ELSEVIER
Cognitive Development 17 (2002) 1323–1343
COGNITIVE DEVELOPMENT

A constructivist model of infant cognition

Leslie B. Cohen*, Harold H. Chaput, Cara H. Cashon

Department of Psychology, University of Texas, SEA #A8000, Austin, TX 78712, USA

Abstract

We propose six Information-Processing Principles (IPPs) that together describe a constructive, hierarchical system by which infants come to understand objects and events in the world around them. We then demonstrate the applicability of these principles to four specific domains of infant perception and/or cognition, (i.e., form perception, object unity, complex pattern perception, and understanding of causal events). In each case empirical developmental changes appear to be consistent with the IPPs. We then present the Constructivist Learning Architecture, a computational model of infant cognitive development. This model is based on the IPPs, and uses self-organizing, neurally based techniques from Kohonen (1997) and Hebb (1949). We then apply the model to the complex domain of infant understanding of causal events, and replicate many of the developmental changes found empirically. Finally, we discuss the applicability of this constructivist approach to infant cognitive development in general.

Keywords: Information-Processing Principles; Constructivist Learning Architecture; Infant cognition

The study of infant cognition has yielded an impressive amount of research over the past two decades. Unfortunately, this research has also produced a diverse, and sometimes contradictory, pattern of results. A close examination of most of the evidence, however, does indicate a coherent view of infant cognitive development — one that describes changes in infant cognition as primarily bottom-up and hierarchical in nature, in other words, a constructive view of infant cognition. In this article, we shall propose such a constructive view by outlining a set of Information-Processing Principles (IPPs) of infant development, detailing

* Corresponding author. Tel.: +1-512-471-0189; fax: +1-512-471-5935.
E-mail addresses: cohen@psy.utexas.edu (L.B. Cohen), chaput@cs.utexas.edu (H.H. Chaput), cashon@mail.utexas.edu (C.H. Cashon).

0885-2014/02/$ – see front matter
PII: S0885-2014(02)00124-7

several studies that illustrate these principles, and presenting a computational model that provides a quantitative instantiation of the view consistent with the empirical evidence.

Although the field of developmental psychology has come a long way since the days of J. B. Watson (1913) and his extreme *tabula rasa* view of newborn infants, the nature–nurture debate continues to be with us in one form or another (see Elman et al., 1996; Haith, 1998; Spelke, 1998). Whereas even the most extreme nativist would admit that development must be governed to some extent by environmental influences and even the most extreme environmentalist would admit that infants must have some inborn capabilities, there still is no agreement about what is inborn and what must be acquired. In fact, to some a constructivist model of infant development may seem a misguided exercise given many relatively recent reports that offer a nativist view. However, within the last few years these nativist views increasingly have come under fire, and alternative explanations are appearing in the literature. The infant cognitive literature now includes a detailed examination of infants' apparent understanding of certain physical laws such as solidity (Cohen, 1995; Cohen, Gilbert, & Brown, 1996), object permanence (Cashon & Cohen, 2000), causality (Cohen & Amsel, 1998; Cohen & Oakes, 1993), and most recently addition and subtraction (Cohen & Marks, 2002). In each case, the newer studies indicate that simpler perceptual and attentional processes can explain the apparent precocious performance of young infants, and that assumptions regarding innate core knowledge or pre-built modules may be unnecessary or even misleading.

The view of infant cognitive development that we propose depicts infants developing their knowledge about the world by way of a continuous interplay between a set of domain-general learning mechanisms and changing environmental experiences. As we have in the past (see Cohen, 1998; Cohen & Cashon, in press), we are labeling this view "an information-processing approach" because in the main, we are attempting to explain how infants process information in their environment and how that processing changes with age.

In some respects our view resembles, or at least has been inspired by, Piagetian theory. For example, we assume that development can be described as a bottom-up, or constructivist process according to which infants initially process simple perceptual units. These simple units then become integrated into more complex, higher order units which themselves become integrated into yet higher order units. Thus, as in Piagetian theory, development follows a hierarchical progression. In addition, this bottom-up approach is embodied in our computational model, which integrates Piagetian constructivism (e.g., Piaget, 1952) with Hebbian learning (Hebb, 1949). The model is consistent with basic principles of neuroscience, chief among them being the Hebbian notion that lower order units (e.g., cell assemblies) that are correlated in their activity tend to become integrated into higher-order units (e.g., phase sequences).

Our constructive view of infant cognitive development has evolved over the last two decades from an attempt to explain and unify several different topics within the areas of infant perception and cognition. These topics range from how infants process a simple 45° angle to how they come to understand a complex causal event. In attempting to find some common developmental pattern we arrived at a set of IPPs, which seem to apply to a wide variety of topics being acquired by the infant over a wide range of ages (Cohen & Cashon, 2001, in press). Although there is evidence that our principles can account for empirical findings in many different domains, we do not mean to imply that these principles represent the only learning mechanism available to the infant.

In the next section, these principles will be presented followed by empirical evidence indicating their applicability to research findings generated from our laboratory, as well as findings from other laboratories. In the subsequent section, we will present a specific computational model based upon these principles. This model will attempt to simulate a particularly complex problem, the development of infants' understanding of a causal event. We shall also be assessing the model in terms of its ability to follow the IPPs we derived from empirical evidence. In the final section of the article, we shall discuss further additional applications of the model and its potential for understanding a wide range of topics within the general area of infant cognitive development.

1. Information-Processing Principles

Our approach to infant cognitive development can be summed up by the following six IPPs. As we mentioned earlier, we arrived at these principles from an examination of developmental changes in several different topics considered to be aspects either of infant perception or infant cognition. Some of the principles may appear to be self-evident; others undoubtedly will be more controversial.

1.1. Infants are endowed with an innate information-processing system

Infants are born neither with a blank slate nor a preponderance of innate core knowledge. Rather, we would argue that infants are born with a system that enables them to learn about their environment and develop a repertoire of knowledge. From the outset, the innate system provides architectural constraints in how this learning may be accomplished. The system is designed to allow the young infant access to low-level information, such as orientation, sound, color, texture, and movement.

