theory of infant reaching perseveration, the current perceptual attributes of the task were dynamically integrated with the infants' previous history of reaching. When the immediate input dominated the field, infants followed the most recent cue. However, when the inputs were not sufficiently strong to overcome previous actions, then infants' movements were shaped by what they had most recently done, even when it was inappropriate for the current task demands (Thelen et al., 2001).²

General Assumptions of the Habituation Model

We now apply these general assumptions of dynamic field models to the specific case of infant habituation. The critical behavior in habituation is the infant's decision to look at a display or to look away from it. In the model, this decision evolves in two, isomorphic and interacting fields that reflect the perceptual variables salient to the looking decision. The first is the activation field. This field represents the two properties of the stimuli: their metrics in perceptual space and their activation strengths. *Perceptual metrics* means that stimuli can thus be characterized on how close or different they are in the field (see Figure 1). For example, a stimulus parameter such as the movement of an object in an event can be defined on a continuum. An object moving horizontally across the display would be "closer" to one moving diagonally

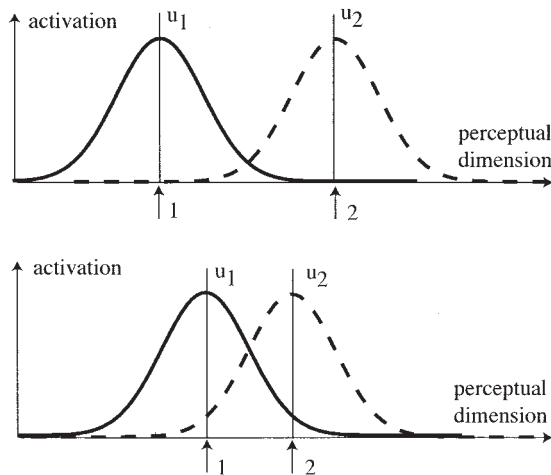


Figure 1. Schematic representation of the field concept. Each field site represents a particular value along a perceptual dimension (e.g., spatial position of stimulus motion, movement direction). Input to the activation field from one stimulus is modeled here by a Gaussian distribution centered over one value (1) of the perceptual dimension (solid line), input from another stimulus by a Gaussian distribution centered over another value (2). The activation field can be represented by the level of activation, u_1 and u_2 , at these two sites. Top panel: When two relevant stimuli are metrically distant, only one of the two activation variable receives input whenever one of the stimuli is presented. Bottom panel: When the two stimuli are metrically close, the second activation variable receives some overlapping input even when the stimulus is centered on the first activation variable's receptive field.

and originating from the same point than it would be to a vertical movement because it shares directionally specific parameters. Closeness is expressed in the model by the overlap of input activation and, potentially, by interaction. Thus, the horizontal input would share more activation with the diagonal stimulus than with the vertical one. The second important dimension on which the stimuli can vary is their activation strength. A more intense or more complex stimulus contributes more activation to the field.

Typical habituation experiments use only a small number of stimuli from the potential number represented by the field. Therefore, we can sample the field by a limited number of model sites, represented by activation variables, u_i ($i = 1, 2, \dots$; see Figure 1). When two or more sites are involved, the input at these sites may overlap to varying degrees, depending on how metrically distinct different stimuli are. Closer sites representing metrically more similar stimuli receive more shared input from a given stimulus (see Figure 1).

The second field is the habituation field. This is a field of inhibition variables. The field receives input from the activation field, not from the stimulus itself. This means that all stimuli produce both activation and habituation responses. This field represents the level of habituation for each type of stimulus. This assumes that habituation is stimulus specific, so that habituation to one particular stimulus does not necessarily lead to habituation to another stimulus. However, if the inputs to the activation field overlap sufficiently, a generalized habituation can result. Again, in practice only a finite number of inhibition variables is needed, v_i ($i = 1, 2, \dots$), one for each model site. The strength of the habituation sites is a function of their input from the activation field.

The activation field drives looking. Whenever any location in the field is activated above a threshold (here defined as zero), looking ensues. When the activation sinks below that threshold, looking stops.

The Specific Model

In the most basic version of the model, we consider an experiment in which infants are habituated to one stimulus and tested with a parametrically different stimulus. In the model, these percepts can be represented as two different sites ($i = 1, 2$). Basic dynamic parameters such as the time scale, resting level, coupling with inhibition field, and noise are identical for these two sites. What can be varied parametrically are the stimulus strength, the duration of time the stimuli are "on" or "off" and the closeness, or overlap, of the two sites. The threshold for looking is set to zero; activation must exceed zero for looking to be elicited. The threshold for the coupling of activation and inhibition is also zero. This means that these processes are mutually interactive even when there is no looking.

Note that we have made two simplifying assumptions. First, we sample the field with only two sites, with varying overlap. In reality, a stimulus field would be much more complex. Thus, the

² Note that as we use it here, the field model is an abstract description of behavior and not anatomically specific to particular brain areas, nor are "sites" in the model descriptions of neurons. However, field model concepts are fully compatible with the dynamics of pattern formation in cortex (e.g., Amari, 1977; Bastian, Schöner, & Riehle, 2003).

only stimulus dimensions we manipulate here in addition to the timing are the intensity and the overlap between the two sites. Second, we have first simulated the most uncomplicated case of a fixed interval of stimulus onset and offset. Many infant experiments used the so-called “infant-controlled procedure,” in which the length of the stimulus durations is a function of infant looking—The stimulus remains on as long as the infant is looking at it. Likewise, the stimulus remains off as long as the infant looks away. As we discuss later, we have little or no data on the time structure of the looks away, in relation to the looks, nor do we know what about the stimulus causes reorientation to it (see Cohen, 1988; Gilmore, 2001). However, even in the simplified case, we expose the time structure of the processes, and we demonstrate that content-free dynamics can produce the basic habituation effect.

Mathematical formulation of the model. The model is given by the following equations:

$$\tau_u \dot{u}_1 = -u_1 + h_u + s_1(t) - c_u \sigma(v_1) + q_u \xi_1(t), \quad (1)$$

$$\tau_v \dot{v}_1 = -v_1 + h_v + c_v \sigma(u_1) + q_v \xi_2(t), \quad (2)$$

$$\tau_u \dot{u}_2 = -u_2 + h_u + s_2(t) - c_u \sigma(v_2) + q_u \xi_3(t), \quad (3)$$

$$\tau_v \dot{v}_2 = -v_2 + h_v + c_v \sigma(u_2) + q_v \xi_4(t), \quad (4)$$

where

$$\sigma(u) = \frac{1}{1 + \exp[-\beta u]} \quad (5)$$

is a nonlinear (sigmoidal) threshold function. Stochastic perturbations are modeled by four independent Gaussian white noise processes, $\xi_i(t)$ ($i = 1, 2, 3, 4$), of zero mean and unit variance.

Stimuli. The inputs, s_1 and s_2 , represent the stimulation supplied to sites u_1 and u_2 . These stimulus parameters vary in time according to the following:

$$s_i(t) = s_i^0 \begin{cases} 1 & \text{while stimulus for activation site } i \text{ present} \\ 0 & \text{no such stimulus present} \end{cases} \quad (6)$$

where $i = 1, 2$. This represents the experimental manipulation. The values are positive when the stimulus exciting that site is activated, and the values are zero when the corresponding stimulus is not active. We can also put parameter values on the durations during which a stimulus, j , is on or off, $\Delta t_{j, on}$ and $\Delta t_{j, off}$.

In some simulations, we assume that the two stimuli are not completely different but share properties. We model the similarity of the stimuli by making both stimulus strengths different from zero when a particular stimulus is applied: One stimulus also excites the other to some determined degree. In that case the stimulus strength, s^0 , is a function both of the identity of the site, i , and of the stimulus, j , that is, forms an input-output matrix: $s_{i,j}^0$.

Parameters. We model different experimental regimes by assigning parameter values to the following variables:

1. $\tau_u > 0$: The time scale of the activation variable. This determines how long it takes to bring about a substantial change of activation. This also fixes the units of time in the simulations. τ_u was 10 s in the simulations.
2. $\tau_v > 0$: The time scale of the habituation (inhibition)

variable. This determines how long it takes to bring about a substantial change in the level of habituation. τ_v was 200 s, or 10 trials of 20 s each.

3. $h_u < 0$: The resting level of the activation variables, relative to the threshold of the sigmoid activation function, which is set to zero. In the simulations, h_u was -1 . This also fixes the units of the activation and habituation variables.
4. $h_v < 0$: The resting level of the habituation variables, again, relative to the threshold of the sigmoid at zero. In the simulations, h_v was -1 .
5. $c_u > 0$: Strength of the inhibitory effect that the habituation variable has on the activation variable. This must be larger than $-h_u$ so that under full habituation the activation falls below zero, leading to no looking. In the simulations, c_u was set to 2.
6. $c_v > 0$: Strength of the excitatory effect that the activation variable has on the habituation variable. This must be larger than $-h_v$ so that above-threshold activation does indeed drive habituation to positive values. In the simulations, c_v was set to 2.
7. $\beta > 0$: Sharpness of the sigmoidal threshold function. The effect of β depends on the ranges within which the activation and habituation variables are allowed to vary. When these ranges are approximately between -1 and 1 , then when $\beta = 10$, the sigmoid is strongly nonlinear. When $\beta = 1$, it is only weakly nonlinear. Set to 10 for the simulations.
8. q_u and q_v : The effective noise strength in activation (u) and inhibition (v) sites. If too large, the dynamics become independent of the input; the stimuli do not effectively drive looking. If too small, looking is completely determined by the input. Noise is applied only in simulations exploring the effects of noise (see Figures 21, 22, 23). The noise strength then was $q_u = 0.025$ and $q_v = 0.005$, reflecting a difference in time scales between the two fields.
9. $s_i^0 = 0$ ($i = 1, 2$): Strength of the input to site i . This strength is relative to the resting level. For a stimulus to be effective in eliciting a response, the stimulus strength must be at least $-h_u > 0$. On the other hand, for a stimulus to be no longer effective when the system is fully habituated, the strength must not exceed $c_u - h_u > 0$.
10. $\Delta t_{j, on}$: Duration during which a particular stimulus, j , is on. Typical value is 10.
11. $\Delta t_{j, off}$: Duration during which a particular stimulus, j , is off. Typical value is 10.

Simulations. The simulation program habit.m (written in Matlab) generates time courses of the different inputs that reflect

different experiments in the habituation paradigm. For instance, the program can simulate the periodic presentation of habituation stimuli followed by a pause (interstimulus interval [ISI]) in a first phase, followed by a test phase in which one or several test stimuli are presented in a particular order, again with appropriate pauses.

Based on these time courses, the model equations are integrated. The program analyzes the resulting temporal patterns of activation, detects periods in which at least one activation variable is larger than a “looking threshold” (typically 0.0 or 0.5), leading to “looking” and periods in which all the activation variables are below the “looking threshold,” considered “not looking,” and then registers the durations of these events.

Results: The Dynamics of Habituation

In the first set of simulations, we show the dynamics of the habituation process itself, without a test phase, and using a single site. Thus, we first ask what happens when an infant sees repeated presentations of a single stimulus or stimulus event.

To interpret the simulations, refer to Figure 2, which also illustrates the form of all of the figures. The top panel of Figure 2 shows the dynamics of the habituating stimulus at a single site by displaying the stimulus, the activation variable, u , and the inhibition variable, v , as time series. The bottom panel of Figure 2 depicts the looking time that is the result of the two coupled

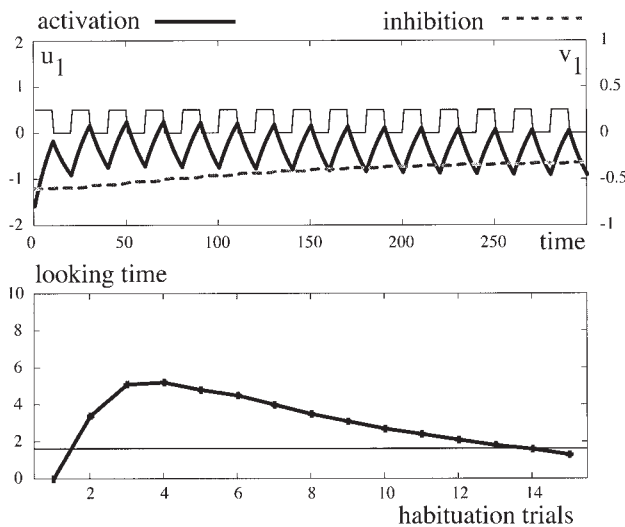


Figure 2. The basic dynamics of habituation. The top panel depicts the stimulus as the thin solid line. The stimulus can have varying strengths, durations, and interstimulus intervals. The thick solid line beneath the stimulus trace is the variable u , which reflects the activation from the stimulus. It rises and falls with the cycle of stimulus presentations of fixed length. The third, dotted line is v , the inhibition variable that is driven by u . This variable changes more gradually, increasing as the number of stimulus presentations increase, and in turn, inhibiting looking. The bottom panel shows the looking time, that is, the time within each stimulus presentation cycle during which activation is above the threshold for looking (zero). After an initial increase, looking time declines with repeated stimulus presentations under the accumulating influence of the growing inhibition strength. The horizontal line in the bottom panel represents the criterion level for habituation of looking times commonly used, half the average of the first three looking times.

processes. Note that looking and not looking are still driven by the stimulus presentation but that the duration of looking declines dramatically. Thus, although the system continues to receive input from the stimulus, the resulting activation from that stimulus gradually declines under the accumulating influence of the growing inhibition strength.³ (In Figures 2–7, the habituation phase consists of the presentation of the stimulus for 10 s and an ISI of 10 s, except in Figure 5, where ISI is 30 s).

Figure 2 displays the results of a simulation at a relatively weak stimulus strength (s^0 ; see Equation 6). Note that inhibition increases with the repeated activation from the stimulus. But this dynamic process also gives rise to complex interactions. Looking time increases initially because at the first trial, activation starts from its resting level but does not decay to its resting level during the subsequent ISI. This means that after the first trial, the activation starts from a nonzero baseline and thus drives more looking. However, the higher activation also engenders a stronger response in the inhibition variable, and after repeated trials, this overwhelms the activation and looking declines. With enough trials, inhibition is so dominant that repeated presentations of the stimuli evoke no interest at all.

The initial increase in looking time is of theoretical interest. Dual-process theorists invoke this initial rise in looking time as evidence for the independent process of sensitization: a generalized arousal mechanism that primes reactivity to the specific stimulus. But note that the field theory is not strictly a dual-process theory, and the initial rise is due to a different mechanism than suggested in dual-process theory. In our account, the activation dynamics alone drive their own increase, and through their coupling to the inhibition field, their own decrease as well. The behavioral results are similar, but the underlying processes are not. The initial rise in responding is an indicator of the underlying dynamics, but it is not the only pattern to emerge; for instance, in Figure 5 the simulation revealed the high initial level of looking followed by a gradual decrease as a result of a longer ISI, and in Figure 6 there is an exponential decrease characteristic of “fast” habituators.

