Title
How infants learn about people and object causal action: An associative account

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Author
Benton, Deon T.

Publication Date
2023

Peer reviewed

How infants learn about people and object causal action: An associative account

Deon T. Benton (deon.benton@vanderbilt.edu)
Department of Psychology and Human Development, Vanderbilt University

Abstract

Causal perception is a cornerstone of early cognitive development. A large database of research attests to the fact that infant causal perception emerges between 6 and 10 months of age. However, it remains unknown how infants learn about the causal properties of more realistic categories such as people and objects. For example, how do infants learn that people can cause other people to act and behave at a distance, whereas inanimate objects require contact to move and act? One answer to this question is that this knowledge is present from birth or shortly thereafter and is underpinned by core knowledge systems. An alternative perspective maintains that infants acquire this knowledge via domain-general associative learning. The goal of the present paper is to demonstrate that this alternative perspective—implemented in a connectionist computational model—is sufficient to explain infants' developing knowledge about people and object causal action.

Keywords: causal perception; causality; computational modeling; artificial neural networks; developmental learning mechanisms

Introduction

Causal perception—or the capacity to "see" and appreciate causal relations—is a key cognitive ability that enables human beings to understand how the world works around them. This is a fundamental ability from a developmental perspective because it supports infants' emerging knowledge concerning the causal relations among and between objects and entities in the world; that is, this capacity allows infants to determine which objects in the world are agents—which cause action and produce effects—and which objects are recipients—which are acted on.

There is now considerable evidence that the ability to perceive causal relations emerges between 6 and 10 months of age[1-2]. In one of the first studies on causal perception, Leslie and Keeble (1987) habituated infants to one of two computer-animated, artificial launching sequences: a direct launching sequence or a delayed launching sequence. In the direct launching sequence, a first object caused a second object immediately to move through direct, physical contact. The delayed launching sequence was similar to the direct launching sequence except that the second object began to move only after a brief delay following contact from the first object. Infants were then tested with the reversal of their respective habituation sequences. The authors reasoned that if 6½-month-olds can perceive cause-and-effect relations, then infants habituated to the direct launching sequence should look longer at the reversal of that event than should infants habituated to and then tested on the reversal of a delayed launching sequence. This is because only the direct-launching sequence entailed a reversal in the causal direction of the event such that the former agent became the recipient, and the former recipient became the agent. This is what Leslie and Keeble (1987) found[4].

These findings were subsequently extended to show that by 9 months of age infants treat launching events differently than triggering events[6], and by 15 months of age infants can perceive causality in causal-chain sequences that consist of three rather than two objects[7]. This research indicates that causal perception undergoes a clear developmental progression during the first two years of life.

Mechanisms supporting infant causal perception

Despite this robust body of research and agreement among researchers about the developmental timeline of infant causal perception, it remains unanswered when and importantly how infants learn about the causal properties of real-world categories such as people and objects. It also remains unknown how infants come to understand that people and objects possess distinct causal properties. For example, human beings can cause other human beings either to act at a distance or through direct, physical contact, whereas inanimate objects tend to require physical contact to act (e.g., to move from one place to another) One explanation for the dearth of research on this issue is that most, if not all, of the research on infant causal perception has used artificial stimuli that bear little resemblance to the kinds of objects and entities that infants and young children may encounter in the real world. This is problematic because a complete understanding of infant causal perception—including when and how this capacity emerges—ultimately will require knowing how infants come to perceive causality in casual events that use artificial as well as realistic stimuli.

One prominent theory that attempted to address this open question is the Core Knowledge perspective[8-9]. The crux of this account is that infants know from birth or shortly thereafter that people and inanimate objects behave differently, and that this knowledge is subserved by a small number of "core" systems. For example, the core system for agents enables infants innately to know that agents are goal-directed, self-propelled, and can cause action in other agents both at a distance as well as through direct, physical contact. In contrast, the core system for objects allows infants innately to know that objects do not move in the absence of physical contact from other objects or agents and are neither self-propelled nor goal directed.