1.2. Infants form higher units from lower units. In other words the learning system is hierarchical

As the infant learns and develops, information that is accessed becomes more and more complex, building upon prior processed information. An assumption underlying this principle is that the ability to process more complex information is the result of being able to integrate the lower-level units of information into a more complex, higher-level unit. That integration is based upon statistical regularities or correlations in activity of those lower-level units. The term "unit" is used as a loose description of information that can be processed independently or coupled with other units to form a more complex whole.

So, for example, an infant may initially process the two lines of a 45° angle as separate units in particular orientations, but because the two lines co-occur in the same relative spatial relationship, even when the angle is rotated, the infant will eventually process the relationship among the lines, that is, the angle rather than the independent lines (Cohen & Younger, 1984).

1.3. Higher units serve as components for still-higher units

The hierarchical nature of the system can account for development beyond the first few months of life. The process of integrating information to form a higher-level unit is itself repeated throughout development. Lower-level information can be integrated into a higher-level unit, which can in turn be integrated into an even higher-level unit, and so on.

To continue with the angle example, after connections have been formed between the two lines to form an angle, several angles and curves could then be integrated to form the complex shape of an object. That object could then be integrated with another object to form an event defined in terms of the relationship between the two objects.

1.4. There is a bias to process using highest-formed units

Whereas the previous principles have described the learning mechanism, or the building of the hierarchy, this fourth principle describes what information an infant will attempt to process after two or more layers of information units have been formed. Specifically, infants will tend to process the incoming information using the highest level available to them. This does not mean that the lower-level information is unavailable, but rather that the most adaptive strategy for an infant is usually to process information at the highest possible level. The next principle describes what happens when that strategy fails.

1.5. If, for some reason, higher units are not available, lower-level units are utilized

A higher-level unit may be unavailable for a number of reasons, but it is often unavailable when the system gets overloaded. Circumstances for overloading the system can vary, but may include complicating the input through the addition of irrelevant material or noise, or converting a simple object or event into a category of objects or events. A corollary of this principle is that if for some reason the system does fall back to a lower level, it will then attempt to learn or accommodate so as to move back up to the next higher level.

1.6. This learning system applies throughout development and across domains

A strength of any theory is that it can account for a variety of findings in a variety of domains. This final principle highlights the robust, domain-general nature of the proposed learning system. Although in this article we are focusing on the applicability of these principles for infants' learning a variety of aspects of cognitive development, we believe this learning mechanism may be quite a bit more general than that. It may, in fact, represent how we as humans become proficient or experts in a wide range of tasks throughout the life span. It just so happens that one of first tasks encountered by young infants is trying to make sense of the immediate physical and social world around them. These principles may help them succeed at that task.

2. Empirical evidence

Evidence for the type of learning mechanism we have described continues to grow. This is true despite the fact that supporting research has not always been conducted from a constructivist perspective. Much of this evidence has been discussed extensively elsewhere (Cohen, 1998; Cohen & Cashon, 2001; Cohen & Cashon, in press). In the present article, we will review what we consider to be the some of best illustrations of the principles described. As we shall show by the variety of topics investigated in these studies as well as the diversity in ages of infants studied, the evidence presented also illustrates the domain-generality of the proposed learning system.

2.1. Early form perception

We have proposed that infants are born with a learning system capable of first processing lower-level information and then integrating lower-level units of information into higher-level information. A prime example was our research on infants' perception of angles (Cohen & Younger, 1984). In a habituation study, infants at 6- and 14-weeks of age were habituated to a picture of a 45° angle. The question being investigated was whether infants would process the two lines that comprised the angle independently of one another or whether they would process the relationship between those lines, in effect, processing the angle. After habituation, infants were tested on a variety of angles that varied in the orientation of the individual lines and/or the relationship between the lines, that is, the size of the angle. Differences were found in the type of information processed by the younger and older infants. Consistent with our principles, the older infants processed the relationship between the lines, whereas the younger infants processed the orientation of each line independently and therefore not the angle as a whole. These findings produced a developmental trend from processing lower-level units of a stimulus to processing the relationship between those units. The results support both IPPs #1 and #2. Infants' processing of lines and angles is clearly hierarchical with older infants processing the relationship between lower-level units.

Additional evidence on infant angle perception comes from Slater, Mattock, Brown, and Bremner (1991). They conducted a similar study with newborns and found that newborns, like our 6-week-olds, responded to line orientation. Slater et al. then made the task more difficult by rotating the angle during habituation. In essence they turned a simple task with one stimulus into a category study. Newborns no longer responded to line orientation. Now they appeared to fall to an even lower level of processing and just responded in terms of the low frequency "blob" at the apex of each angle (Cohen & Cashon, in press). Thus, the angle research also provides support for IPP #5, the fallback to a lower level of processing when the system gets overloaded.

2.2. *Object unity*

As suggested by the principles, this process of hierarchical integration is not limited to extremely simple stimuli such as angles. Once infants are capable of integrating simple lines into an angle, they should be equipped to use that angle as a part of a more complex form or object (IPP #3). Research on object unity provides additional support for this notion. The classic study by Kellman and Spelke (1983) showed that 4-month-old infants were capable of perceiving two ends of a moving rod as a single unified entity, even though the rod's middle section was hidden by the presence of a box (occluder). After being habituated to this partially hidden rod translating left and right behind the occluder, 4-month-old infants were found to look longer at a display with two broken rod pieces than one complete rod. These results were taken to mean that the infants had perceived (or possibly inferred) that there was one continuous rod moving during habituation and they looked longer at the two-rod test display because it was novel.

Kellman and Spelke's study spawned a series of experiments investigating the *development* of infants' ability to perceive object unity. These studies, which raise questions regarding a nativist view, are consistent with our idea that development involves the integration of lower-level information. One set of findings with newborns (Slater et al., 1990) and another set with 2-month-olds (Johnson & Náñez, 1995) illustrate the progression from piecemeal to integrative processing. Unlike the 4-month-olds in Kellman and Spelke, Slater et al. (1990) found that newborns did not perceive a unified rod. Instead they seem to have perceived the top and bottom portions of the rod as separate. Johnson and Náñez (1995), however, found that 2-month-olds were at an intermediate stage. They reported that their infants could perceive the rod as one object but only if the occluder was made more narrow than in the original Kellman and Spelke experiment.