How do infants actually behave? Many infant studies do not report the full time course of habituation. In those who do, the manner of reporting the data makes answering this question difficult. In their early studies, Caron and Caron (1968, 1969) report trial-by-trial results and show clear evidence of such an initial rise. So do Kaplan et al. (1990), but only with their most complex stimuli. Many other studies also show rises in looking time before habituation (e.g., Baillargeon et al., 1985, Figure 3; Kellman &

³ We have held certain parameter values constant across the different simulations, except where noted: $h_u = -1.6$; $h_v = -0.6$; $\tau_u = 15$; $\tau_v = 300$; $\beta = 5$; $c_u = c_v = 2$; threshold for looking = 0.0; initial conditions (except for “fast habituators”) for activation variables = h_u ; initial conditions for habituation variables = h_v ; duration of stimulus-on interval = 10 s; duration of stimulus-off interval = 10 s (except Figure 5). The number of habituation trials was set at 15 or 4, the number of test trials was 0 or 2 of each stimulus). Noise strength is the standard deviation of stochastic Gaussian white noise = 0, except in Figures 21, 22, and 23 in which noise is changed to noise levels $q_u = 0.025$ and $q_v = .005$. Figures 2–6 show a single pair of activation and habituation variables from the activation of only one site. In Figure 2, the stimulus strength is weak, $s^0 = 3$.

Spelke, 1983, Figures 4, 6, 8, 10, 12, and 14; Kellman, Spelke, & Short, 1986, Figures 3 and 6; Gibson, Owsley, & Johnson, 1978, Figure 4; Johnson & Aslin, 1995, Figures 3 and 4; Spelke et al., 1994, Figures 5, 6, 7, 9, and 11). However, it is probable that in these latter studies, the increases are artifacts of the way that the data are presented. As Cohen and Menten (1981) demonstrated, the common practice of backward averaging used in these studies produces this artificial peak. Backward averaging aligns all of the infants' looking time data on the trial at which they reached criterion. This could range from 4 to more than 10 trials. Infants who habituate in a few trials usually have long initial looking times and a rapid decline, although those who need more trials to habituate show a more gentle decline. When the two types of data are averaged backward, the fast habituators' long looking times are emphasized in the last few trials before criterion and produce such an artifactual peak. Similarly, however, forward averaging may also obscure the initial increase, because this can smooth away the different rates of increase for infants who habituated either quickly or slowly. (Because of these individual differences, Cohen and Menten caution about using group data to infer habituation processes.) Baillargeon's (1987a) results are especially interesting in this regard. In her Figures 4 and 5, only the "slow habituators" showed the initial rise; they also did not look longer at the impossible event. We return to this issue in the next sections.

Stimulus strength. The stronger the stimulus, the more activation during the stimulus onset. Looking times are thus longer for stronger stimuli. However, stronger activation also drives stronger inhibition, so that the habituation curves for stronger and less strong stimuli tend to converge over time. Compare Figure 3 with Figure 2. Figure 3 is the same simulation as that shown in Figure 2 but with an increased stimulus strength. Looking time increases, but it also declines with repeated trials. The simulation agrees with experimental results that show increased looking time to stimuli with increased spatial frequency, in which complex stimuli are assumed to be stronger (Caron & Caron, 1968, 1969; Kaplan et al., 1990).

One prediction of the model is that at larger stimulus strengths, the activation variable is driven at higher rates and the initial rise

period is shortened. This is not seen in either the Caron and Caron (1968) or the Kaplan et al. (1990) data. Caron and Caron found equal rise times to both 12×12 and 2×2 patterns. In contrast, Kaplan et al. found an initial increase, a signature of dual-process theory, only with complex stimuli. However, note that these are rather subtle effects in the model, which may be obscured by averaging group data. Finally, in Figure 4, we further increased the stimulus strength. This produces very long looks initially. We also see that if the stimulus is strong enough, the system hardly habituates at all; after 15 trials, the infant has not yet reached criterion. The activation simply swamps an accumulated inhibition. In other words, even with the strong inhibition produced by the stimulus, the activation variable overrides it.

ISI. According to the field model, the timing of stimulus onset and offset is critical. Because the looking field relaxes when the stimuli are off, longer intervals between stimulus presentations should affect both the initial rise of the looking variable and the habituation process itself. Thus, when the second stimulus is presented at, say, 30 s after the first, it encounters a lower existing activation level than when it occurs 10 s after the first. This simulation is presented in Figure 5. Compare Figure 5 with Figure 3: Both have stimulus strengths of $s^0 = 4.5$, but the second figure has three times the ISI. The initial rise is dampened and overall activation levels are less. However, this also results in weaker inhibition, so that the long-term effect of a longer ISI is more looking time after a large number of trials, as shown in Figure 5. Bashinski, Werner, and Rudy (1985) varied the ISIs (10, 20, and 30 s) for infants habituated to a 12×12 checkerboard. Longer ISIs indeed dampened the initial activation and led to initially decreased looking. However, Bashinski et al. did not continue with enough trials to test the second prediction of the model. Additional experiments are needed to see whether the short ISI group would have continued to habituate at a faster rate after seven or eight trials and whether the longer ISI group would continue to show a more gradual decline.

Although the durations of the habituating events and/or the ISIs are sometimes reported in infant habituation studies, these time dimensions are rarely considered experimental parameters and, thus, are not manipulated in experiments. However, the model and simulations suggest these are critical to the effects that are found.

Fast habituators. The model offers a new explanation for the individual differences seen in rates of habituation based on a simple difference in the initial level of activation before encountering any stimulus. In Figures 6 and 7, we have simulated a "fast" habituator. We increased the initial activation level of u to $h_u + 1.5$, which is an increase over the default resting value of $h_u = -1.0$. Because the stimulus (here set at a weak level of $s^0 = 3$) encounters an already activated system, the stimulus activation drives an initial high level of looking. High activation produces high inhibition, so that under this condition, the system becomes saturated, and looking times immediately begin to decrease rather rapidly. Infants quickly reach the usual criteria for habituation, which is a looking time of 50% of the average of the first three trials. These infants would be considered "efficient" information processors.

There is, in addition, an interaction between the initial state of the system and the stimulus strength. Figure 7 illustrates the increase in trials needed to reach criterion for "fast habituators,"

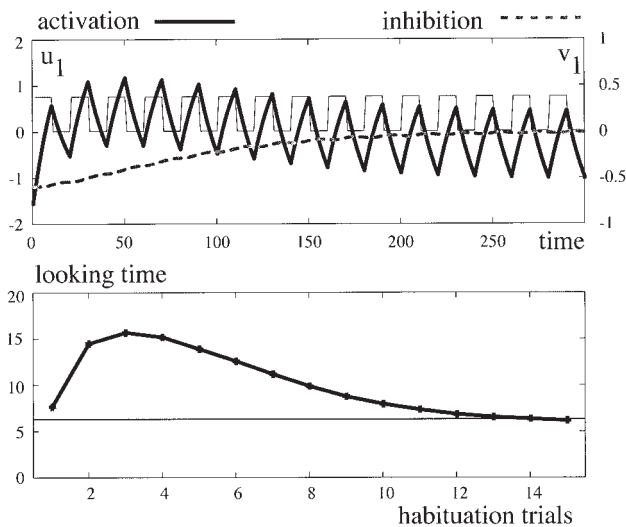


Figure 3. Stimulus strength is increased to $s^0 = 4.5$.

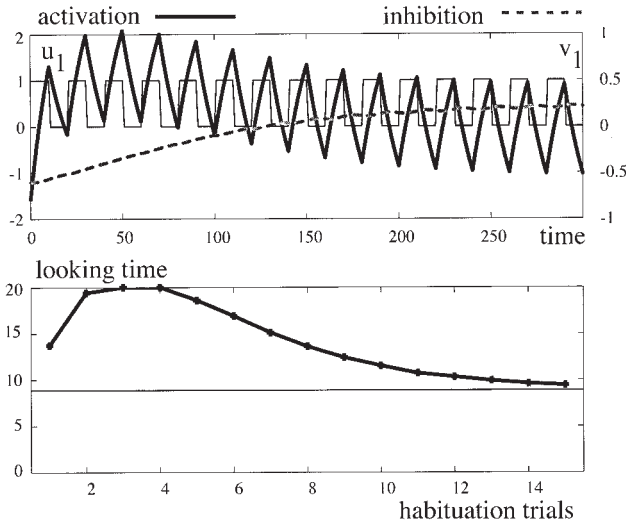


Figure 4. Stimulus strength is increased to $s^0 = 6$. Very little habituation results.

when the stimulus strength is increased to the medium level of $s^0 = 4.5$.

There is some good evidence supporting the dynamic account. Consider Figure 4 in Baillargeon (1987a), which shows that the fast habituators have longer looks during the 6 trials preceding criterion and decline rapidly, although the slow habituators have lower levels of looking and show fluctuating looking times. The data reported by Bornstein and Benasich (1986) are more compelling because they did not do the backward averaging. In their Figure 1, they show three patterns of habituation: “exponential decrease,” “increase–decrease,” and “fluctuating” (which shows an initial decrease before an increase). The exponential decrease group have initial high looking levels and rapid decline, as simulated in our Figures 6 and 7. The other two groups have much

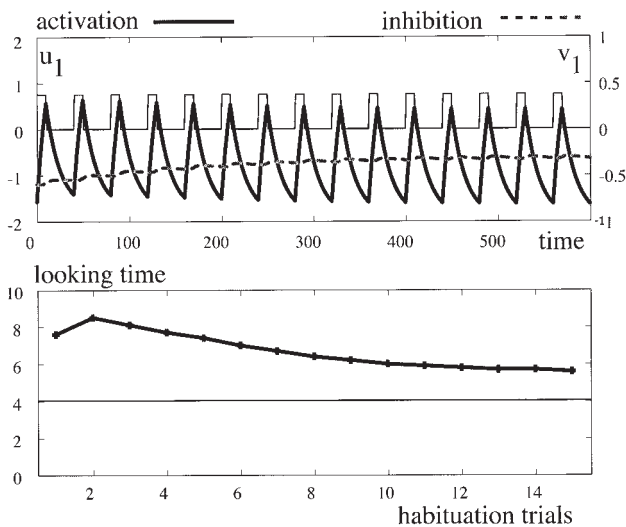


Figure 5. The interstimulus interval is increased to 30 instead of 10. There is less habituation. Medium stimulus strength ($s^0 = 4.5$).

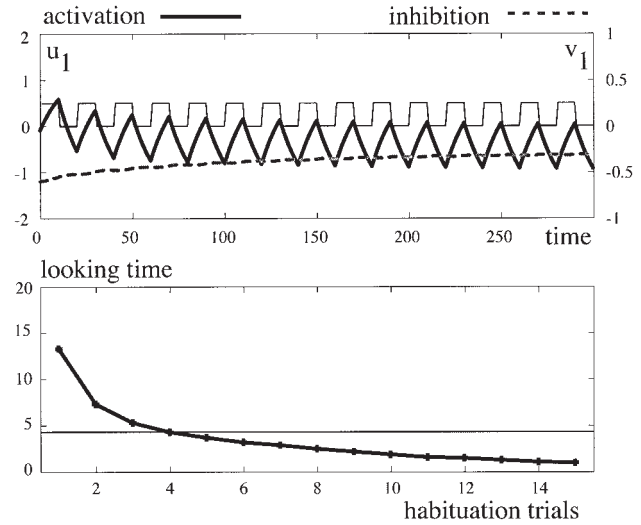


Figure 6. Simulation of “fast habituator.” Weak stimulus strength ($s^0 = 3.0$) but initial activation level increased to $h_u + 1.5$.

lower initial looking times and initial rises more similar to our Figures 2–4. In accord with our theoretical predictions, the exponential group averaged 4.6 trials to criterion whereas the other two groups averaged 7.2 and 7.8 trials, respectively. Thus, individual differences in habituation can be modeled by a simple adjustment of the resting level activation value. This creates conditions for which the system needs more or fewer trials to reach habituation criterion. An infant behaving like the simulation in Figure 6 would be given the dishabituating stimulus after 4 habituation trials, whereas one following the parameters set in Figure 2 would not be dishabituated until after 13 trials. Individual differences emerged strictly from the initial conditions and the dynamics of habituation that follow this initial look. In our simulations, the fast and slower habituators had no other differences; that is, their subsequent

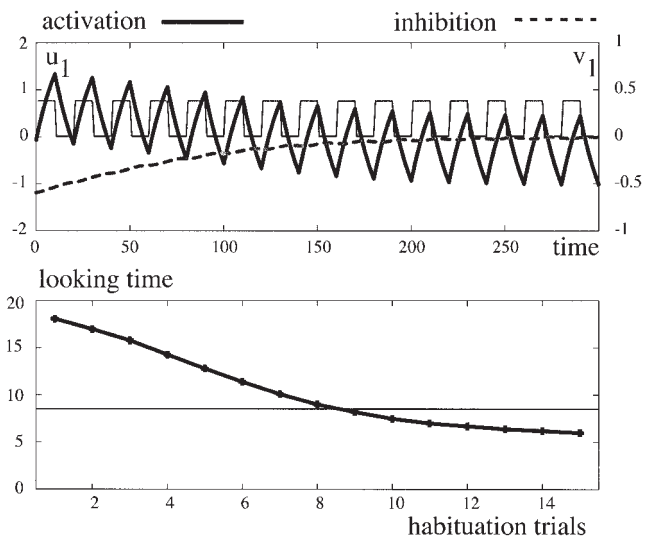


Figure 7. Interaction of fast habituator and increased stimulus strength. Initial resting level, $h_u + 1.5$, stimulus strength is $s^0 = 4.5$.

processes were identical. We show later how these differences in initial levels of activation can also explain differences in dishabituation.

Dishabituation: Simulations of the Alternating Test Paradigm

In infant habituation experiments, infants are almost always dishabituated to two, alternating stimuli. One is usually identical to or similar to the test stimulus; the other is the manipulation of interest. Here we present several simulations of dishabituation, varying the strengths and the differences of the dishabituating stimuli.

Novelty effect. As the first illustration of the model’s treatment of dishabituation, we show the most simple case for which all stimulus strengths remain the same, but there is a completely novel stimulus introduced as a test. This is also an extreme case. In real experiments, it is unlikely that both habituating and dishabituating stimuli would be exactly the same strength and with no overlapping features. In Figure 8, we demonstrate the novelty effect by first habituating the system with 15 repetitions of a medium stimulus strength ($s^0 = 4.5$), at one site. The activation and looking time decrease as inhibition grows. Then we present two alternating test stimuli. The novel stimulus and associated activation and inhibition variables are represented in the bottom panel. In Figure 8, the familiar stimulus comes first and there is no dishabituation to it, but there is some continued looking. Indeed, the first response to the familiar stimulus continues the previous dishabituation curve. Introduction of the novel stimulus, as shown in the bottom panel, gives rise to increased looking, indicated by the gray curve. The second stimulus activates a new site with no accumulated inhibition, and thus looking time increases.