Support for these two systems was garnered by Spelke, Phillips, and Woodward (1995). This study used the logic of the launching studies to investigate whether 7-month-old infants understand that people, but not inanimate objects, can cause action at a distance in other objects. Using a between-subjects design, infants were habituated to one of two live events that involved real, three-dimensional objects and

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

people. In one event, which corresponded to the People condition, a person entered the stage from the left and traveled a short distance until it disappeared behind a centrally located screen. A second, initially half-covered person then emerged from behind the screen and exited the stage to the right. The second event, which corresponded to the Inanimate Object condition, was identical to the first one except that it used two inanimate objects rather than two people (i.e., a red box with a jagged top edge and a blue cylinder). Infants then saw two test events without the screen three times in alternation. In the Contact test event, the first person or object moved toward and contacted the other person or object at the center of the stage. It should be noted that this event is similar to the direct launching event described earlier except that it used real people or objects. In the No Contact test event, the first person or object moved toward but ultimately stopped short of the other person or object. The results revealed that infants in the Inanimate Object condition looked longer at the No Collision test event than at the Collision test event. In contrast, infants in the People condition looked equally long at both test events. These findings were interpreted to mean that infants possess core knowledge about people and objects.

Although the Core Knowledge perspective can account for infants' looking patterns in Spelke et al. (1995), this theory has notable shortcomings. First, although the Core Knowledge account assumes that infants possess innate knowledge about people and object causal action, Spelke et al. only tested 7-month-olds. This means that infants' looking behavior could have been based on seven months of real-world, extensive causal experience with people and objects rather than on innate knowledge. A second limitation concerns aspects of the stimuli that were used. Prior to habituation all infants were familiarized with the people and object stimuli. Infants assigned to the Inanimate Object condition were shown two inanimate objects—one to the right of and the other to the left of a central screen. Infants assigned to the People condition were shown two people that stood on either side of the screen. Crucially, the people, but not the inanimate objects, danced and wiggled slightly throughout the familiarization phase. This means that the two conditions differed not only in terms of the stimuli that were used (i.e., people vs. inanimate objects) but also in terms of important animacy cues (i.e., only the people danced and wiggled). If infants learned that things that move in small but perceptible ways can cause action at a distance in other things—rather than that people, per se, cause action at a distance in other people—then infants would be expected to show heightened looking when two objects, but not when two people, acted at a distance. This means that it is impossible to know from this study alone whether infants' responses to the Contact and No Contact test events across experimental conditions was due to these animacy cues or to innate causal knowledge about people and objects.

Given these limitations, the goal of the simulations presented here is to test an alternative mechanistic account for how infants come to know that people and objects possess distinct causal properties. The crux of my argument is that infants' knowledge about people and object causal action derives from an associative-learning mechanism that links salient perceptual features (e.g., eyes or legs) with distinct kinds of causal action (e.g., contact causality or action-at-a-distance causality), perhaps in conjunction with a small number of inherent perceptual biases, such as biases for face-like stimuli over non-face-like stimuli and for movement over non-movement, among others. These biases serve to emphasize certain regions of the perceptual array and to deemphasize other regions. This mechanism begins when infants first notice—either incidentally or intentionally—that a link exists between certain salient perceptual features (e.g., eyes, legs, etc.) and different kinds of causal action (e.g., action-at-a-distance causality and contact causality). Infants may come to notice and then encode this link based on the fact that some perceptual features tend to co-occur with different kinds of causal action across time and space. For example, infants may learn that entities with legs can cause other entities with legs to act both at a distance and on contact. Infants may be attuned to these relations in the first place—that is, that legs "go with" two kinds of causal action—based on an innate or early emerging bias to attend to movement over non-movement. Infants may come subsequently to notice that legs tend to be included in these movement events. This act of noticing that contact causality and action-at-a-distance causality tend to include legs may serve to establish a nascent link between legs and these two kinds of causality, which becomes ever strengthened as the components of the relation (i.e., legs and the two kinds of causality) are experienced repeatedly together. The consequence of this developmental mechanism is that the presence of one of the features alone (e.g., legs) will, with time, trigger an expectation for the other feature (e.g., action-at-a-distance causality), even if the other feature is not physically present. This means that if one (e.g., legs), but not both, of the features is physically present, this may cause infants to show heightened or increased looking, as if to be looking for or expecting the second feature (i.e., legs). This may explain why infants looked longer at the No Collision test event relative to the Collision test event in the Inanimate Object condition but not in the People condition; the No Collision event (i.e., action-at-a-distance event) triggered an expectation for legs that ultimately was not met (because the objects lacked legs).