Taken together, these findings illustrate the integrative aspect of development with respect to the processing of objects. The findings with 2-month-olds also serve to illustrate the idea that infants will process the highest level of information available to them (IPP #4), but will utilize the next lower level if necessary (IPP

#5). It appears that the original width of the occluder made the task of perceiving object unity too difficult, possibly because of the added difficulty of bridging a wide gap in the visible portions of the rod. But when the occluder was narrowed, 2-month-old infants were now able to perceive object unity.

Recently, Eizenman and Bertenthal (1998) attempted to replicate Kelman and Spelke but with the occluded rod rotating like a propeller about its center point rather than translating behind the occluder. In this case, 4-month-olds showed no signs of perceiving the rod as a unified object. In fact, there was a trend for them to look longer at the solid rod. Eizenman and Bertenthal (1998) found that it was not until 6 months of age that infants showed a significant preference for the broken rod during the test trials. From our point of view, a rotating rod would be much more category-like than a translating rod. Infants would be seeing the rod in many different angles and distances relative to the occluder. Under these circumstances, we might well expect an information overload with 4-month-old infants regressing back to an earlier level of processing (IPP #5). In fact, taken as a whole, this collection of experiments on object unity becomes a prototypic example of our IPPs.

2.3. Complex patterns or objects

Evidence that the complexity of information infants can process increases hierarchically over age also comes from a series of studies by Younger and Cohen (1983, 1986) on infants' perception of imaginary line-drawn animals. Compared to the angles, or even the rod and box displays discussed earlier, the stimuli used by Younger and Cohen were quite complex. The animals to be processed had different kinds of tails, legs, feet, ears, and body shape. Thus, is it appropriate that the ages studied would include not only 4 months, but also 7 and 10 months. The question Younger and Cohen asked was whether infants would process these animals as a set of independent features or whether they would process the correlations among the features.

The authors found a developmental progression consistent with our IPPs. They found that 4-month-old infants were insensitive to correlations among the features of the animals. They processed the animals solely as a group of independent features (IPP #1). Seven-month-old infants, on the other hand, were quite different. They had no trouble processing these rather complex animals at a higher level, in terms of the correlations among the features (IPP #2).

But the story gets even more interesting. Seven-month-old infants were able to process the correlation among features, but only when all the features were correlated. In one of the experiments reported by Younger and Cohen (1986), the task was made more difficult. Some of the animals' features were correlated, but other features varied independently. In essence they turned an animal identification task into an animal categorization task. Now the 7-month-olds had a great deal of difficulty even habituating to the stimuli. When forced to habituate, the 7-month-olds dropped down (IPP #5) to the level of the 4-month-olds and processed only independent features. Ten-month-olds, on the other hand, had no difficulty responding in terms of the correlations even in this category task.

Taken as a whole these findings on angle perception, object unity, and complex object perception, support the IPPs presented earlier. They provide evidence that infants are endowed with a hierarchical information-processing system; that lower units are integrated into higher units; and that those higher units can serve as elements of yet higher units. They also provide evidence that the system can become overloaded, and when it does, the infants fall back to a lower level of processing.

2.4. Causal events

So far, all of the evidence we have considered has been restricted to infants' processing of either a simple moving object, such as a translating rod, or a two dimensional pattern. But what about a more complex situation in which individual moving objects become elements in yet a higher order event? According to IPP #6, our proposed learning mechanism should be domain-general. It should also apply to dynamic interactive events as well as to simple events or static patterns, albeit with infants at somewhat older ages. Research on infants' understanding of causality provides an initial test of this domain-generality claim. Infant causality research also illustrates further how lower-level units can be used to form higher-level units (IPPs #2 and #3), and how an overload in the system can produce a drop to a lower level (IPP #5).

Most of the research on infants' understanding of causal events has examined infants' reactions to a simple launching event in which one object moves part way across a screen, hits a second object and stops, "causing" the second object to move the rest of the way across the screen (see Michotte, 1963 for research with adults). Leslie (Leslie, 1984; Leslie and Keeble, 1987) presented this simple causal event along with other non-causal events to 6.5-month-old infants. The non-causal events differed from the causal event in that they included either a spatial gap, a temporal delay, or both a delay and gap between the two objects. The infants seemed to process these events on the basis of their causality; that is, they responded similarly to the non-causal events, but differently to the causal event. Based upon these results, Leslie (1988, 1995) argued that infant causal perception must be hard-wired, modular, and subject to little or no developmental change.

Other research, though, tends to refute these claims. This research has shown that infant causal perception develops gradually, is hierarchical, and follows the same IPPs as other aspects of infant perception and cognition. Cohen and Amsel (1998), for example, examined infants at 4, 5.5, and 6.25 months of age. They found a three-step hierarchy. Four-month-old infants responded to the continuous versus non-continuous movement of the objects; 5.5-month-old infants responded primarily to the spatial and temporal relationships between these moving objects; and 6.25-month-old-infants responded to the causal versus non-causal nature of the overall event. Furthermore, Oakes and Cohen (1990) found that when realistic objects were used rather than the simple squares used by Leslie, 6.5-month-old infants became overloaded and no longer could respond on the basis of causality. Even more telling, Cohen and Oakes (1993) showed that 10-month-old infants, who could respond to causality even with realistic objects, fell back to the lower spatial–temporal relationship level when confronted with a category task in which the objects changed from trial to trial (IPP #5). Thus, the development of infants' understanding of causality provides yet another demonstration first, that the system is hierarchical and second, that it will fall back to a lower level when it becomes overloaded.