The magnitude of the response depends, however, on the order of the stimuli. In Figure 9, the same conditions hold, but the new

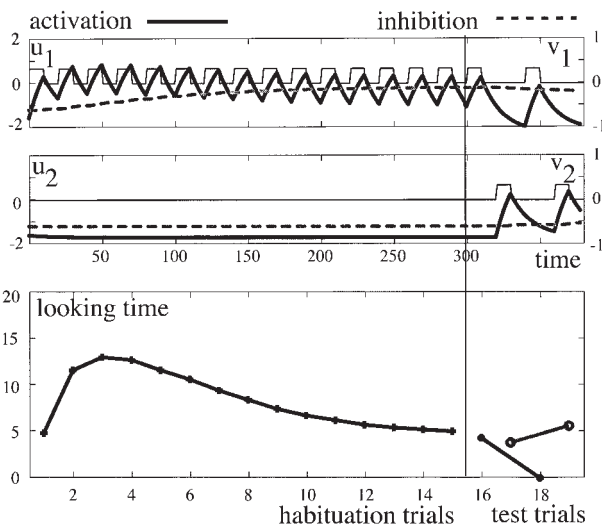


Figure 8. Novelty effect. After 15 habituation trials to a stimulus that activates one neural site (top panel), the second site is activated (bottom panel). When on test the familiar is first, there is a small difference between looking at novel (open gray circles connected by line) and familiar stimulus (solid black circles connected by line).

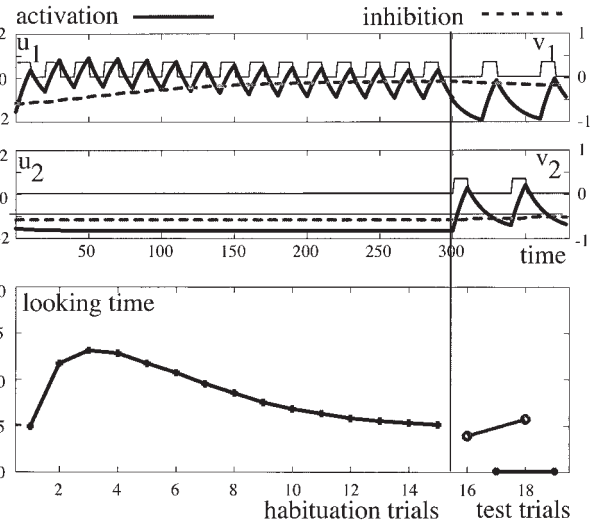


Figure 9. Novelty effect. After 15 habituation trials to a stimulus that activates one neural site (top panel), the second site is activated (bottom panel). When the novel is presented first on test, there is a bigger difference between looking times to novel (open gray circles) and familiar (solid black circles).

stimulus is shown first. It elicits an increased level of looking compared with the familiar stimulus, which is even more deeply habituated as the activation level drops below the resting level because of the longer ISI in which there is no further activation.

Note that in these simulations with independent sites, we do not accomplish dishabituation, as the novel site cannot compete with the overall decrement in activity. We do get a difference in looking time, favoring the novel, with the difference increased when the novel is first. Indeed, many habituation experiments do not report whether there is dishabituation for the test stimuli, but only the difference between Test 1 and Test 2. We later see that true dishabituation to novelty can come about either when the two sites overlap and/or when there is an additional “boost” in activation during the tests. Many habituation experiments meet these additional conditions: There is almost always similarity or “overlap” between habituating and dishabituating stimuli, for example.

If experimenters are comparing, as they often do, the differences in dishabituation to the two test stimuli, they may or may not find the order effect: Here we saw a big difference between novel and familiar if the novel stimulus comes first and a smaller difference if the familiar stimulus comes first. Such an order effect is indeed reported (Baillargeon et al., 1985), but because the stimuli are not defined as “familiar” and “novel” but as “plausible” and “implausible,” the meaning of this effect is discounted. When we recast such experiments in terms of “familiarity” and “novelty” (e.g., Cashion & Cohen, 2000; Schilling, 2000), these order effects become understandable.

Familiarity effect. Experimental evidence converges to show that after only a few habituation trials, infants prefer the familiar stimulus. In Figure 10, we simulate this. The conditions are identical to those in Figure 8, but we give only four habituation trials. In Figure 10, the familiar stimulus is presented as a test first. There is more looking to the familiar because after only a few trials, the system is not deeply inhibited and the test stimulus meets an

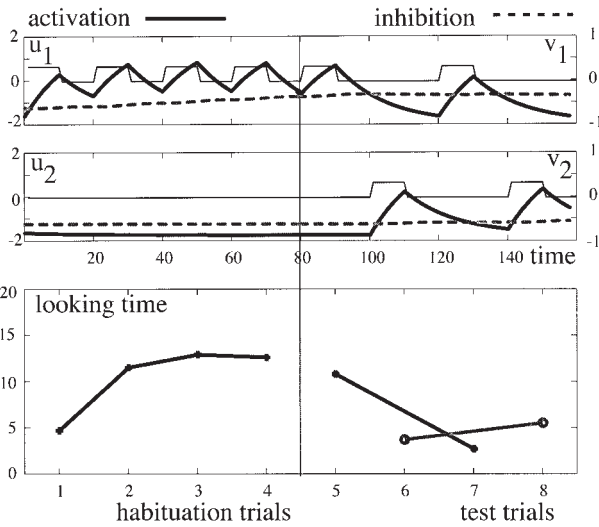


Figure 10. Familiarity effect. After four habituation trials, the second site is activated. On test, the familiar stimulus is presented first and there is more looking to familiar (solid black circles connected by line) than to novel (open gray circles connected by line line).

already highly activated site. Conversely, the novel stimulus gets less looking because it encounters a much less activated system, the same reason that during habituation, looking has an initial rise.

In these simulations, stimulation is not continued until the model reaches criterion, as is obvious from the continued high levels of looking at the last presentation of the habituation stimulus. In that sense, these are simulations of “familiarization” rather than of “habituation.” In the next section, we always present the habituation stimulus until looking time reaches the criterion level and thus model “habituation,” properly speaking.

Finally, we have an order effect for familiarity as well, as illustrated in Figure 11. In this simulation, we use only four habituation trials again but now present the novel stimulus first. As before, the novel stimulus meets an inactivated system. The familiar stimulus hits a system after a longer ISI than in Figure 10 and because there are only four habituation trials, the activation and inhibition are not sustained over the interval and looking is depressed. The difference between familiar and novel is much reduced over the other order.

Note again in Figures 10 and 11 that, in both orders, although there is a difference between novel and familiar, neither stimuli are truly dishabituating. In the case of the fewer habituation trials, inferences are made about preferences based on differences in looking time, without actually dishabituating the infants.

Modeling the Drawbridge Experiments

Thus far we have presented the basic dynamics of habituation experiments with some simplifying assumptions: The habituating and test stimuli are of the same strength, and the novel stimulus shares no features with the habituating one. In the next sections, we make more realistic assumptions to simulate a set of well-known, influential but contested experiments based on Baillargeon’s drawbridge. (Indeed, Baillargeon (1987a) was recently cited as No. 12 in “20 studies that shook up child psychology”;

Dixon, 2003.) The studies are not only important, they are clear, programmatic, and thoughtful. Here we account for the “classic” effects first reported by Baillargeon et al. (1985) and by Baillargeon (1987a), variations reported in Baillargeon (1987b) and Baillargeon (1991), as well as for the manipulations of the design by Rivera et al. (1999) and Schilling (2000). We demonstrate that simple, plausible changes in the parameters of the habituation model can account for all of the effects in these experiments and their control conditions. Results attributed to infants’ conceptual knowledge and reasoning can all be understood by the nonlinear effects of the habituation dynamics.

Recall that in the canonical version of the drawbridge studies, the habituation event is a screen rotating 180° with no block in view (see Figure 12). There are two, alternating dishabituating events, both with the block in view before the screen starts its rotation. One event was considered impossible because the screen continued to rotate 180° and apparently passed through the block. The other is the possible event in which infants saw the screen stop at 112° (Baillargeon, 1987a; 120° in Baillargeon et al., 1985), apparently stopped by the block. Infants looked longer at the impossible event. In the control condition, infants were habituated to the screen as before and then shown either 180° or 112° rotation without the block. In this condition, infants showed no preference. The authors concluded that infants looked longer at the impossible event, not because they preferred the longer rotation, but because they were surprised or puzzled when the screen did not stop for the solid block. All of the experiments in this series follow this basic logic.

Critical to all this class of possible–impossible experiments is that the test stimuli differ from the habituating stimuli only to the degree that indicates the possibility of the event, but not so much that the other presumably nonsalient events can trigger dishabituation. Thus, in the control conditions, it was important to demonstrate that just the change from 180° rotation to 112° (or other

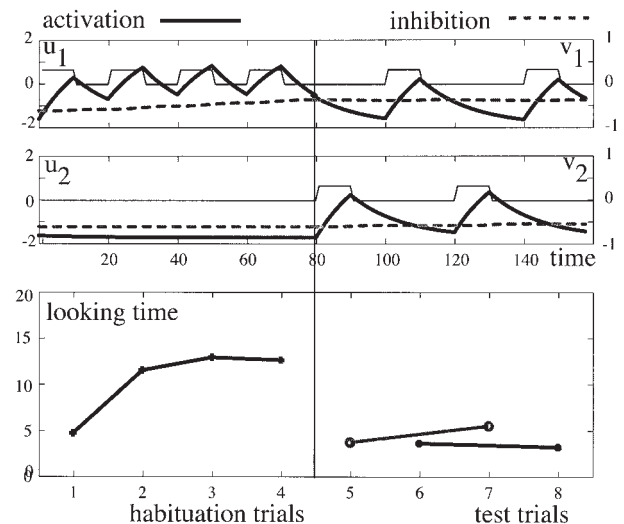


Figure 11. Familiarity effect. After four habituation trials, the second site is activated. On test, the novel stimulus is first and there is less difference between familiar (solid black circles connected by line) and novel (open gray circles connected by line).

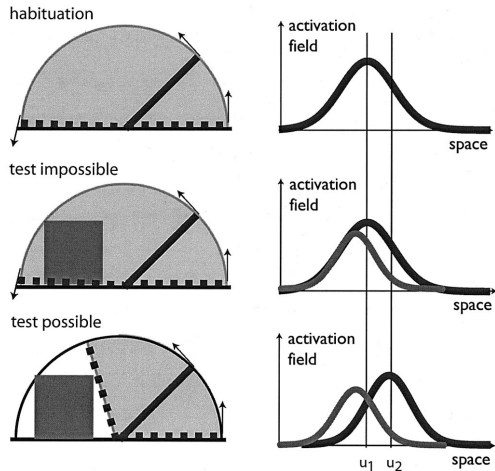


Figure 12. Left panel: The stimuli used in the generic drawbridge paradigm are illustrated. Imagine the infant viewing the display from the right side. The amount of drawbridge motion is reflected by the angle covered by the shaded area. The habituation stimulus (top left) is a 180° motion of the drawbridge. The impossible test stimulus has the same 180° motion in the presence of a block (visible only during the initial phase of drawbridge motion). The possible test stimulus has 112° motion stopping short of the block. Right panel: An activation field driven by these stimuli is sketched. The perceptual dimension depicted here is the space covered by the drawbridge motion and the block. The 180° motion generates broad input (black line), whereas the 112° motion covers a smaller patch of space centered further to the right (bottom right). The spatial position of the block generates input (gray lines in bottom two graphs) that overlaps more with the 180° than with the 112° motion. When the field is sampled at two sites by activation variables u_1 and u_2 , representing the two types of motion, both receive overlapping input from either motion stimulus as well as from the block.

amounts of rotation) was insufficient to lead to dishabituation, and that only the conceptual surprise was effective. In short, the experiments are designed to keep all conditions in the critical test stimulus similar to the habituating stimulus, except those of “cognitive” importance.

Researchers may invent experiments to create “violations of expectancy” but infants experience them as visual events. What are the characteristics of these displays that can lead to differential looking after habituation? First, we show that the effect depends on the addition of the block during both test stimuli. The block, a bright color, and sometimes with a face or design, is an interesting new addition to the visual scene and boosts the strength of activation over the habituation stimulus. This is critical, as we shall see, in producing differences in looking. Moreover, infants typically look at the block for varying amounts of time before the test, further boosting the activation of the system. Second are the perceptual characteristics of the two screen rotations used in habituation and in test. The conditions are different in that the more the screen rotates, the more visual input it provides to the infant. Activation is higher for 180° than for 112°. They also have different stopping points and different velocity profiles, so we assign them different “event” values. At the same time, the screen rotations overlap perceptually as they traverse the same path—for the first 90° or 112° they look the same. Thus, their activations

partially overlap. Third, the timing of these perceptual events matters, especially the number of habituating stimuli, which regulates the depth of habituation, and the order of the test stimuli, which varies the ISIs and thus, the decay of activation.

Simulations of Baillargeon (1987a) and Baillargeon et al. (1985). We begin with the most contentious of the drawbridge experiments, Baillargeon’s (1987a) claim that infants as young as 3.5 and 4.5 months old can use sophisticated reasoning about the permanence and impenetrability of objects. Critics of the drawbridge studies contend that the impossible event garners more looking not because infants detect a violation but because it is more familiar to the habituating event (both screens rotate 180°; Bogartz et al., 2000; Cashon & Cohen, 2000; Schilling, 2000). Our simulations reproduce this familiarity effect, but we also show how fragile it is. In the experiments, 4.5-month-old infants looked more at the familiar (impossible) event but only when they saw it first on test. The 3.5-month-olds preferred the familiar (impossible) event, but only when they were classified as “fast” habituators; “slow” habituators did not detect a difference. These effects fall out of the dynamics of habituation.

Figure 13 illustrates the simulation for the 4.5-month-old infants who got the familiar (impossible) stimulus first.⁴ For this “canonical” version, we set the initial activation value to $h_u = 0.6$, higher than the level used in the generic simulations. Some initial activation is warranted because, in the experiment, infants played with the box and had an initial “familiarization” trial lasting from 10–30 s. The top panel shows the behavior of Site 1, representing the 180° rotating screen, and the middle panel shows the behavior of Site 2, representing the 112° event (see Figure 12 for how these activation sites are positioned in the field). During habituation, the activated site responds to the input by decreasing looking time, reaching criterion in eight trials. Note that because the rotating screens are similar for part of their path, Site 2 shares some activation with Site 1. At test, both sites are boosted to represent the influence of the block. Here Site 1, the familiar stimulus, is the first test, alternating with Site 2. The bottom panel gives the looking time.

This simulation shows a strong familiarity effect. The familiar, 180° test stimulus—boosted by the block—encounters an already activated Site 1, and together, activate a large amount of looking.

⁴ Activation Site 1 receives more input from the 180° stimulus than from the 112° stimulus, whereas the reverse is true for activation Site 2 (see Figure 12). Both sites receive some additional input when the block is present.