Present simulations

The goals of the present simulation studies were twofold. First, Simulation 1 modeled infants' looking behavior in Spelke et al. (1995). This means that models, like the infants, were assigned either to the People condition or to the Inanimate Object condition. Models were then habituated to and tested on events in which people or objects caused action in other people or objects either at a distance or through direct, physical contact. To examine what predictions, if any, the network (and by extension, the theory) makes about younger infants' knowledge about people and object causal action, “younger” networks were tested in Simulation 2. Simulation 3 extended the first two simulations by exploring how the model treated “modified people” (i.e., people with object parts) and “modified objects” (i.e., objects with people parts). The model’s behavior in this new situation will serve as a key behavioral prediction that can be tested in future research. The broader aim of the present simulation studies is to provide a proof-of-concept that an associative learning mechanism—instantiated in an artificial neural network—is sufficient to explain infants’ developing knowledge about people and object causal action. This would not disconfirm the Core Knowledge perspective but rather would suggest that innate knowledge may not be necessary to explain infants’ causal knowledge about people and objects.

Simulation 1

The aim of Simulation 1 was to model Spelke et al. (1995) and to demonstrate that an associative-learning mechanism—implemented as an artificial neural network—is sufficient to account for the results.

Method

Network architecture

I used a three-layer, feedforward, simple-recurrent network (SRN)[11] with an autoencoder component[12]. SRNs are applied to sequential problems in which a network’s current activity is conditioned on its previous activity. Autoencoders learn to reproduce on the output layer the pattern of activation that is presented on the input layer. The model was trained using the backpropagation algorithm[13]. The learning rate, momentum, and weight decay were set to, respectively, 0.05, 0.9, and 0.001. Weights were initialized to small random values (sampled uniformly between ±0.1). Finally, the activation of the output units was set according to a sigmoid activation function, whereas the activation of the hidden units was set according to a RELU activation function to prevent the gradients from vanishing as error is backpropagated across time.

The model consisted of three layers (Figure 1). The input and output layer consisted of 5 “banks” of units and the hidden layer consisted of two groups of hidden units. Two of these banks of units—which consisted of 40 units—were used to represent the people and objects. People and objects were represented as orthogonal patterns of activity to ensure that the network’s responses at test were based on learned associations between particular features and different types of causal action rather than on the particular features of a given person or object. Patterns of activation were presented in both banks of units to simulate the fact that two people (or objects) were present.

In addition to these banks of units, two other banks of units represented whether a given person or object possessed legs. These banks consisted of two units each. If the first unit in this bank was set to “on” (i.e., its value clamped to 1) and the second unit in this bank was set to “off” (i.e., its value clamped to 0), this indicated that the person or object to which this bank of units corresponded possessed legs. However, if the first unit in this bank was set to “off” and the second unit set to “on,” this indicated that the person or object to which this bank of units corresponded did not possess legs. In the simulations reported here, only the people possessed legs. More realistic simulations might explore whether the present results change when some proportion of the objects possesses legs and some proportion of the people do not possess legs. These 4 banks of units instantiated the autoencoder component of the network. In other words, for these banks of units the network’s task was to learn to recreate the pattern of activity presented in each of these groups in the corresponding output groups through an intermediate group of hidden units. The number of hidden units in the hidden layer was smaller than that in either the input or output layers. This “feature” of the model forces the network to develop a more compact representation of the input that is sufficiently reliable to reproduce the pattern of activation along the input layer at the output layer[12].