3. The Constructivist Learning Architecture (CLA)

The IPPs and supporting studies constitute a particular approach to infant cognition, one that focuses on hierarchical changes in development. But the material presented so far does not go into detail about the nature of the learning process that produces those hierarchical changes. In this section we would like to move to the next step and attempt to incorporate these principles in a specific computational model. A computational model can verify the principles at work in a particular domain, help flesh out the details of the learning process, and eventually produce predictions about infant cognitive development that can be tested empirically.

The CLA is a connectionist model of infant cognitive development that was built using the IPPs as a general design specification. Unlike many other computational models, CLA is not intended to be an "existence proof" showing that something *can* be learned. Rather, CLA is an attempt to model the constructive process of cognitive development as observed in infants. Although there are computational models that capture some of the observed principles to varying degrees (see, for example, Shultz & Cohen, 2002), there is no one model that addresses the entire complement. CLA addresses all six IPPs, and does so in a way that is simple and neurally plausible. Given the general applicability of the IPPs across various domains, a computational model based on these principles would have an equally broad scope.

CLA uses Kohonen's (1997) Self-Organizing Maps (SOM), an unsupervised learning system based on neuroscientific principles. CLA uses a SOM for each layer of unit, and connects multiple layers hierarchically using a Hebbian (1949) learning algorithm. The lowest level organizes simple perceptual stimuli into a set of prototypes positioned within a topographical map. The next level organizes the information represented by the lower-level map into yet another map of prototypes. Levels may continue to be added as needed. Thus, CLA is a neurally based, bottom-up, simple-to-compound system that is consistent with the evidence for constructivist learning presented in this paper and elsewhere. (See Chaput & Cohen, 2001 for an earlier version of this model.)

We do not claim, however, that all things learned by infants can be addressed with this model, nor are we suggesting that CLA is a complete neurological model of learning in the infant brain. There are many things unknown about cognitive development, and many things that are known and not addressed by CLA. We do not intend CLA to be a complete solution, but focused on one particular learning technique that we believe is prevalent in infant development.

3.1. The development of causal understanding

To test CLA, we applied the model to a domain of cognitive development that had a body of experimental data we could use for comparison and was sufficiently complex to truly test the limits of the model. The domain we chose was infants' understanding of causality, that is, infants' ability to respond to events as causal versus non-causal. Understanding causality is a serious challenge for CLA or any learning system. For one, causal events are not static images or feature sets, but dynamic events that occur over time. More importantly though, causal understanding is a question that has perplexed psychologists and philosophers for decades, a puzzle that, in its modern form, dates back to Hume (1777/1993) and Kant (1794/1982). Applying CLA to the understanding of causality would not only be a rigorous test of the model, but could contribute something valuable to a long-standing psychological and philosophical debate.

As we mentioned earlier in this article, Michotte (1963) explored causal perception in adult humans using simplified billiard-ball collisions, called launching events (see Fig. 1 for a schematic of these events). Michotte found that adults presented with a simple direct launching event would describe the event as causal. He could then alter the likelihood that subjects would identify the event as causal by altering two components of the event. He could either introduce a delay at the moment of the collision, or a gap between the two balls at the point of collision, or both. He found that manipulating the temporal component of the event (i.e., increasing the delay) or the spatial component of the event (i.e., increasing the gap) reduced the likelihood that the subjects would identify the event as causal.

Leslie (1984) used this same approach to identify causal perception in infants. He used a habituation paradigm to determine whether infants were attending to the spatial and temporal components of the events, or the causality of the events. Leslie placed the four events on a theoretical 2 × 2 grid (see Fig. 2) and tested infants

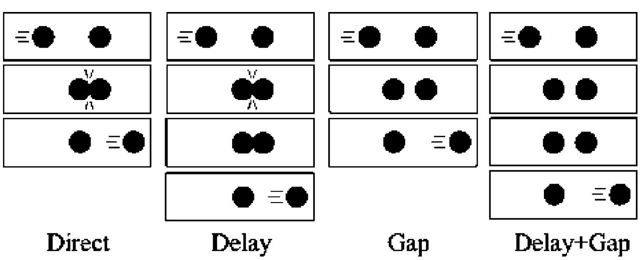

Fig. 1. The four launching events used as stimuli for our model.

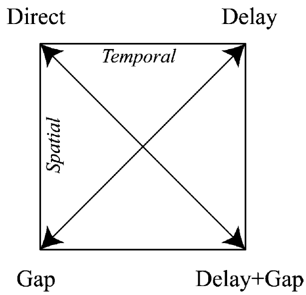

Fig. 2. The four launching events placed on a 2 × 2 grid. The horizontal axis represents the temporal feature of the event, with no Delay on the left, and a Delay on the right. The vertical axis represents the spatial feature of the event, with no Gap at the top, and a Gap at the bottom.

across the diagonals of the grid. In other words, some infants were habituated to a Direct event and tested on a Delay + Gap event (or vice versa). Other infants were habituated to a Delay event and tested on a Gap event (or vice versa). Leslie reasoned that, if infants had a component view of the events, if they were just responding perceptually in terms of the spatial and temporal properties of the events, then dishabituation along one diagonal should be similar to dishabituation along another diagonal, because both pairs of events involve a change in both components. However, if infants had a causal view of the events, then dishabituation along the Direct-to-Delay + Gap diagonal should be greater than along the Delay-to-Gap diagonal, because only the Direct event is causal. Leslie found that 6.5-month-old infants did, in fact, respond to the events in accordance with the causal view.

Further, Cohen and Amsel (1998) used a similar paradigm and found interesting developmental changes in infants' responses to causality. Rather than test across the diagonals, Cohen and Amsel compared dishabituation between Gap and Direct events with dishabituation between Gap and Delay events. If infants were responding in terms of causality they should have dishabituated more when going from the Gap to Direct event, as that difference includes a change in causality. Cohen and Amsel found that 6.25-month-olds did respond to the events causally, thus replicating Leslie. However, they also found that 5.5- and 4-month-olds responded to the events in a way more consistent with a component view. Four-month-old infants were sensitive mainly to the duration of movement in the event, whereas 5.5-month-old infants were also sensitive to the spatial and temporal components. Neither age responded in terms of the causality. These results provide an indication of a developmental shift in causal understanding that progresses from a component view to a higher-level causal view.