For this, and the following simulations, we set the global parameters of the model as follows, except where noted: $h_v = -1.6$; $h_u = -0.6$ (younger infants, $h_u = -0.8$; older infants $h_u = -0.4$); $\beta = 5$; $c_u = c_v = 2$; $\tau_u = 15$; $\tau_v = 600$; initial conditions for activation sites = $h_u + 1.5$, for inhibition sites = h_v ; threshold for looking = 0.5; duration of stimulus-on interval = 10 s; duration of stimulus-off interval = 10s. Noise strength = 0. Habituation criterion = half the average looking time of the first three trials; 3 pairs of alternating test trials. Note that τ_v , the inhibition time scale, is slower than in the earlier simulations so that we can give three pairs of test trials without inhibiting the system completely. We also raised the looking threshold slightly.

The event input to the 180° rotation was 3.24, and to the 112° rotation was 2.40; the rotations share a common input, 0.72; the boost was 1.2; for control conditions, 0.5; and for “extra boost,” 1.3.

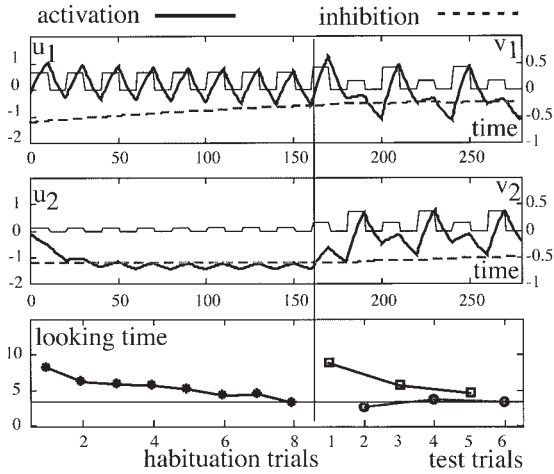


Figure 13. Simulation of the drawbridge experiment. Baillargeon's (1987a) 4.5-month-old infants, familiar (impossible) event first, with preference for familiar. The top and middle panels show stimulus traces (thin lines), activation levels (thick lines) and inhibition levels (dashed lines) at two sites of a neural field as a function of time. The bottom panel shows looking time within each stimulus cycle as a function of the number of stimulus cycles ("trials"). In test trials, solid squares mark looking time to familiar stimuli; open gray circles mark looking time to novel stimuli. For the canonical experiment simulated in this figure: During habituation, Site 1 (top panel) receives the common input + 180° event input, whereas Site 2 (middle panel) receives common input only. During the "familiar" test, Site 1 receives the common input + 180° event input + boost; Site 2 receives the common input + boost. For the "novel" test, Site 1 gets the common input + boost, whereas Site 2 gets common input + 112° event input + boost. Simulated infants needed 8 trials to reach criterion.

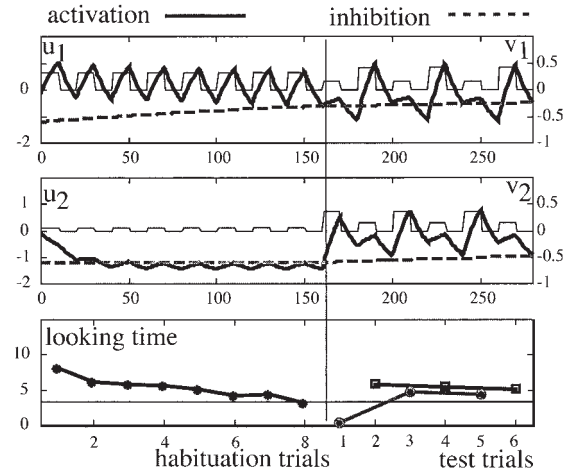


Figure 14. Simulation of the drawbridge experiment. Baillargeon's (1987a) 4.5-month-old infants, novel event first.

that the level of initial activation determines the rate of habituation. Higher initial levels of activation lead to more looking initially, so that the habituation criterion is reached earlier (see Figures 5 and 6), whereas lower initial levels of activation lead to less looking

In contrast, the novel stimulus, coming second, starts from much lower activation and garners little looking. It looks like the infants prefer to look at the impossible event.

But a very different picture emerges when we simply reverse the order of the test stimuli, as shown in Figure 14. The habituation phase is identical. However, in the interval when infants see the novel display first, activation at the familiar Site 1 declines. Now the familiar stimulus does not provide enough summed activation to provide a large boost in looking, and there is no difference between novel and familiar. Infants do not distinguish between possible and impossible. Small changes in timing give big differences in results.

Such an order effect, in which a looking time advantage for the impossible (familiar) stimulus emerged only when that stimulus was presented first during test was reported in all studies that tested for it (Baillargeon, 1987b; Baillargeon et al., 1985; Rivera et al., 1999). In all of these cases the possible event is novel, that is, differs more from the habituation stimulus than the impossible event. Dual-process theory would, therefore, predict the opposite outcome: Presentation of the novel stimulus first should lead to sensitization and thus enhance looking at the familiar (impossible) event when it comes second. As in the experiment, the combination of the two test orders gives an overall familiarity effect (see Figure 15A, Table 1).

Why would slow habituators not look longer at the impossible than the possible display (Baillargeon, 1987a)? We showed above

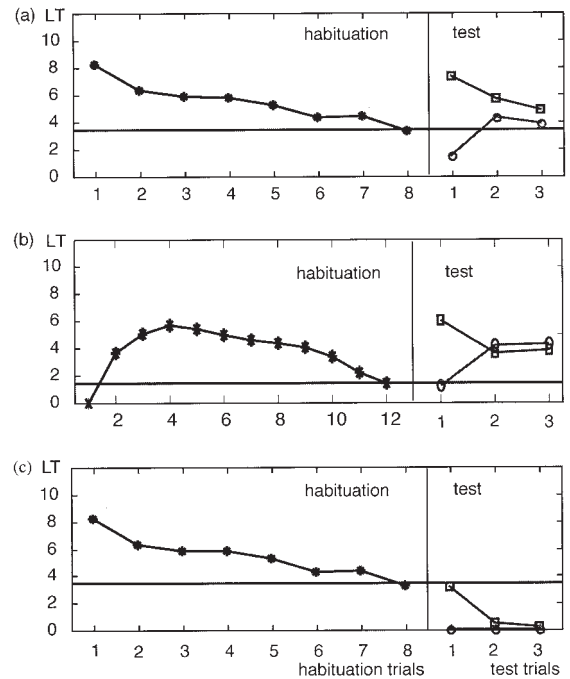


Figure 15. A: Looking time averaged across the two orders of presentation (familiar event first, or novel event first) of 4.5-month-old infants as a function of stimulus cycle. Overall familiarity effect. B: In the same condition, looking times of "slow" habituators, which start with a lower initial level of activation ($= h_{i0}$). This leads to an initial rise of looking time and more looking time to criterion (12 trials). On test, the familiarity effect is strongly reduced. C: Control condition. Boost reduced to 0.5 at test. Reduced familiarity effect over Panel A. Squares = familiar; circles = novel; LT = looking time.

Table 1
 Summary of Simulated Mean Looking Times (LT) at Dishabituation, Averaged Over Three Test Trials and Two Orders

Experiment	Condition	180° impossible (familiar)	112° possible (novel)	Difference, 180°–112°	LT advantage in exp.	Figure
Baillargeon (1987a)	Canonical, 4.5-month-olds	5.91	3.18	2.73	180°	15(a)
	Control	1.26	0	1.26		15(b)
	3.5-month-old, “slow”	6.88	4.59	2.29		15(c)
Schilling (2000)	Long habituation, Condition B	2.87	3.23	–0.35	112°	16
Rivera et al. (1999)	With block	8.83	4.86	3.97	180°	17(a)
	Without block	4.72	0.91	3.81	180°	17(b)
		157° impossible (familiar)	112° possible (novel)	Difference, 157°–112°		
Baillargeon (1991)	80% occlusion of block; 4.5-month-olds	4.53	3.20	1.33		18(a)
		135° event	112° event	Difference 135°–112		
	50% occlusion of block; 4.5-month-olds	3.47	3.00	0.47		18(b)
	50% occlusion of block; 6.5-month-olds	2.13	1.19	0.93		18(c)
	50% occlusion of block with added block; 6.5-month-olds	3.05	1.86	1.18	135	18(c)
Baillargeon (1987b)		Horizontal impossible	Vertical impossible	Difference, impossible–possible		
	Sliding box	6.71	4.11	2.60	horizontal	20(a)
	Rotating box	Analogous	Analogous	2.60	vertical	
		Hart test event	Soft test event	Difference, hard–soft		
	Impossible event 157°	3.13	2.48	0.65	hard	20(b)
	Possible event 112°	1.37	1.31	0.06		20(c)

Note. It is impossible to compare the simulated looking time values with the actual looking times in experiments (exp.) because time differences in experiments are only statistically meaningful for any particular experiment and control. For example, in Baillargeon’s (1987a) Experiment 2, a 4-s difference between tests was statistically insignificant, whereas in Baillargeon’s (1991) Experiment 1, a 4.2-s difference was significant. Likewise, in Schilling (2000) and Rivera et al. (1999), looking time differences of 3 s or less were significant. We reproduce all the directions and metric changes of these effects.

initially and more time needed to reach the habituation criterion (see Figures 2 and 3). As a consequence, at criterion the level of inhibition that has been built is lower for fast than for slow habituators (comparable levels of inhibition exist, for instance, at Trial 4 in Figure 6 and at Trial 14 in Figure 2). At criterion, slow habituators are therefore more deeply habituated than fast habituators. In dishabituation, this reduces the familiarity preference, leading to a smaller looking-time advantage for the familiar “impossible” display (see Figure 15B and Table 1). Note that an account for slow habituation that is based entirely on slower encoding would predict the opposite, that is, more familiarity preference.

What about the control condition in which infants saw only the two different screen rotations without the block? Baillargeon (1986, 1987a, 1987b) argued that the control revealed no preference, but we show that the effect comes from the missing box. For Figure 15B, we provided only a very small boost at test. The familiarity effect is nearly gone (see also Table 1). The strong box input is thus necessary to recover differential looking. The simulations in Figures 15A and 15B reproduce closely Baillargeon’s (1987a) Figure 2.

Now we look at the younger, 3.5-month-old infants. Overall, they displayed no preference for novel or familiar. But the effect reemerges when Baillargeon (1987a) divided the groups into “fast” and “slow” habituators. The fast habituators showed a pattern of response similar to the older infants. Because younger infants

habituate more slowly, it is reasonable to assume that the 3.5-month-old group had some infants who habituated slowly and some who behaved like 4.5-month-olds. The model also explains why the familiarity effect only emerges in the fast (or older) habituators. In Figure 15C, we simply reduced the initial habituation level to $h_u = -0.8$ to simulate a younger infant. This means that it takes longer for the habituation field to build up and decrease looking, a reasonable assumption. Now it takes 12 trials for the infant to habituate to criterion, and the strong familiarity preference and order effects seen in the faster habituators are reduced (see Table 1). When these two groups—fast (older) and slow (younger)—are combined as in the 3.5-month-olds, the net effect might well be no overall main effect for familiarity and no order effect. In Baillargeon (1987a), about 20%–25% of the infants failed to complete all of the test trials, mostly because they were fussy. These may have been the fully habituated babies. Indeed, more slow habituators than fast ones dropped trials, consistent with the theoretical predictions that these infants are more deeply habituated. Note that there is a suggestion of a novelty effect in Baillargeon’s (1987a) slow habituators (her Figures 4 and 5), suggesting these infants are indeed more fully habituated.

Recently Schilling (2000) used the drawbridge paradigm to produce a familiarity to novelty shift. Infants given seven habituation trials preferred the familiar, 180° event, whereas those given 12 trials shifted to look more at the novel 112° event. Likewise,

when we extended the simulated habituation trials for the fast habituators to 20, we produced a novelty preference (see Figure 16).

Overall, our simulations closely reproduce the experimental and control data reported in Baillargeon (1987a). What about the first drawbridge study (Baillargeon et al., 1985)? Methods were similar but the infants were older—5.5 months. In this, the authors also reported a preference for the impossible event, suggesting that these two ages processed the displays similarly. Note that 5.5-month-olds had no order effect in the experimental condition (in contrast to the infants who were a month younger), but the authors did report an order effect in their control condition. This is difficult to interpret because the statistical effect of order (as an interaction with event) in the 1987 study emerged when condition (experimental or control) was included in the analysis while in the 1985 study, the authors did not include condition in a single analysis of variance but analyzed the two conditions as two separate experiments. If they had been combined, an overall order by event effect may have been detected.

Confirming some assumptions. We made two important assumptions for these simulations: first, that the habituating and test drawbridge events shared some overlap but also that the larger, 180° rotation provided more input; second, that the addition of the block boosted activation. There is some experimental support for these assumptions.

Rivera et al. (1999) conducted two experiments to test the hypothesis that infants in the drawbridge studies preferred the impossible event because the increased movement in the 180° event elicited more looking. To test infants' initial preferences, Rivera et al. did not habituate the infants but after three familiarization trials (to the box and each screen event), proceeded directly to eight alternating test trials, thus replicating the test trials in Baillargeon's (1987a) experimental and control conditions. With the block present, infants looked longer at the 180°, impossible event, but there was also an order effect, with a greater preference for the larger rotation when it came first. Without the block, Rivera et al. found a similar 180° preference and order effect, favoring the 180° rotation first.

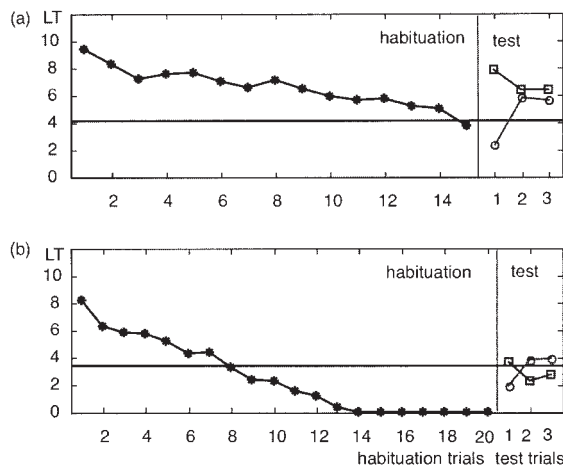


Figure 16. A: Simulation of 3.5-month-old habituators ($h_u = -0.8$). Infants needed 12 trials to reach habituation criterion and have a reduced familiarity effect. B: Familiarity-to-novelty shift. The number of habituation trials was set to 20. Squares = familiar; circles = novel; LT = looking time.

We simulated Rivera et al.'s (1999) experimental results, using our original parameters, including the assumption that the 180° rotation activates looking more than the 112° rotation. We simulated the four pairs of test trials. Figure 17 reports averages over both orders for the block and no-block experiments, showing a large looking preference for the 180° rotation (see Table 1). Both conditions also give the predicted order effect, with a large advantage for seeing the 180°, impossible, event first (with block: looking time differences, familiar first = 5.7, novel first = 2.2; without block, familiar first = 5.5, novel first = 2.1). We also simulated (see Figure 17) the overall reduced looking to both stimuli with repeated trials over the test phase which is reported in Rivera et al. (1999) and in Baillargeon et al. (1985).