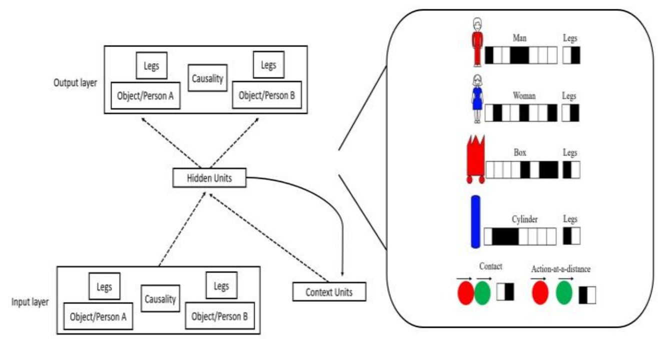

Figure 1. Schematic of the model used in these simulations.

The final bank of units represented the location of the object or person currently in motion. Note that although two of the input groups represented distinct objects or people, there is a single motion path. This means that on each timestep a single bit in the motion layer was active according to the current position of the first or second object. The networks’ task was to predict the correct active bit of the motion layer occurring at the next timestep. Activity presented before the midpoint of this path (i.e., before the 5th bit of the motion path) represented the motion of the object or person on the left and activity presented after this midpoint encoded the motion of the object or person on the right. Note that in the same way that nothing in the real world explicitly demarcates people from objects, the networks used here demarcated people from objects based on the fact that they engaged in different kinds of causal action and possessed distinct perceptual features (e.g., legs).

Finally, the hidden layer consisted of 12 units. Note that the hidden layer was connected to a corresponding group of “context” units. These context units encoded the pattern of activity that was presented along the hidden at the previous time step, which enabled the model to remember information from the past and use it to guide future learning[11].

Training

Pretraining. All networks in this simulation received 600 epochs of pretraining experience. The pretraining phase was meant to correspond to the "real world" experience with which infants presumably entered Spelke et al.'s (1995) study. There were three kinds of events that networks experienced during this phase. In one event (N = 16), a first person (i.e., an entity with legs) would cause a second person to move in the absence of physical contact. In other words, on the first three timesteps, the first three bits of the motion layer were active (i.e., positions 1-3). This corresponded to the motion of the person on the left. On the next three timesteps, the last three bits of the motion layer were active (i.e., positions 6-8). This corresponded to the motion of the second person. The fact that no motion (or activity) occurred at the 4^{th} and 5^{th} bits of the motion layer indicated that the first person caused the second person to move in the absence of physical contact. In another event (N = 16), one person caused another person to move through direct, physical contact. This event was similar to the first event except that the first person contacted the second person (i.e., the 4^{th} and 5^{th} bits of the motion layer were sequentially activated, along with the remaining bits). The final set of events (N = 16) were similar to the People Contact events except that objects were used (i.e., the stimuli did not have legs). The pretraining phase implemented the idea from the theory outlined above that people can engage in multiple forms of causality (i.e., contact as well as action-at-a-distance causality), whereas objects only engage in contact causality.

Habituation. Similar to Spelke et al. (1995) networks were randomly assigned to the People condition (N = 64 networks) or to the Object condition (N = 64 networks). The habituation events were similar to the pretraining events except that motion was absent during this phase and a new set of unrelated people or objects were used. I chose not to model motion during this phase to simulate the fact that infants could not determine whether a contact or no contact event was being shown during habituation in Spelke et al. (1995) due to the fact that a central screen obscured which of the two causal actions was being shown during habituation. Given that motion was absent during this phase, networks simply had to reproduce the pattern of activity in each of the input groups in the corresponding output groups. Networks assigned to the People condition saw habituation events that used people (i.e., stimuli with legs), whereas networks assigned to the Object condition saw habituation events that used objects (i.e., stimuli without legs). Networks were habituated to 4 distinct events involving 4 different people or objects. I chose to habituate networks to multiple events, rather than to a single event as was done in Spelke et al. (1995), to determine whether a given network's response to the habituation or test events was due to the particular habituation or test events.

Testing

Following habituation, networks were tested on 4 Contact and 4 No Contact test events in alternation (as was done in Spelke et al. [1995]), and the networks' response to the test events was assessed. We used sum-squared error (SSE) as a measure of "looking time"[14]. Larger errors indicate a larger discrepancy between what the network observes (the pattern of activity across the output layer) and what it expects (the target information across the output layer). The contact events were identical to the pretraining contact events and the no-contact events were identical to the pretraining action-at-a-distance events. A network's response to each event was averaged over each of the corresponding test events.