3.2. Design

The challenge then is to model this apparent part-to-whole development of causal understanding within CLA. Fig. 3 provides a schematic overview of the

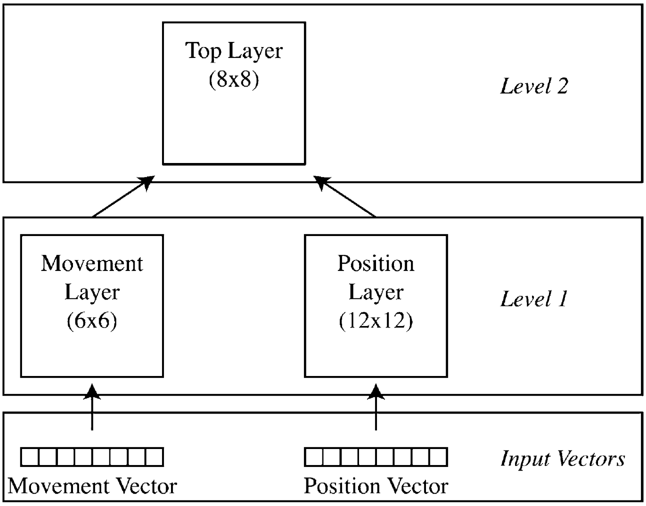

Fig. 3. A schematic of the layer setup used in our CLA model of an infant's development of causal understanding. The bottom two vectors are the input, the middle two layers are the (spatial and temporal) component layers, which are both projected onto the Top Layer.

different layers of the model. We started with a representation of the launching events that captured only the different spatial and temporal components. The events were represented using two input feature vectors to represent the two components of the launching events. The first input, called the Position Input, captured the spatial information by representing the position of the two balls at any point in time. The second, called the Speed Input, provided the temporal information of the events. It represented the speed of each ball at any point in time. These two parallel input vectors presented the stimuli to the model using features that we know even young infants can perceive (Cohen & Amsel, 1998).

Each of these input vectors projects to a corresponding layer, the Position Input to the Position2 Layer, and the Speed Input to the Speed2 Layer. These second level layers represent, as prototypes, the different positions and speeds separately over the course of an entire event. So, for example, different areas of the Position2 Layer would represent the balls being far apart, close together, or touching. Different areas of the Speed2 Layer would represent both balls stationary, or only the first ball moving, or only the second ball moving.

Finally, the Top Layer of the model receives inputs from both the Position2 and Speed2 Layers. Its task is to represent the event as a whole and hopefully, after training, be able to distinguish causal from non-causal events.

During training, each event was presented to the model at a frequency determined by our crude estimate of its relative frequency in nature. Based on these estimates, Direct events were presented 85% of the time and each of the other events (Delay, Gap, and Delay + Gap) were presented 5% of the time. We feel that the frequency of experiencing these events, while certainly not the only information used by infants, plays a crucial role in the acquisition of causal understanding. If the world consisted mostly of Delay events, for example, infants would develop a radically different view of causality.

3.3. Results

As each layer received more training, different prototypes in each layer came to represent the launching events at different points in time. If we combine all of the activations from all the time steps, we get a composite activation graph showing all the prototypes that were activated during the event. Fig. 4 shows the composite activation in the Speed2 Layer for all four events. Although there is a clear distinction in the activation patterns between events that have a delay and events with no delay, there is no distinction between Gap events and non-Gap events. In other words, certain prototypes have specialized to represent the presence or absence of a Delay, but not of a Gap. Conversely, Fig. 5 shows the exact opposite pattern for events in the Position2 Layer. There are specialized prototypes for the presence and absence of a gap, but not a delay. The lower-level layers are working as they were designed, reflecting the components of the event.

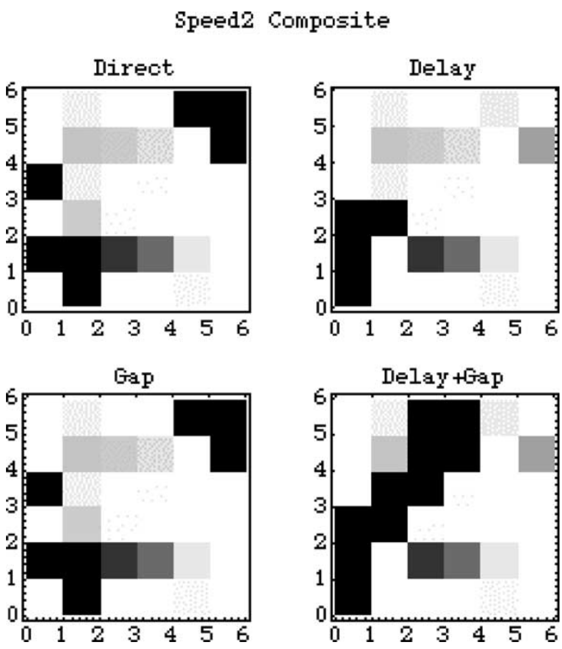

Fig. 4. The composite activation of the Speed2 Layer for all four launching events (after training). This layer is sensitive to the presence of a Delay, but not a Gap.

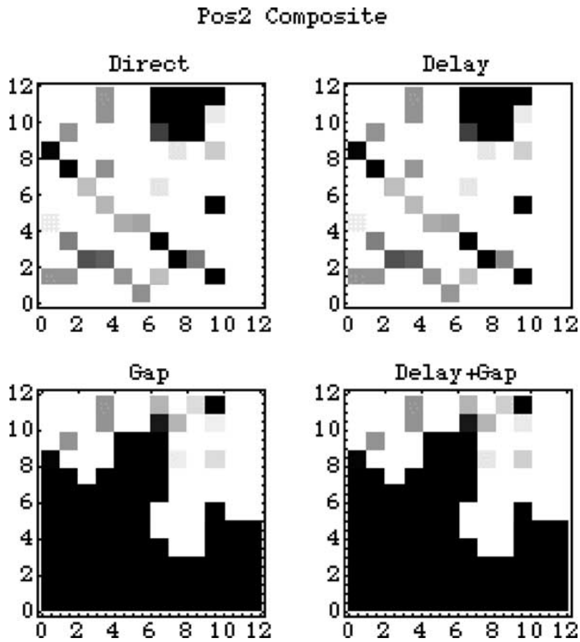

Fig. 5. The composite activation of the Position2 Layer for all four launching events (after training). This layer is sensitive to the presence of a Gap, but not a Delay.