The Rivera et al. (1999) experiments make perfect sense theoretically. Without a long habituation period, the alternating test trials activate both sites but do not allow for much inhibition to build up, especially because the effective ISI for each site is twice as long. The 180° rotations provide stronger activation and continue to do so throughout the pairs of trials, especially when they come first and that site gets and keeps an activation advantage. Because activation is high and habituation is low, the boost is not needed to maintain looking. However, there are two small differences between the simulations and experiments. First, in the simulations without a boost, looking to the 112° event goes to zero and it does not in the experiment. Again, this is a function of our looking threshold. Second, Rivera et al. actually found a switch to 112° looking preference in later trials when that rotation comes first, and we did not. We think this is a function of the difference between the infant-controlled procedure used in experiments and our fixed trials in simulation (see the Discussion section, below.) Specifically, infants in their experiment had very long looks in their first test trials (because they were not habituated) and thus became more deeply habituated than in our fixed-trial simulations. This led to a later novelty switch in the experiments. Again, this is theoretically consistent.

There is also evidence for our assumption of an activation boost with the addition of the block. In their careful parametric study of effects, Bogartz et al. (2000) found that the presence of the block did enhance looking but only in those infants who were "further along in processing the rotating screen component" (p. 421)—that is, more deeply habituated. This is entirely consistent with the Rivera et al. (1999) results and our simulations. When the system is highly activated, looking is maintained without the block, but when habituation becomes more dominant, the boost is needed to recover looking. Cashon and Cohen (2000) also found an overall looking preference for trials that included a block. These results are also consistent with recent studies demonstrating that infants are very sensitive to the simple quantity of stimulus presented to them (Clearfield & Mix, 1999). Thus, we feel confident in our assumption that the block added intrinsic interest to the display and increased the activation of the test stimulus in a nonspecific way.

Infant Reasoning or Habituation Dynamics?

Baillargeon has used the basic drawbridge design in a series of other experiments to demonstrate infants' high levels of representation of objects and their reasoning about object properties. We now show that these experiments are easily simulated by the habituation model with minimal and plausible adjustments to the

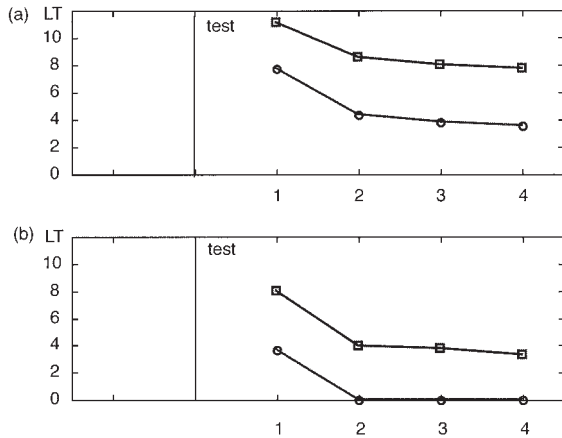


Figure 17. A: Simulation of Rivera et al. (1999), Experiment 1. There was no habituation, and four alternating test trials with boost (block). More looking to 180° event (squares) than at 112° event (circles). B: Same as in Panel A, but with no boost at test. LT = looking time.

parameters used to simulate the first experiments. As before, we view the task as a series of perceptual events subject to basic habituation dynamics.

Simulating Baillargeon (1991). We first address the studies reported in Baillargeon (1991). In the first set, infants were habituated to the usual 180° rotating screen and tested on a possible and impossible event. In the possible event, the screen rotated and stopped at the box, as before, a 112° rotation. In the impossible event, the screen rotated through either the top 80% or the top 50% of the box. (This corresponds to screen rotations of 157° and 135°, respectively.) She found that the 6.5-month-old infants looked longer at the 80% occlusion of the box but not at the 50%. Infants who were 4.5 months old did not look longer at either event. Thus, she reasoned that the older infants could represent the existence, height, and location of the hidden object, but only when the screen passed through most of the box. In the 50% condition, infants had limited ability to represent the object or to “compute the screen’s stopping point” (Baillargeon, 1991, p. 24). Note again that the impossible event is more familiar because it overlaps more with the 180° habituating event.

To simulate this effect, we proportionately reduced the input to the 180° site on the test trials, reflecting less motion of the screen. We first consider Baillargeon’s (1991) Experiment 2, in which 4.5-month-old infants, who detected a violation in the canonical procedure, were tested in the 80% condition. Figure 18A simulates the results with the 180° screen rotation input reduced to 2.94. With just this small decrease in input, the familiarity effect is noticeably reduced, compared with the conditions shown in Figure 15 (see Table 1). Thus, the young infants seem incapable of detecting the 80% violation. The familiarity effect is diminished even more dramatically when we reduced the input to stimulate the 50% condition (see Figure 18B and Table 1).

Now we turn to the 6.5-month-old infants in the same 50% situation. Older infants habituate more quickly. To simulate older infants, we raised h_u to -0.4 , creating stronger habituation. Here we see the familiarity effect significantly reduced compared even to 4.5-month-old infants (see Figure 18C vs. Figure 15A). When

the box moved only 50% of its arc, even older infants lost the ability to detect the impossible event.

In a second set of experiments, Baillargeon (1991) then gave infants a second box, placed next to the rotating display. She reasoned that the second box, which remained always in view, would act as a reminder of the block that becomes occluded by the rotating screen. Alternatively, infants could compare the heights of the blocks and predict when it should stop. Indeed, providing the extra block improved 6.5-month-old infants’ performance on the 50% rotation condition such that they favored the impossible event, the more familiar 135° rotation over the possible 112° rotation. Adding the second block increased looking time differences from 1.0 s to 5.4 s. The added block seemed to help infants reason about the occluded object.

Again, this effect is simply reproduced by considering the always-visible block as an extra boost of activation to the system during the dishabituation trials. Accordingly, we ran the same simulation as in Figure 18 (the older infants at 50% rotation), but

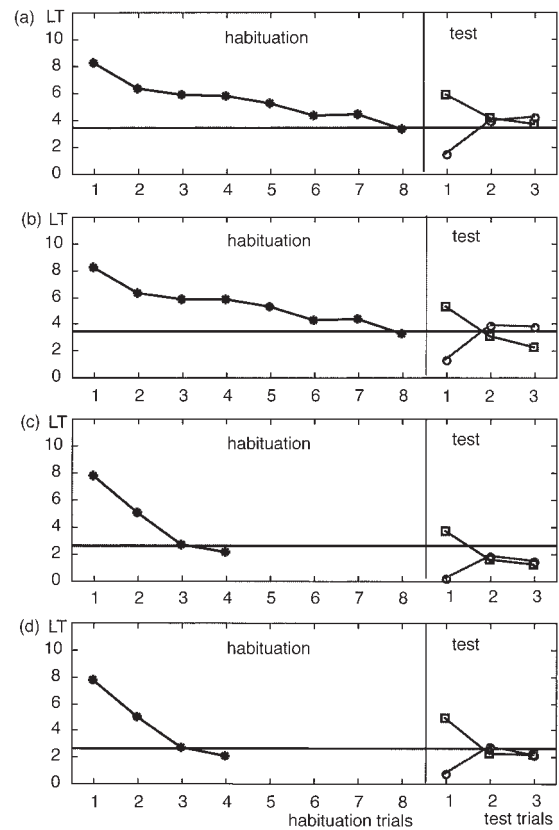


Figure 18. A: Simulation of the 80% rotation in 4.5-month-old infants. The event input at test reduced to 2.94, leading to reduced familiarity effect. B: 50% rotation; same as in Panel A, but event input at test was reduced to 2.74. Familiarity effect is more strongly reduced. C: Simulation of 50% rotation in 6.5-month-old infants. Same as in Panel B, but resting level is increased to $h_u = -0.4$ to model older infants. This reduces the amount of looking needed to criterion as well as to a partial restoration of the familiarity effect. D: Simulation of 50% rotation in 6.5-month-olds with two blocks. Same as in Panel C, but with boost increased at test to 1.3. Further restoration of familiarity effect compared with Panel C. Squares = familiar; circles = novel; LT = looking time.

we increased the boost from 1.2 to 1.3. This magnifies the familiarity effect from a looking difference of 0.939 without the boost to 1.18 with the extra boost (see Figure 18D and Table 1).

Simulating Baillargeon (1987b). Now we tackle the experiments with the most plausibly “cognitive” content. In her Experiment 1, Baillargeon (1987b) compared the typical rotating screen event with a “sliding” screen event. In the new event, the screen rotated 90°, similar to the first half of the rotating screen event (see Figure 19). But then the screen stopped and the upright screen slid backward, away from the viewing infant. For each screen condition, there were two test events using a long, thin box that could be positioned either vertically, like a tower, or horizontally, like a sheet. For the sliding box, the impossible event was when the horizontal box was behind the screen because it would block the sliding motion of the screen. The vertical box was possible, in

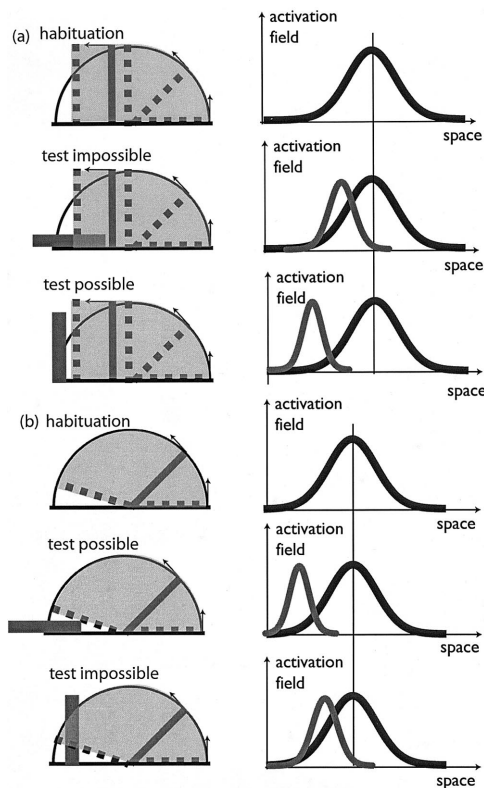


Figure 19. Illustration of events in slide versus rotation experiments (left column) together with corresponding illustrations of inputs to an activation field representing the space covered by motion and boxes (right column). The stimulus would be seen by the infants from the right. The sliding or rotating screen motion (shaded in light gray or dark gray) is similar across habituation and test trials. During test trials, a box either lies horizontally or stands vertically; (A) the sliding screen the horizontal box is impossible; (B) with the rotating screen, the vertical box is impossible. The field receives input representing the space covered by the screen motion, which is similar during habituation and test trials. The box provides input that overlaps more with the motion input when the box shares space with the movement (impossible) than when the box is spatially separate (possible). The site at which the field is sampled (vertical line) is already activated by the habituation stimulus when the box provides more overlapping input in the impossible conditions than in the possible conditions.

contrast, because the box could fully slide before contact the box. For the rotating screen, the opposite was true: The horizontal box was so thin that the screen could rotate to nearly 180°, but the tower-like block would realistically get in the way of the screen (see Figure 19). Seven-month-old infants looked longer at the impossible events in both conditions, suggesting that they could reason about the location and size of the occluded object.

In this experiment, we cannot rely on familiarity preference to explain the differences in looking times because the movements of the screens are identical in both test trials, and both are slightly less than in habituation. Rather, we invoke the different visual properties of the boxes as the infant sees them during test trials in relation to the movements of the screens in the habituation trials. Specifically, the boxes share to varying extent the same space as that covered by the moving drawbridge or slide. First consider the slide condition (see Figure 19A) The figure illustrates on the right the input generated by the three stimulus conditions for an activation field defined over the space covered by slide and box. The horizontal box occupies more of the space covered by the moving slide than the vertical box, leading to more input to the field sites already activated by the slide motion. When the field is sampled by a single activation site (vertical line), the horizontal box contributes more input than the vertical box. The opposite holds true for the drawbridge (see Figure 19B). Here, the vertical box covers more of the same space as the moving drawbridge than the horizontal box and thus contributes more input to already activated parts of the field.

To simulate these differences, we increased activation at test for the horizontal box in the slide condition and the vertical box in the rotation condition. In addition, to account for the slightly reduced movements of the screen at test, we reduced the 180° event input by 0.3. Figure 20A shows this simulation for the slide condition. There is more looking to the horizontal (impossible) event (see Table 1). Note that the rotating condition is exactly the same as the canonical condition, with the horizontal box similar to the possible and the vertical box, the impossible, events. We have already demonstrated a looking advantage to the vertical (impossible) event.

We also uncovered a strong order effect, with increased looking to the horizontal box when it came first (looking time differences: horizontal first = 2.1, vertical first = 1.7). Indeed, this exactly replicates Baillargeon’s (1987b) effect: “Infants in the sliding screen condition who saw the horizontal-box event first looked reliably longer overall than those who saw the vertical-box event first” (p. 189). Although Baillargeon accepts her predicted main effects, she ignores these, and other, order effects. Indeed, order effects cannot be explained by reasoning or knowledge accounts; if infants “really know” something, the order of the test should not matter. But order effects arise naturally as a function of the nonlinear habituation dynamics and depend on the sensitive balance between the strength of activation and inhibition, and the parametric nature of the test stimuli.

Our final simulation of the drawbridge series addresses Baillargeon’s (1987b) Experiment 2. This study tested 7-month-olds’ understanding of the compressibility of objects: Would they understand that the drawbridge could squish a soft object as it rotated but could not pass through a solid one. For the “impossible” experiment, infants were habituated to two events in which a screen rotated 157° and compressed two, soft, ball-shaped objects

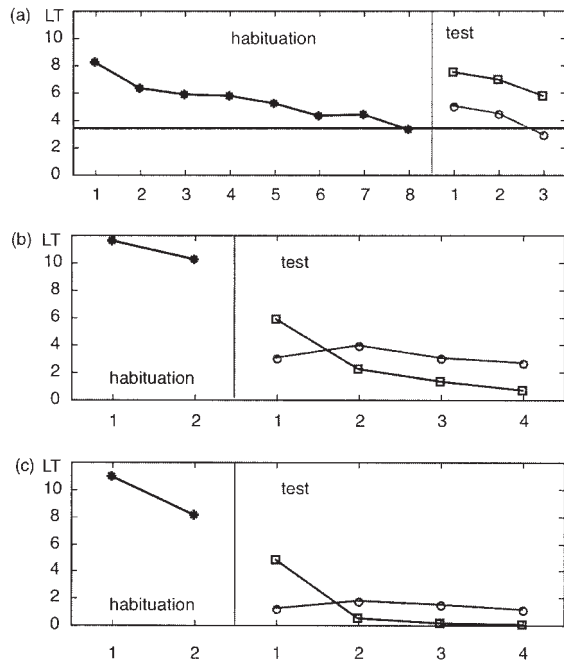


Figure 20. A: Slide condition. Both sites have identical input during habituation and test (practically a single neuron model). The two tests have different boosts. Common input = 0.72; event input for 180° event = 3.24; for 112° event = 2.40; 180° event at test = 2.94. Boost 1 (horizontal box) = 1.2; Boost 2 (vertical box) = 0.5. During habituation, both sites get the common input + the 180° event input. For Test 1, the horizontal box, both sites get the common input + the 180° test event input + Boost 1. For Test 2, the vertical box, both sites get the common input + the 180° event input Test 2 (possible). B: Soft versus hard: Impossible condition. Site 1 has the event input for soft; Site 2 has the event input for hard. The boost of activation from the object is included in the common input for both habituation and test because infant sees an object in both. Common input is 1.66. Event input (soft) = event input (hard) = 3.0. During habituation, Site 1 gets common input + event input (soft) and Site 2 gets just common input. There are only two habituation trials. During Test 1 (soft, looking times marked by squares), Site 1 gets common input + event input (soft) and Site 2 gets just common input. During Test 2 (hard, looking times marked by circles), Site 1 gets common input and Site 2 gets common input + event input (hard). C: Soft versus hard: Possible condition. Same as Panel B, but common input is 1.46 instead of 1.66, reflecting shorter drawbridge motion. LT = looking time.

made of stuffing or plastic bags. For the test, the screen rotation was the same, but in one test, a hard block was substituted for the soft object, making it impossible for the screen to complete the full rotation. In the “possible” experiment, the screen rotated only 112°. Again, a hard and soft object alternated for test, but both events were possible. There were just two habituation trials.

When shown the impossible event, the infants looked longer at the hard than the soft object. In the possible experiment, infants tended to look longer at the soft object, although this effect was not conventionally significant.

In these simulations, we assumed older, faster habituating infants than the canonical studies ($h_u = -0.4$), and we reduced the habituation trials to only two, so the infants were not at habituation criterion. Furthermore, we assigned differential inputs to the ro-

tating screen because infants saw more screen rotation in the impossible experiment. We used two sites, one for the soft event and one for the hard, such that the site received added activation (“3”) only when the infant saw the specific object in the trial. In other words, for the impossible experiment, during habituation and the soft object test, Site 1 had input from the common screen motion plus the soft event. Site 2 had the common screen motion but no event input. For the hard event test, Site 1 had only screen movement and Site 2 had both motion and event input. The logic was the same for the possible experiment, but the screen rotation input was less.

Figures 20B and 20C show how infants might react differentially to soft and hard events. After two habituation trials to soft events, with the “impossible” screen rotation, infants showed a small looking advantage to the hard event (see Figure 20B and Table 1). With only the reduction in screen rotation (“possible” screen rotation), this difference was noticeably reduced (see Figure 20C and Table 1). Baillargeon (1987b) provided no information about the actual looking times for this experiment, so we do not know how 7-month-old infants react to only two habituation trials. Her Figure 4, her data, and our model indicate that infants were still strongly habituating during the test trials. Recall that screen rotations are the same for both habituation and test trials. We suggest, then, that infants presumed understanding of object permanence arises from a slight novelty boost from the hard object during continued habituation combined with the extra activation from the larger screen rotation, which enhances the small novelty effect. An appropriate control experiment would be to habituate with the hard object and test with the soft and hard.

A coherent, parsimonious account. In her response to the challenges to the drawbridge experiment, Baillargeon (2000) warned against using any one experiment to dismiss the many demonstrations of early infant competence. (Aslin, 2000, made a similar argument.) She also cautioned about piecemeal explanations:

To demonstrate at this point that young infants cannot represent occluded objects, a scientist would need to show that, for each report in Table 1 [A list of 30 experiments showing 3- to 7-month-old infants can represent fully hidden objects] explanations based on perceptual factors provide a more parsimonious interpretation of the finding. Furthermore, a scientist would need to consider explanations across sets of related experiments. Arguments that Finding 1 is best attributed to Factor A, Finding 2 to Factor B, Finding 3 to Factor C, and so on, where A, B, and C are all generated post hoc with little or not attempt at a theoretical integration, could never be favored over accounts that provide a single coherent explanation of large bodies of findings. (Baillargeon, 2000, p. 448)

We agree with Baillargeon (2000) that a coherent theory is needed to understand a large body of data, and we believe the dynamic field theory can meet the challenge. For all experiments using visual habituation, whatever they are called, infants watch repeated events and then are tested with slightly different events. This looking and not-looking is governed by general processing dynamics. We have demonstrated here that such dynamics, with simple assumptions about the properties of the events, can provide a unified explanation for four of the 30 studies listed in Baillargeon (2000). Would the remaining 26 be exempt from these considerations? Whatever the status of infants’ abilities to represent objects, we believe that the habituation dynamics create the effects.

Discussion

We have presented a model of infant habituation based on a simple and plausible mechanism. We constructed the model on well-known facts about habituation itself: Preferences for familiarity and novelty depend on the depth of habituation, strength and similarity of the stimuli, and the timing of the stimulus events. Using this model, we explored the dynamics of the process of habituation and dishabituation and showed they could simulate well-known experimental results.

Because we could manipulate the model parameters precisely, the modeling was useful for pointing out the complex and subtle interactions of activation and inhibition as they evolve over time and with different stimulus regimes. Small changes in model parameters such as stimulus strength, stimulus timing, and initial conditions made a big difference in whether infants dishabituate or not and in whether they looked longer at novel or familiar events. We saw these effects despite several simplifying assumptions about the nature of the stimuli, the timing of looking and looking away, and the role of stochastic processes. In real experiments, these factors would tend to make the outcomes more, not less, difficult to interpret as they would amplify the nonlinearities inherent in these coupled processes. We discuss the implications of these assumptions next.

Assumptions and Limitations

Fixed stimulus intervals. In the model, stimulus onset and offset are set as fixed parameters. This is a simplifying assumption because most habituation studies now use the so-called “infant-controlled procedure,” in which the stimulus is left on as long as the infant continues looking, usually defined as not looking away for 2 s or more. When the infant does look away, experimenters provide some attention-getting mechanism to bring the infant’s gaze back to the display for the next habituation trial. These attention-getters range from calling the infants’ names; to a curtain going up and down; to bright visual displays, sometimes with sound. At the infants’ reengagement, the habituating stimulus is presented again, until looking away commences, and so on, until the infants’ looks reach some habituation criterion. Thus, the timing of looks and looks away is controlled by the infants’ own behavior. It is important to note that although looks to the stimulus are recorded, looks away are almost never considered.

As we saw in the model simulations (and in experiments by Kaplan and colleagues; see Kaplan, Werner, & Rudy, 1990), not only is looking time itself critical, but so are the intervals between stimulus presentations in the habituation phase, between habituation and tests, and between the test phases. The build-up and decay of activation and the corresponding rates of inhibition are all strictly time dependent. Indeed, when infants look a long time, activation and the concurrent habituation both build rapidly. Very strong initial activation, however, can maintain looking for a long time. In such a case, the fixed interval model may turn off the stimulus while the infant is still looking. Conversely, infants with very short looks may be looking away while the stimulus is still on.

In Figure 21, we present a first effort to model the infant-controlled procedure by adding a stimulus to return the infants’ gaze to the display after they have looked away, so that the initiation and termination of the trials is not fixed. Here we give the

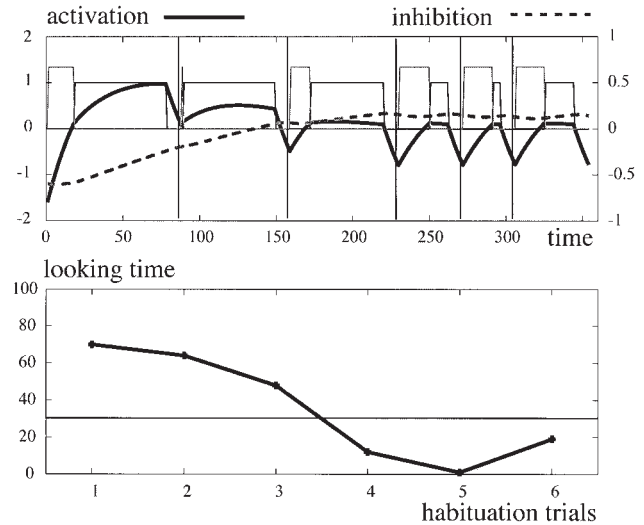


Figure 21. Infant-controlled procedure. Trial begins with an “attraction stimulus” (open gray rectangles) at strength 4. When infant passes looking threshold, habituation stimulus (open black rectangles) given at strength 3. $\tau_u = 30$.

model two stimuli. The first is the “attraction” stimulus (shown as rectangles in light gray at the beginning of the trials). These are set to 4 and are highly activating. When the system reaches the threshold for looking (activation = 0.1), the attraction stimulus is switched off and the habituation stimulus is turned on (activation strength = 3) for no more than 60 s. To accommodate the increased overall activation the time scale was slowed to $\tau_u = 30$; other parameters are the same as in Figure 2. Under this regime, initial activation is high because of both activating stimuli, but inhibition also grows rapidly. The attracting stimulus still recovers looking to the display, but the habituating stimulus itself gets less and less looking time. Looking time and trial length both decrease, which is more similar to procedures most often used now.

However, taking the model to this next step of a strictly infant-controlled procedure is not so straightforward because we do not yet fully understand what determines the initiation and duration of looking away nor what draws the infant back to looking. For example, the field model suggests that the ISI may be a function of the previous activation, as well as each infant’s threshold and particular activation–inhibition dynamics. If looking and looking away are determined by a balance between activation and habituation, there should be a relationship between the duration of the looks at and the durations of the looks away. A simple one is that the longer the look, the more excitation, thus, the more inhibition of looking. However, as the model also suggests, excitation and inhibition build and decay on different time scales, yet unknown. Thus, the effects of varying levels of excitation may not be manifest in the looks away until some time lag, also unknown.

But there are additional complications in the infant-controlled procedure that we know little about. One possibility is that looking away is driven by some competing visual stimulus, as there is always something else to look at. Then the process would involve a decision between two stimuli of differing degrees of salience in two different locations, both with their own dynamics of activation

and inhibition. Equally unknown is the effect of the “attention getter” on the coupled dynamics of activation and inhibition during the ISI or how the attention stimulus reinforces or competes with the habituating stimulus. Even when the attention-getting event is very different from the habituating stimulus, it may have a more general activating function, as dual-process theory suggests. This would, in effect, provide an already-refreshed system for the habituating stimulus. Because the effective ISIs between attention-getting stimuli are quite long (looks plus looks away) infants may not habituate to them. At the very least, the attention-getting events provide complex and competing cycles of activation and inhibition.

Again, we have little or no data that speak to these issues. Looking time is measured, but looking-away time is not, although these data should be available or can be readily collected. However, if the dynamics are as complex as we suspect, uncovering the time structure may need much longer runs of looking and looking away than are usually allowed. The role of the attention-getters—visual or auditory—has not been considered in experiments. The implicit assumption is that because they are the same in all conditions, they should not mask experimental effects. However, given the subtle and nonlinear nature of the coupled processes, this assumption may not be warranted.

These issues regarding what turns looking on and off were discussed by Cohen and Menten in 1981 and more recently by Gilmore (2001), who constructed a linear model to simulate the time structure of habituation. Although he was able to produce a pattern of looking and looking away, Gilmore’s best simulation could only produce one long look and then a series of fixations of the same duration, a kind of “one-trial” habituation. Although our model was more successful in reproducing a realistic pattern of habituation, we agree with Gilmore that progress in modeling now depends on gathering more data about the time course of memory and its decay, as well as what pulls attention back to the habituating stimulus.

Representations of the stimulus. The second simplifying assumption in the dynamic field model was to represent both the habituating and dishabituating stimuli as two discrete sites in a continuous field. This simplification allowed us to assign values for the relative strength of the stimuli and their relative similarity in terms of overlap. Changes in these two dimensions alone or together resulted in different outcomes in terms of rate of habituation and response to novelty.

In reality, of course, stimuli are much more complex, and how infants represent the features and events they see is of core importance. Over the years, researchers have tested infants in the habituation paradigm on an enormous range of stimuli, from simple static displays to understand perceptual processes (e.g., Caron & Caron, 1969; Clearfield & Mix, 1999) to elaborate events (Spelke, Katz, Purcell, & Erlich, 1994) with presumptive rich cognitive content. Indeed the heart of the debate is what these stimuli actually mean to the infant. Are infants merely watching colorful objects in their visual fields, or are they making inferences from the displays to what they “really know” about how the world works?

Infants provide this answer by a single behavioral measure: changes in looking time. So it is central to the current debate to understand the parametric nature of the dishabituating stimuli. How different is novel and how similar is familiar? Investigators

interested in the psychophysics of infant vision have used habituation to characterize just how much difference infants can detect in nearly every aspect of the visual scene: edges, motion displays, patterns, orientation, flicker frequency, three dimensions, contrast, and so on (reviewed recently in Kellman & Arterberry, 1998). These meticulous experiments are designed and conducted to be devoid of any “cognitive” content. Displays are precisely specified, and content neutral, so that the exact nature of the differences is known and controlled, and researchers generally do not make any inferences about the infants’ conceptual understanding of them.

In contrast, when habituation is used to ask about what infants know, not what they see, the displays are often very complex, the events are not well-specified, and the magnitudes of the differences between habituation and tests are not quantified. The information gathered by the vision researchers is not brought to bear. Indeed careful parametric experiments by Dannemiller and colleagues (Dannemiller, 1998; Ross & Dannemiller, 1999), using preferential looking have shown that color contrast, luminance, and movement all compete to attract infants’ attention. This means that the addition of a red block, or the movements of a hand gloved in gold lame, are not trivially related to strength of the habituating and dishabituating stimuli.

The model pointed to the critical importance of more parametric studies of stimulus distance and intensity, parallel to, and in combination with, studies that manipulate the time structure of the habituation and dishabituation events. For example, in the more “cognitive” style experiments, once infants are habituated to one stimulus, they are rarely tested on parametric differences in dishabituating stimuli.⁵ (One suspects that some of this is done in pilot experiments that go unreported.) This is even more an issue when complex events are used, as stimulus movement is certainly part of its activating value, as shown explicitly by Rivera et al. (1999).

Along with more experiments probing just what differences make a difference, we must extend the model to probe the interactions of time and stimulus parameters when the stimuli are multidimensional and extend in time. Field concepts can be extremely useful in this regard. In a continuous field of n dimensions, stimulus qualities can be expressed by location in the field, distance and overlap with other stimuli, relative intensity, and a metric of “sharpness” or spread over sites in the field (see Erlhagen & Schöner, 2002; Schutte et al., 2003; Spencer, Smith, Thelen, 2001; Thelen et al., 2001). Time of onset and offset can also be defined, as well as the dynamics of activation and decay. If the stimuli are not too complex, experimenters can precisely define the feature space and use habituation experiments to map detectable differences at different ages. This would motivate further experiments combining features to ask about generalization among several stimulus qualities (e.g., Cohen, Gelber, & Lazar, 1971), early category formation (e.g., Quinn & Johnson, 2000), and the role of movement.