But the composite activation graphs for the Top Layer (in Fig. 6) show no such symmetry. While the lower levels had prototypes that specialized for the components of the launching events, the Top Layer has prototypes that specialize for the events themselves. To identify which prototypes have specialized for which events, we can take the composite for a given event and subtract out the composites for the other three events. The results of this subtraction can be seen in Fig. 7. As the figure shows, there are specialized prototypes that activate exclusively during a Direct event, although there are no such specialized prototypes for Delay and Gap events. This shows that the Top Layer has created a unique representation for the causal event that does not exist for other non-causal events. (The specialized prototypes that appear for the Delay + Gap event also fit our predictions and will be discussed in more detail below.)

To compare our results with looking times found by Leslie (1984) and Cohen and Amsel (1998), we looked at the difference in activation between pairs of events. The results of these comparisons for the events used by Leslie can be seen in Fig. 8. We found that the difference in activation is significantly greater between the causal (Direct) and non-causal (Delay + Gap) events ($M = 9.66$, S.D. $= 1.16$) than between the two non-causal events (Delay and Gap) ($M = 5.38$, S.D. $= 1.00$), $t(6) = 5.58$, $P = .002$ (two-tailed), $d = 4.28$. This pattern of results is consistent with the difference in looking times found by Leslie, who considered his results

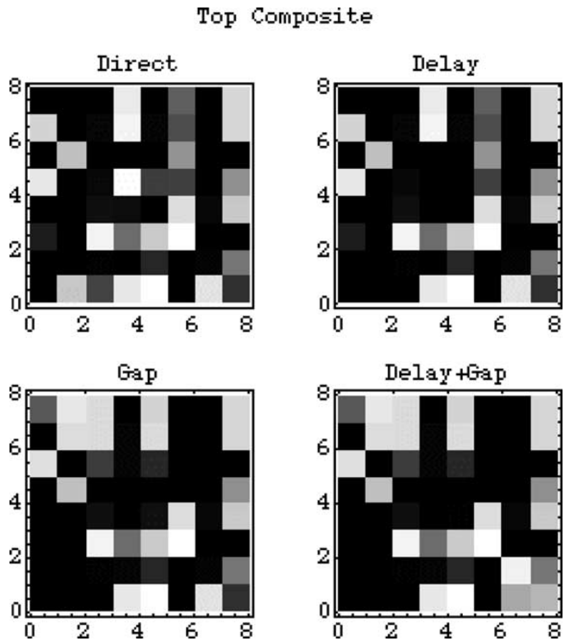

Fig. 6. The composite activation of the Top Layer for all four launching events (after training). This layer has difference representations for all four events.

to be evidence for infant causal perception. A similar pattern of results was found by Cohen and Amsel (1998), who also took their findings as evidence for infant causal perception. Again, our model produced activation levels consistent with the empirical data. As seen in Fig. 9, the activation difference is significantly greater between the causal (Direct) and non-causal (Gap) events ($M = 8.20$, S.D. $= 0.56$) than between the two non-causal events (Delay and Gap) ($M = 5.38$, S.D. $= 1.00$), $t(6) = 4.93$, $P = .005$ (two-tailed), $d = 2.82$.

The model's treatment of the Delay + Gap event was an unexpected surprise. Rather than just building specialized prototypes for the Direct event, the Top Layer also built specialized prototypes for the Delay + Gap event (see Fig. 7). In effect, the model did not just represent events in terms of "causal" and "not causal," but in *degree* of causality. A special representation was created for the Delay + Gap event because it is "less causal" then an event with just a Delay or Gap by itself. This suggests that the events actually lie along a continuum, ranging from causal to non-causal to very non-causal. Such a continuum was first proposed by Leslie and Keeble (1987). Fig. 10 is a reproduction of their "spatiotemporal continuity." It is wholly compatible with our results, placing the Direct and Delay + Gap events at the extremes, while placing the Delay and Gap events somewhere in the middle.

The model also provides an indication of stage-like development. In Fig. 11, we can see the composite activation for the Direct event in the three layers at different

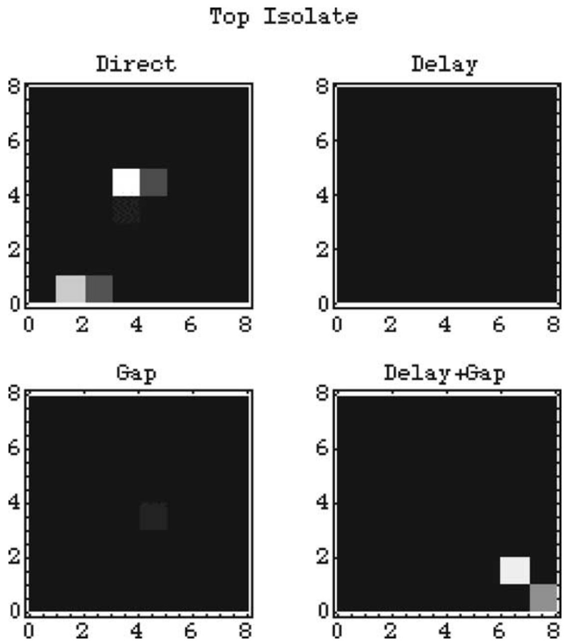

Fig. 7. The activity of the Top Layer for each of four launching events with the activation of the other three events subtracted out. There are specialized prototypes for the Direct and Delay + Gap events, but not for the Delay and Gap events.