Noise. In real habituation experiments, only averaged group data are shown. But infants are notoriously variable in this task,

⁵ Baillargeon (1991) is an exception. She varied the amount of rotation on the test and showed differential outcome.

and very few infants actually look like the mean data. Thus our assumptions in the model are likely too simplistic.

As an exercise to mimic more realistic conditions, we used the same parameters as the simulations in Figure 2–4, but with the noise increased. Figures 22, 23, and 24 show these simulations as 10 individual runs and the mean superimposed, as a function of the increasing strength of the stimulus.

The good news is that the overall patterns of habituation remain. The bad news for experimental interpretation is that there is a dramatic effect of the stimulus strength on the degree of individual variability. When the events are presumably compelling to the infants, we might expect reliable group effects in which the mean values represent the differential outcomes.

Individual infants' performance. But what happens when the displays are intrinsically less activating? Here there is considerable variability due to noise, and presumably, any effects detected would be more fragile. The interaction between stimulus strength and variability has not been investigated, but as with other unknown parameters, may have important effects on outcomes and their interpretations.

Comparisons With Other Formal Models of Habituation

We began this enterprise to adapt dynamic field concepts, which have proved useful for understanding infant reaching, to another important process, visual habituation. We have demonstrated the power of a specific dynamic field model to explain habituation both in a theoretical (abstract) sense and also when applied to a large set of theoretically important data. Here we compare our dynamic model to other models of habituation proposed over the last 30 years. Sirois and Mareschal (2002a, 2002b) provide a more extended review of habituation models.

As noted in the introduction, there have been two major lines of reasoning about habituation. One is derived mainly from animal

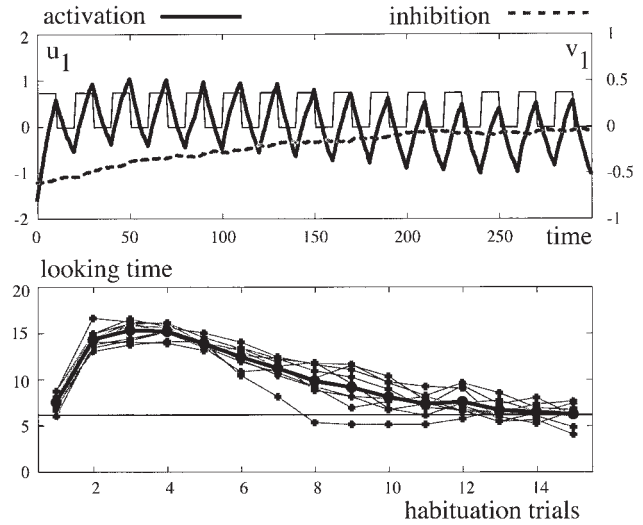


Figure 23. Same as Figure 3, but noise added (noise level $q_u = 0.025$ and $q_v = 0.005$).

experiments and is exemplified by Thompson and Spencer's (1966) dual-process theory. The second comes from Sokolov (1963) and dominates thinking about human infant habituation. Most contemporary models are versions of these main types.

Habituation Models in the Tradition of Dual-Process Theory

Recall that dual-process theory is about the content-free and general processes that lead to response decrement and recovery. Briefly, Thompson and Spencer (1966; see also Groves & Thompson, 1970) proposed that stimuli provoke both specific inhibition and general arousal, and they suggested specific neural mechanisms that may produce known habituation effects. Responses are

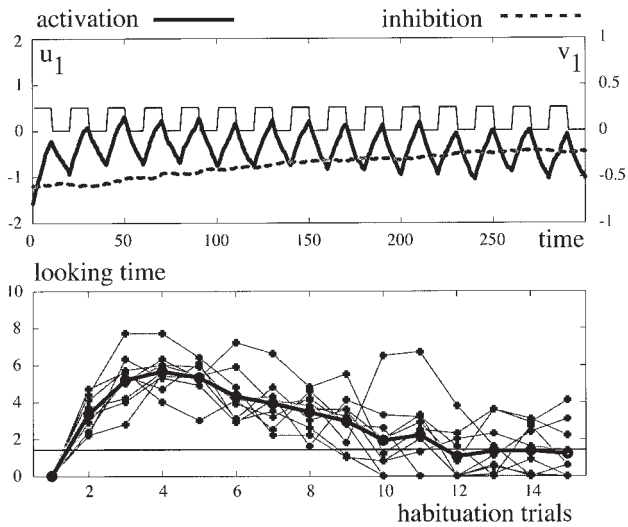


Figure 22. Same parameters as in Figure 2, but noise added (noise level $q_u = 0.025$ and $q_v = 0.005$). Top panel: Activation, inhibition, and stimulus traces are shown for only a single run. Bottom panel: Looking times from 10 runs (crosses linked by thin black lines) and their mean values (solid circles, thick black line) are shown.

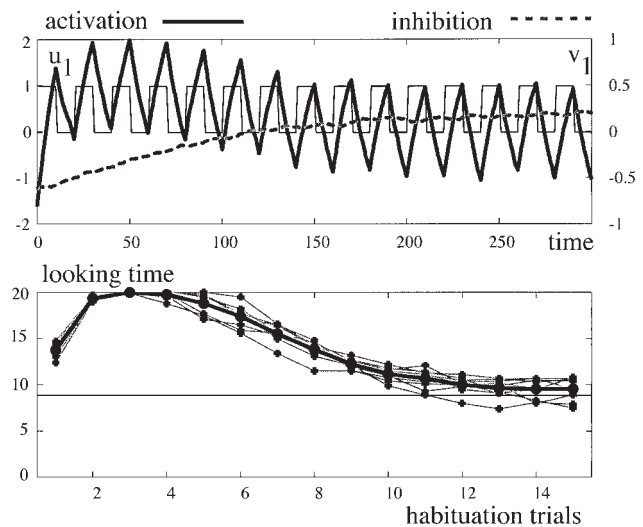


Figure 24. Same as Figure 4, but noise added (noise level $q_u = 0.025$ and $q_v = 0.005$).

determined by the sum of two independent pathways. One is specific to the stimulus and decreases responding. The other channel is sensitizing. At weak stimulus strengths, inhibition is greater than excitation, leading to habituation. But with a novel, strong stimulus, excitation wins out and the response is restored.

Stanley (1976) offered a computer simulation based on dual-process theory. Stanley's circuit contains two links between the stimulus input and the behavioral response. One is the direct route, reflecting the specific decrement with repeated presentations of the stimulus. This is independent of the stimulus strength. The second, indirect route, is mediated by a internal state neuron, representing the generalized activation of the system. The direct route has its synaptic strengths weakened over continued stimulation, although the indirect route is strengthened proportional to the strength of the stimulus. The indirect process is assumed to be faster than the direct one, so for strong stimuli, there will be an early phase of increased response followed by a late phase of decreased response, a signature of dual-process accounts. For weak stimuli, the initial strong response is not seen. Thus, Stanley's (1976) model successfully replicates the time course of habituation as a function of stimulus strength. So does a similar model by Wang and Arbib (1995; Wang, 1995), who also added a "recovery" period and showed restored responding after a delay, similar to experimental results.

These dual-process-based models have serious weaknesses for dealing with infant habituation. They do not address dishabituation and its complexities. Because the stimulus-specific pathway is only inhibitory, the models do not build up representations in the Sokolovian sense. When the novel stimulus reexcites the network indirectly through its increased arousal value, it does not specifically correspond to the input. In contrast, the dynamic field model correctly simulates both habituation and dishabituation. Moreover, because two sites can be activated simultaneously and directly during both habituation and dishabituation, we have a mechanism for establishing relative overlap between them. Although we invoke two processes—excitation and inhibition—they are directly linked to stimulus properties.

In their recent reviews, Sirois and Mareschal (2002a, 2002b) called dual-process-inspired models "function estimators" because they only reproduce the exponential decrease of responses. We agree that the models are insufficient to fully characterize infant habituation, but we are sympathetic to the general process assumptions of dual-process theories, which are not incompatible with the dynamic field ideas.

Sokolov-Inspired Models

The second class of formal habituation models follow Sokolov (1963). Recall that Sokolov proposed that infants pay attention to a stimulus while they are constructing an internal representation of it. When the representation is complete, they look away. Their interest is rekindled only as the new stimulus is different from the old one and therefore doesn't match the constructed representation. Thus, the major concern of Sokolovian-style models is building the internal representation, basically making the internal representation look like the stimulus.

A symbolic model. The most explicit version of a Sokolov (1963) account is Simon's (1998) simulation of numerical competence using a version of ACT-R cognitive architecture, a sym-

bolic production system. Simon sought to reproduce Wynn's (1992) experiment demonstrating early numerical abilities but without building in a "counting" mechanism. To do this, he created a series of elaborate rules for representing the location, identity, and other attributes of objects in the world. When objects are occluded, the infants creates a mental "index" of those objects. When the new display differs from the infant's current index, the infant compares the state of the world in memory to that actually seen. Infants look longer at events, such as secretly added, removed, or changed objects, when they do not match the internal predictions because the processing time is longer for mismatches than for matches. Thus, there is a preference for the impossible display.

Simon bills his model as a "violation of expectancy" account, and indeed, habituation is not involved at all. The longer looking for unexpected events is a function entirely of the mental matching process, which is predetermined by many built-in rules. There is no attempt to model the time course of looking in repeated presentations. Individual differences and the complexities of familiarity and novelty shifts are ignored. Furthermore, the model does not generalize to habituation when nothing is occluded. Thus, we still need a general process to account for all of the habituation studies in which the stimulus is changed at test without any tricks or assumptions of possibility and impossibility.

Connectionist models. Because such symbolic architectures are brittle and domain specific, connectionism has become the most common theoretical framework for Sokolov-inspired habituation models. Here theorists construct a representation of the stimulus by changing the weights in a neural network through a learning process. When the output of the network stabilizes, it has completed the representation. Looking depends on a comparison between the output of the network and the stimulus, decreasing as the output increasingly resembles the stimulus.

The most simple connectionist models of infant habituation use standard feed-forward neural networks with a hidden layer. For instance, Schafer and Mareschal (2001) modeled Stager and Werker's (1997) speech sound discrimination task using such a network. Stager and Werker showed that 8-month-olds were capable of phonemic discriminations whereas 14-month-olds were not. Schafer and Mareschal's network learns to reproduce the input (specific associations) by using the input signal to train the output units. The network compresses information into an internal representation and then expands it to reproduce the original input. The output is then compared with the original stimulus to compute a network error. The greater the error, the authors assume, the greater the looking time. To simulate the experiment, Schafer and Mareschal trained the network with an artificial language, more extensively for 14-month-olds than for 8-month-olds. Then, they habituated the network to a single association (habituation, defined as decreased error) and tested the network with a novel association. They showed that the "younger" networks had better discrimination of the novel pair than the "older" networks. The authors speculated that more language exposure actually makes it more difficult for the network to learn the habituation pair. Younger networks are more flexible, they suggest, whereas older networks, because of their history, cannot learn so well. This makes the age shift in discrimination not so much a change in cognitive strategy as suggested by Stager and Werker but, rather, a function of the process of language-learning itself.

How well does this and similar connectionist networks (e.g., Shultz & Bale, 2001) reproduce what researchers know about infant habituation? First, the network accounted for the decrease of responding over time, but this is inevitable, given that the learning rule depends on decreasing error. Similarly, the network shows a novelty effect because when a new stimulus is introduced, the error goes up again. The model does not account at all for a familiarity effect, as the authors recognize: There is no mechanism for increased looking to a familiar stimulus. There is also an issue with the mapping of error to looking time as a large error can lead to faster learning (Sirois & Mareschal, 2002b). Finally, the results on phonemic discrimination are counter to the well-established fact that older infants habituate faster than younger ones to the same stimuli.

Associator networks are a class of recurrent neural networks that learn to make a one-to-one map from the input to the output. These networks consist of interconnected units, capable of mutually adjusting their weights until they settle on stable values representing the input. Kohonen (1977) has used an associator network to simulate habituation by adjusting the weights so that the units inhibit each other and the output decreases. Here the goal is not to copy the input, but rather, to reduce the activation to zero, in a sense, learning to habituate. The Kohonen network is called a “novelty filter” because novel stimuli would not be inhibited and therefore could reactivate looking. The novelty filter captures the decrease of responding over time but cannot show a familiarity-to-novelty shift, because as in the other network models, the response is always decreasing.

Hybrid Models

Two recent models incorporate features of both dual-process and Sokolovian aspects of habituation. Gilmore’s (2001) model uses a single variable, “memory strength,” that controls looking in an algorithm similar to a dual-process architecture. Initially, a new stimulus leads to looking. When memory strength reaches an upper threshold, the system “looks away.” Looking is restored if the stimulus is present (direct path) and the memory strength is weak (indirect path). If memory strength reflects the degree of completion of the internal representation, then Gilmore’s model contains both theoretical constructs. He accounts for the decay of responses, and for alternating periods of looking and looking away. However, the model so far has only one memory variable, so it cannot handle stimulus specificity, generalization, or the familiarity to novelty switch. Moreover, Gilmore does not simulate a real data set. A much simplified model, which essentially fits a time-function template to looking time curves, was used however, by Gilmore and Thomas (2002) to characterize individual differences.

Finally, Sirois and Mareschal (2002a) proposed a network habituation model that uses both an autoassociative memory—a “cortical sheet”—and a novelty filter, presumably the subcortical element. The authors consider this a dual-process model: The first subsystem builds up the familiar representation and the second filters it. By making the autoassociative memory learn faster than the novelty filter, responses to constant stimuli decrease faster than responses to variable ones, which the authors interpret as a shift from familiar to novel.

However, unlike classic dual-process models, Sirois and Mareschal (2002a) cannot account for the initial rise in looking seen in

many infant habituation studies as their network uses only a decreasing error signal. We have no information regarding how it handles parametric differences in habituating and dishabituating stimuli, nor the subtle order and boost interactions we describe above. Moreover, the authors have not yet tested the model against the data from a real infant habituation experiment.

In sum, there are several promising models on the table. The most successful seem to need some sort of dual process. We believe that the dynamic field model has many strengths compared with the others:

1. The model incorporates elements of both dual-process theory and Sokolovian comparisons within a single dynamic process. In particular, stimuli can be represented both in terms of metric distances and in activation strength, resulting in nonlinear interactions similar to those in real experiments.
2. The model simulates all the known effects of both the habituation curves and the subtleties of dishabituation, including the familiarity-to-novelty shift, boosts, and order effects, as well as individual differences.
3. We have simulated a large and theoretically important set of experiments. We show how the dynamics of the events produce outcomes previously ascribed to deliberate cognitive strategies.
4. The model is a general process account, potentially able to explain a wide range of data sets. Connectionist models, in contrast, are designed to construct highly specific mental pictures.