points in training. Although the lower Position2 and Speed2 Layers begin to form immediately, and start to settle at about 4500 epochs, it is not until this point that the Top Layer begins to form. This is a consequence of CLA's hierarchical design. Higher layers are organizing patterns of activity in lower layers, but they

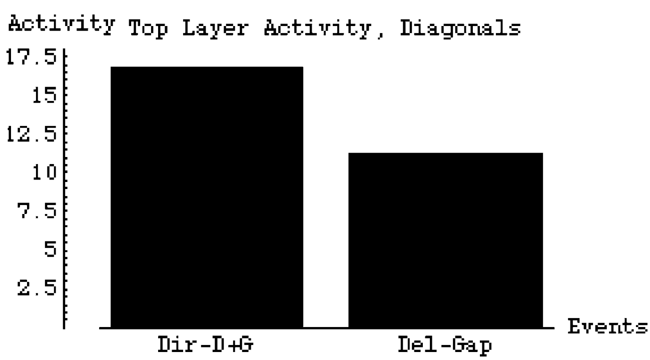

Fig. 8. The difference in activation between the Direct and Delay + Gap events compared to the difference in activation between the Delay and Gap events. As Leslie (1984) found in 6.5-month infants, the model shows a greater difference with a change in causality than without one.

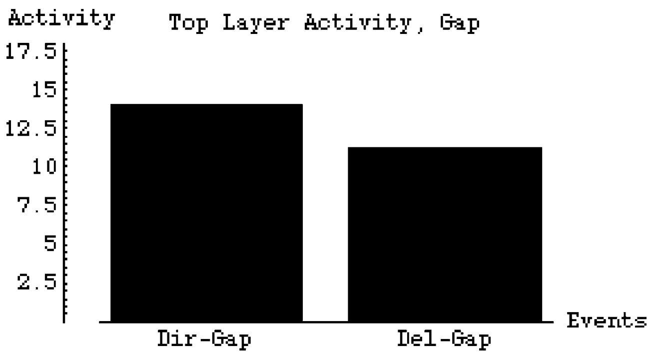

Fig. 9. The difference in activation between the Direct and Gap events compared to the difference in activation between the Delay and Gap events. Again, like Cohen and Amsel (1998) found in infants, the causal difference is significantly greater than the non-causal difference.

Fig. 10. Leslie and Keeble's (1987) "spatiotemporal continuity" with our own event names. Similar to their hypothesis, CLA representation of different events is compatible with this continuum, with the Direct event at one end, the Delay + Gap event at the other, and the Delay and Gap events somewhere in the middle.

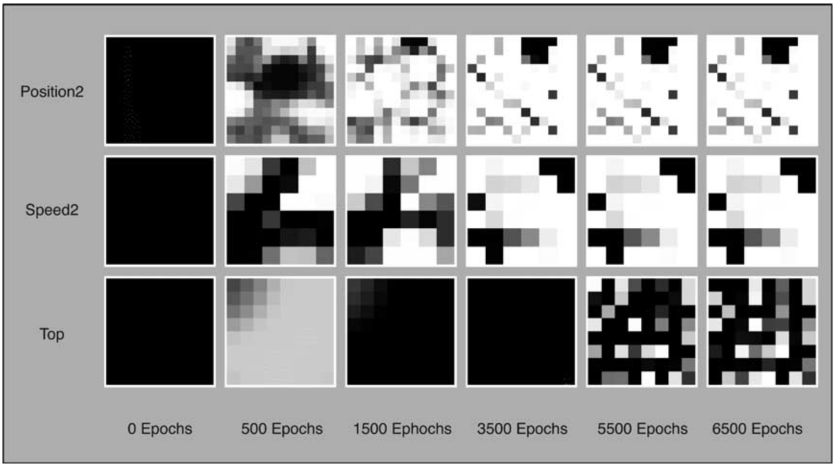

Fig. 11. Activation composites for the three layers during the Direct event at different points in learning. The lower layers (Position2 and Speed2) start forming a representation immediately, and settle at about 4500 epochs. It is not until this point that the Top Layer starts to form a representation.

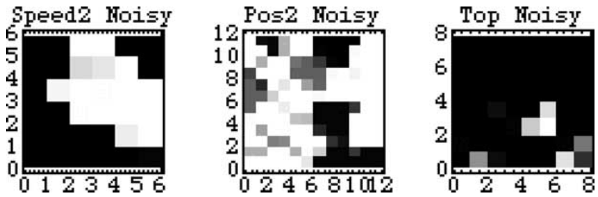

Fig. 12. The activation composite of the "noisy" launching event for the Speed2, Position2 and Top Layers. The Top Layer shows very little activation for the noisy event. But there is substantial (albeit diminished) activation at the lower layers, indicating that the overloaded system is falling back to a component representation of the event.

cannot organize anything until the lower layers are stable. Even though CLA uses a continuous learning system, we see what appear to be stages of development, which is consistent with Cohen and Amsel's (1998) results that infants at different ages process the same events differently.

Finally, there is also evidence of "fallback" in the model. Recall that IPP #5 says that, should higher units be unavailable, lower units should still be available to process the stimulus. Higher units might become unavailable if the system is overloaded by interference or categorization problems (among other things). To test this, we created a Noisy launching event, in which the speed varied throughout the event, and the position of the moving balls would pause and jump. (Imagine that the billiard balls are oval.) This is a novel event for the trained model, and we expected to see some activation at the lower levels, but limited activation at the higher levels. That is just what we found, as seen in Fig. 12. Whereas the lower layers are activated by the event (albeit to a lesser degree), the Top Layer's activation is greatly diminished. This result is consistent with the study we reported earlier by Cohen and Oakes (1993). They also found that infants dropped from a causal to a non-causal level of responding when the task was too complex.

4. General discussion

We have presented a set of six IPPs, which help to describe and interpret infants' cognitive development in several domains. These principles suggest that much of infant cognitive development is both constructive and hierarchical. A building block approach is proposed in which simple informational units processed at one level become integrated into more complex units at the next level, which in turn become integrated into yet more complex units at a third level, and so on. Furthermore, once infants have access to multiple levels they will tend to process information at the highest possible level unless the system is somehow overloaded. In that case infants will drop down to a lower level of processing and attempt to rebuild the system.

We presented empirical evidence from four different domains of infant perception and cognition to verify these principles. The domains ranged from simple form perception, to object unity, to complex pattern perception, to an understanding of causal events. In each case we provided evidence that developmental changes in these domains followed our proposed constructive, hierarchical pathway. In each case we also provided evidence that an overload in the system would produce a drop to a lower level of processing.

We then used these principles to build the CLA, a computation model of cognitive development that is neurally plausible and psychologically grounded. We applied the model to the development of causal understanding in infants, a particularly challenging domain. The model conformed well to the IPPs. It demonstrated that learning to understand causal events follows a hierarchical progression of levels and that presenting a heretofore unknown event produces a fall back to a lower level. Even more important, the model reproduced the results of Leslie (1984), Cohen and Amsel (1998), Leslie and Keeble (1987), and Cohen and Oakes (1993) including the unanticipated finding that the top level of the model produced a continuum of causality.

We believe we have demonstrated that exploring infant cognitive development using the IPPs and the CLA is a promising approach, one that has much potential. Clearly, there is more work to be done. We plan to continue to refine our model of causal acquisition, so that we may better understand the specific mechanisms at work. This will involve not only modeling what is known, but using the model to generate predictions as well. Using a simulation to guide empirical research, rather than just explain it *post hoc*, is a powerful use of a quantitative model. We also plan to explore many other domains of infant cognition, both by understanding them in terms of the IPPs, and by modeling them using the CLA.

By focusing on the development of knowledge in infants, rather than just the presence of knowledge, we are beginning to put together a more comprehensive and detailed picture of infant cognition. And from this picture, we are building a more accurate model of a central mechanism in cognitive development. Our belief is that understanding what an infant knows starts with an understanding of how an infant learns.

Acknowledgments

This research was supported in part by NIH grant HD-23397 awarded to Leslie B. Cohen.

References

Cashon, C. H., & Cohen, L. B. (2000). Eight-month-old infants' perception of possible and impossible events. *Infancy*, *1*, 429–446.

Chaput, H. H., & Cohen, L. B. (2001). A model of infant causal perception and its development. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 182–187). Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, L. B. (1995, March). *How solid is infants' understanding of solidity?* Symposium paper presented at the Society for Research in Child Development Meeting, Indianapolis, IN.

Cohen, L. B. (1998). An information-processing approach to infant perception and cognition. In F. Simion & G. Butterworth (Eds.), *The development of sensory, motor, and cognitive capacities in early infancy* (pp. 277–300). East Sussex: Psychology Press.

Cohen, L. B., & Amsel, G. (1998). Precursors to infants' perception of causality. *Infant Behavior and Development*, *21*, 713–731.

Cohen, L. B., & Cashon, C. H. (2001). Infant object segregation implies information integration. *Journal of Experimental Child Psychology*, *78*, 75–83.

Cohen, L. B., & Cashon, C. H. (in press). Infant perception and cognition. In R. Lerner, A. Easterbrooks, & J. Mistry (Eds.), *Comprehensive handbook of psychology. Volume 6, Developmental psychology. II*. Infancy. New York: Wiley.

Cohen, L. B., Gilbert, K. M., & Brown, P. S. (1996, April). *Infant's understanding of solidity: Replicating a failure to replicate*. Poster presented at International Conference on Infant Studies, Providence.

Cohen, L. B., & Marks, K. S. (2002). How infants process addition and subtraction events. *Developmental Science*, *5*, 186–201.

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, *29*, 421–433.

Cohen, L. B., & Younger, B. A. (1984). Infant perception of angular relations. *Infant Behavior and Development*, *7*, 37–47.

Eizenman, D. R., & Bertenthal, B. I. (1998). Infants' perception of object unity in translating and rotating displays. *Developmental Psychology*, *34*, 426–434.

Elman, J., Bates, E., Karmiloff-Smith, A., Johnson, M., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, *21*, 167–179.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Hume, D. (1777/1993). *An enquiry concerning human understanding*. Indianapolis, IN: Hackett.

Johnson, S. P., & Náñez, J. E., Sr. (1995). Young infants' perception of object unity in two-dimensional displays. *Infant Behavior and Development*, *18*, 133–143.

Kant, I. (1794/1982). *Critique of pure reason*. Translated by Wolfgang Schwarz. Aalen: Scientia.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483–524.

Kohonen, T. (1997). *Self-organizing maps*. Berlin: Springer.

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception*, *13*, 287–305.

Leslie, A. M. (1988). The necessity of illusion: Perception and thought in infancy. In L. Weiskrantz (Ed.), *Thought without language* (pp. 185–210). Oxford: Oxford University Press.

Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition*. (pp. 121–141). Oxford: Clarendon Press.

Leslie, A. M., & Keeble, S. (1987). Do six-month-olds perceive causality? *Cognition*, *25*, 265–288.

Michotte, A. (1963). *The perception of causality*. New York: Basic Books.

Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development*, *5*, 193–207.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Shultz, T. R., & Cohen, L. B. (2002, April). *Modeling age differences in category learning*. Symposium paper presented at International Conference on Infant Studies, Toronto

Slater, A. M., Mattock, A., Brown, E., & Bremner, J. G. (1991). Form perception at birth: Cohen and Younger revisited. *Journal of Experimental Child Psychology*, *51*, 395–405.

Slater, A., Morison, V., Somers, M., Mattock, A., Brown, E., & Taylor, D. (1990). Newborn and older infants' perception of partly occluded objects. *Infant Behavior and Development*, *13*, 33–49.
Spelke, E. S. (1998). Nativism, empiricism, and the origins of knowledge. *Infant Behavior and Development*, *21*, 181–200.
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*, 158–177.
Younger, B. A., & Cohen, L. B. (1983). Infant perception of correlations among attributes. *Child Development*, *54*, 858–867.
Younger, B. A., & Cohen, L. B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development*, *57*, 803–815.