Conclusions

When a 4-month-old infant is seated before a display, the same mechanisms of visual processing, habituation, and dishabituation are engaged, regardless of what the experimenter believes the display is about. According to our model, an interesting visual display should produce an increase of looking until the inhibitory dynamics win out, when infants disengage. We believe that these processing dynamics cannot be ignored. Those who claim their displays are just about content—be it number, object properties, or causality—must, we also believe, tell us at what point a display ceases to obey basic visual and attention processes and crosses a line into pure meaning. Perhaps this can be done, but the challenge has so far not been addressed.

We see a parallel with our recent treatment of the well-known “A-not-B error.” Since the task was first described by Piaget (1937/1954), researchers have assumed that it was about infants’ developing an object concept. Our models and experiments (Diedrich, Highlands, Spahr, Thelen, & Smith, 2000; Diedrich, Thelen, Smith, & Corbetta, 2001; Smith, Thelen, Titzer, & McLin, 1999; Thelen et al., 2001) have shown that it is about the dynamics for reaching for targets. As in habituation, repeated reaches (or looks) create a time-based memory for the ongoing activities that biases the system. When the infant is given the test (reach for *B* or look at the new display), the behavior that emerges at that time is the dynamic product of the child’s immediate history. In the A-not-B task, we discovered that this history included their visual

behavior of looking at the targets and also their bodily memory of reaching itself and its postural context. Manipulations that had nothing at all to do with object representations, such as weighting the infants' arms or changing their postures, created infants who looked as though they had or did not have an object concept. The dynamics of reaching applied whether infants were reaching for hidden objects or visible ones. By similar logic, the dynamics of habituation apply whether infants are viewing cognitively neutral dots on a background or meaning-laden displays depicting abstract notions of containment or support.

The danger of ignoring the nature of habituation itself is that experiments can take on a particular ad hoc quality that is always open to misinterpretation. For example, there are a nearly infinite number of displays that adults might interpret as being about containment or support or object permanence. The model showed that the balance between familiarity and novelty preferences can be changed by rather subtle differences in the stimuli or the timing of the events. If the parameters are not constrained by theory or mechanism, it is possible to adjust the stimulus displays to produce nearly any outcome one desires. What are the limits on the parameters? How are they decided upon? How are they modified? Can other experimenters design similar experiments with opposite results? The recent challenges to the drawbridge experiment (Bogartz et al., 2000; Cashon & Cohen, 2000; Rivera et al., 1998; Schilling, 2000) demonstrate that this danger is real.

Habituation remains one of the most intriguing and useful windows to the mental processes of infants. The dynamic model points out some ways that researchers can systematically increase the understanding of habituation itself and thus how it changes during development. We suggest that until these dynamics are more thoroughly understood, the procedure must be used with care.

References

- Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27, 77–87.
- Aslin, R. N. (2000). Why take the cog out of infant cognition? *Infancy*, 1, 463–470.
- Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month-old infants. *Cognition*, 23, 21–41.
- Baillargeon, R. (1987a). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology*, 23, 655–664.
- Baillargeon, R. (1987b). Young infants reasoning about the physical and spatial characteristics of a hidden object. *Cognitive Development*, 2, 179–200.
- Baillargeon, R. (1991). Reasoning about the height and location of a hidden object in 4.5- and 6.5-month-old infants. *Cognition*, 38, 13–42.
- Baillargeon, R. (2000). Reply to Bogartz, Shinsky, and Schilling; Schilling; and Cashon and Cohen. *Infancy*, 1, 447–462.
- Baillargeon, R., Spelke, E., & Wassermann, S. (1985). Object permanence in 5-month-old infants. *Cognition*, 20, 191–208.
- Bashinski, H., Werner, J., & Rudy, J. (1985). Determinants of infant visual attention: Evidence of a two-process theory. *Journal of Experimental Child Psychology*, 39, 580–598.
- Bastian, A., Schöner, G., & Riehle, A. (2003). Preshaping and continuous evolution of motor cortical representations during movement preparation. *European Journal of Neuroscience*, 18, 2047–2058.
- Bogartz, R. S., Shinsky, J. L., & Schilling, T. H. (2000). Object permanence in five-and-a-half-month-old infants? *Infancy*, 1, 403–428.
- Bogartz, R. S., Shinsky, J. L., & Speaker, C. J. (1997). Interpreting infant looking: The Event Set \times Event Set design. *Developmental Psychology*, 33, 408–422.
- Bornstein, M. H. (1989). Stability in early mental development: From attention and information processing in infancy to language and cognition in childhood. In N. A. Krasnegor (Ed.), *Stability and continuity in mental development: Behavioral and biological perspectives* (pp. 147–170). Hillsdale, NJ: L Erlbaum.
- Bornstein, M. H., & Benasich, A. A. (1986). Infant habituation: Assessments of individual differences and short-term reliability at 5 months. *Child Development*, 57, 87–99.
- Caron, R. F., & Caron, A. J. (1968). The effects of repeated exposure and stimulus complexity on visual fixation in infants. *Psychonomic Science*, 14, 78–79.
- Caron, R. F., & Caron, A. J. (1969). Degree of stimulus complexity and habituation of visual fixation in infants. *Psychonomic Science*, 10, 207–208.
- Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy*, 1, 429–446.
- Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, 10, 408–411.
- Clearfield, M. W., & Westfahl, S. M.-C. (in press). Familiarization in infants' perception of addition problems. *Journal of Cognition and Development*.
- Cohen, L. B. (1988). An information-processing approach to infant cognitive development. In L. Weiskrantz (Ed.), *Thought without language* (pp. 211–228). Oxford, England: Clarendon Press.
- Cohen, L. B. (2002). Extraordinary claims require extraordinary controls. *Developmental Science*, 5, 211–212.
- Cohen, L. B., Gelber, E. R., & Lazar, M. (1971). Infant habituation and generalization to differing degrees of stimulus novelty. *Journal of Experimental Child Psychology*, 11, 379–389.
- Cohen, L. B., & Menten, T. G. (1981). The rise and fall of infant habituation. *Infant Behavior and Development*, 4, 269–280.
- Colombo, J., Frick, J. E., & Gorman, S. A. (1997). Sensitization during visual habituation sequences: Procedural effects and individual differences. *Journal of Experimental Child Psychology*, 67, 223–235.
- Colombo, J., Mitchell, D. W., O'Brien, M., & Horowitz, F. D. (1987). The stability of habituation during the first year of life. *Child Development*, 57, 474–487.
- Dannemiller, J. L. (1998). A competition model of exogenous orienting in 3.5-month-old infants. *Journal of Experimental Child Psychology*, 68, 169–201.
- DeLoache, J. S. (1976). Rate of habituation and visual memory in infants. *Child Development*, 47, 145–154.
- Diedrich, F. J., Highlands, T., Spahr, K., Thelen, E., & Smith, L. B. (2001). The role of target distinctiveness in infant perseverative reaching errors. *Journal of Experimental Child Psychology*, 78, 263–290.
- Diedrich, F. J., Thelen, E., Smith, L. B., & Corbetta, D. (2000). Motor memory is a factor in infant perseverative errors. *Developmental Science*, 3, 479–494.
- Dixon, W. E. J. (2003). 20 studies that shook up child psychology. *Developments: Newsletter of the Society for Research in Child Development*, 46, 6, 13.
- Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, 109, 545–572.
- Fantz, R. L., & Fagan, J. F. (1975). Visual attention to size and number of pattern details by term and preterm infants during the first six months. *Child Development*, 46, 3–18.
- Frick, J. E., Colombo, J., & Saxon, T. F. (1999). Individual and developmental differences in disengagement of fixation in early infancy. *Child Development*, 70, 537–548.
- Gibson, E. J., Owsley, C. J., & Johnston, J. (1978). Perception of invariants

- by five-month-old infants: Differentiation of two types of motion. *Developmental Psychology*, 14, 407–415.
- Gilmore, R. O. (2001, April). *Models of the learning process that habituation indexes*. Paper presented at the annual meeting of the Society for Research in Child Development, Minneapolis, MN.
- Gilmore, R. O., & Thomas, H. (2002). Examining individual differences in infants' habituation patterns using objective quantitative techniques. *Infant Behavior & Development*, 25, 399–412.
- Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review*, 77, 419–440.
- Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, 21, 167–179.
- Hunter, M., & Ames, E. (1988). A multifactorial model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 5, p. 69–95). Norwood NJ: Ablex.
- Hunter, M., Ames, E., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, 19, 519–529.
- Hunter, M. A., Ross, M., & Ames, E. (1982). Preferences for familiar and novel toys: Effects of familiarization time in 1-year-olds. *Developmental Psychology*, 18, 519–529.
- James, W. (1899). *Principles of psychology* (Vol. 1). New York: Holt.
- Johnson, S. P., & Aslin, R. (1995). Perception of object unity in 2-month-old infants. *Developmental Psychology*, 31, 739–745.
- Kaplan, P. S., & Werner, J. S. (1986). Habituation, response to novelty, and dishabituation. Tests of a dual-process theory. *Journal of Experimental Child Psychology*, 42, 199–217.
- Kaplan, P. S., Werner, J. S., & Rudy, J. W. (1990). Habituation, sensitization, and infant visual attention. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 6, pp. 61–109). Norwood, NJ: Ablex.
- Karmel, B. Z. (1969). The effect of age, complexity, and amount of contour on pattern preferences in human infants. *Journal of Experimental Child Psychology*, 7, 339–354.
- Kellman, P. J., & Arterberry, M. E. (1998). *The cradle of knowledge: Development of perception in infancy*. Cambridge, MA: MIT Press.
- Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, 15, 483–524.
- Kellman, P. J., Spelke, E. S., & Short, K. R. (1986). Infant perception of object unity from translatory motion in depth and vertical translation. *Child Development*, 57, 72–86.
- Kohonen, T. (1977). *Associative memory: A system-theoretical approach*. New York: Springer-Verlag.
- Kopecz, K., & Schöner, G. (1995). Saccadic motor planning by integrating visual information and pre-information on neural, dynamic fields. *Biological Cybernetics*, 73, 49–60.
- McCall, R. B., & Kagan, J. (1967). Attention in the infant: Effects of complexity, contour, perimeter, and familiarity. *Child Development*, 38, 939–952.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J., & Cook, M. (Trans.). (1954). *The construction of reality in the child*. New York: Basic Books. (Original work published 1937)
- Pinker, S. (1994). *The language instinct*. New York: Morrow.
- Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy*, 1, 31–46.
- Rivera, S. M., Wakeley, A., & Langer, J. (1999). The drawbridge phenomenon: Representational reasoning or perceptual preference? *Developmental Psychology*, 35, 427–435.
- Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, 1, 491–507.
- Rose, S., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology*, 18, 704–713.
- Ross, S. M., & Dannemiller, J. L. (1999). Color contrast, luminance contrast and competition within exogenous orienting in 3.5-month-old infants. *Infant Behavior and Development*, 22, 383–404.
- Schöner, G., Dose, M., & Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16, 213–245.
- Schöner, G., Kopecz, K., & Erlhagen, W. (1997). The dynamic neural field theory of motor programming: Arm and eye movements. In P. G. Morasso & V. Sanguineti (Eds.), *Self-organization, computational maps and motor control, psychology series* (Vol. 119, pp. 271–310). Elsevier-North-Holland.
- Schafer, G., & Mareschal, D. (2001). Modeling infant speech sound discrimination using simple associative networks. *Infancy*, 1, 7–28.
- Schilling, T. H. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. *Infancy*, 1, 389–402.
- Schutte, A. R., Spencer, J. P., & Schöner, G. (2003). Testing the dynamic field theory: Working memory for locations becomes more spatially precise over development. *Child Development*, 74, 1393–1417.
- Shultz, T. R., & Bale, A. C. (2001). Neural network simulation of infant familiarization to artificial sentences: Rule-like behavior without explicit rules and variables. *Infancy*, 2, 501–536.
- Simon, T. J. (1998). Computational evidence for the foundations of numerical competence. *Developmental Science*, 1, 71–78.
- Sirois, S., & Mareschal, D. (2002a). Infant habituation: A review of current computational models and a new proposal. In J. A. Bullinaria & W. Lowe (Eds.), *Proceedings of the 7th Neural Computation and Psychology Workshop: Connectionist models of cognition and perception* (pp. 90–103). Singapore: World Scientific.
- Sirois, S., & Mareschal, D. (2002b). Models of habituation in infancy. *Trends in Cognitive Sciences*, 6, 293–298.
- Slater, A., Morison, V., & Somers, M. (1988). Orientation discrimination and cortical function in the human newborn. *Perception*, 17, 597–602.
- Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting: The task dynamics of the A-not-B error. *Psychological Review*, 106, 235–260.
- Sokolov, E. (1963). *Perception and the conditioned reflex*. New York: Pergamon.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605–632.
- Spelke, E. S., Katz, G., Purcell, S. E., & Erlich, S. M. (1994). Early knowledge of object motion: Continuity and inertia. *Cognition*, 51, 131–176.
- Spelke, E. S., & Newport, E. L. (1998). Nativism, empiricism, and the development of knowledge. In R. M. Lerner (Ed.), *Handbook of child psychology. Vol. 1: Theoretical models of human development* (5th ed., pp. 275–340). New York: Wiley.
- Spencer, J. P., Smith, L. B., & Thelen, E. (2001). Tests of a dynamic systems account of the A-not-B error: The influence of prior experience on the spatial memory abilities of 2-year-olds. *Child Development*, 72, 1327–1346.
- Stager, C. L., & Werker, J. F. (1997, July 24). Infants listen to more phonetic details in speech perception tasks than in word-learning tasks. *Nature*, 388, 381–382.
- Stanley, J. C. (1976, May 13). Computer simulation of a model of habituation. *Nature*, 261, 146–148.
- Thelen, E., Schöner, G., Scheier, C., & Smith, L. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Brain and Behavioral Sciences*, 24, 1–33.
- Thelen, E., & Smith, L. B. (1994). *A dynamic-systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

- Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*, 16–43.
- Wang, D. (1995). Habituation. In M. A. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 441–444). Cambridge, MA: MIT Press.
- Wang, D. L., & Arbib, M. A. (1992). Modeling the dishabituation hierarchy in toads: The role of the primordial hippocampus. *Biological Cybernetics*, *67*, 535–544.
- Wynn, K. (1992, August 27). Addition and subtraction by human infants. *Nature*, *358*, 749–750.
- Wynn, K. (1995). Infants possess a system of numerical knowledge. *Current Directions in Psychological Science*, *4*, 172–177.

Received April 2, 2002

Revision received February 20, 2003

Accepted May 12, 2005 ■

Instructions to Authors

For Instructions to Authors, please consult the January 2006 issue of the volume or visit www.apa.org/journals/rev and click on the “Instructions to authors” link in the Journal Info box on the right.