Results

Figure 2 shows the networks' mean "looking times" to the Contact and No Contact test events across the People and Inanimate Object conditions. The data revealed that the difference in "looking time" (i.e., difference in SSE) to the No Contact and Contact test events was greater for models assigned to the Inanimate Object condition (N = 64) ($M_{\text{No-Collision}}$ = 65.58, $SD_{\text{No-Collision}}$ = 10.61; $M_{\text{Collision}}$ = 51.37, $SD_{\text{No-Collision}}$ = 4.88) than models assigned to the People condition (N = 64) ($M_{\text{No-Collision}}$ = 54.26, $SD_{\text{No-Collision}}$ = 4.28; $M_{\text{Collision}}$ = 55.83, $SD_{\text{No-Collision}}$ = 4.13), t(62) = 77.56, p < .0001. Crucially, this result replicates exactly infants' looking behaviors in Spelke et al. (1995).

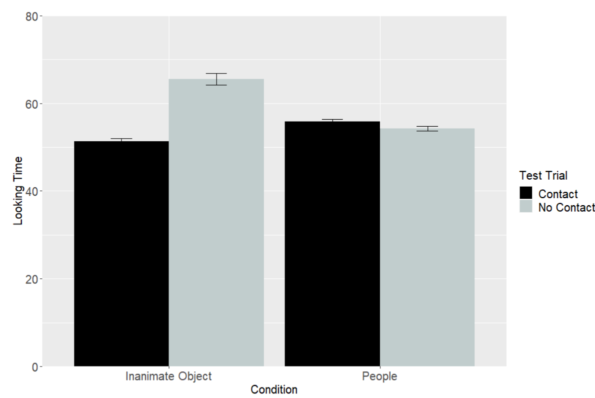

Figure 2. Networks' mean "looking time" (i.e., SSE) to the Contact and No Contact test events across conditions.

Discussion

Simulation 1 replicated infants' looking responses in Spelke et al. (1995) such that the difference in "looking time" (i.e., sum-squared error) to the Contact and No Contact test events was greater for networks assigned to the Inanimate Object condition than for networks assigned to the People condition. This result suggests that Core Knowledge systems may not be necessary to explain infants' developing knowledge about people and object causal action. This is because an associative learning mechanism—implemented in an artificial neural network—was sufficient to account for infants' looking behavior in Spelke et al. (1995).

However, an open question concerns whether "younger" networks assigned either to the People condition or to the Inanimate Object condition respond to the Contact and No Contact test events in the same way as "older" networks (i.e., the networks in Simulation 1). If infants' causal knowledge about people and objects is underpinned by an associative-learning mechanism, then younger networks (and by extension, younger infants)—who have had much less time to encode associations between the features of objects and entities and different kinds of causal action—should show less heightened "looking" when an object moves in the absence of contact compared to older networks with more experience with the relevant associations. This would provide insight into whether an associative learning mechanism—as it is instantiated into an artificial neural network model—predicts that infants' causal knowledge about people and inanimate objects should undergo a developmental progression. This is important to demonstrate because the Core Knowledge perspective makes no such development prediction given that infants' causal knowledge about people and inanimate objects is assumed to be present from birth (or very shortly thereafter). Thus, the goal of Simulation 2 is to explore what predictions, if any, the model makes for "younger" networks (and by extension, for infants younger than 7 months of age in Spelke et al. [1995]).

Simulation 2

The aim of Simulation 2 was to explore whether an associative learning mechanism—as it is implemented in an artificial neural network—predicts a developmental progression in infants' knowledge about people and object causal. To examine this question, networks "younger" than those tested in Simulation 1 were tested in Simulation2.

Method, Training, and Testing

Simulation 2 was similar to Simulation 1 except that networks received 300 (rather than 600) epochs of pretraining experience. In addition, the learning rate, weight decay, and number of hidden units were set, respectively, to .001, .01, and 8. Recall that in Simulation 1, these parameters were set, respectively, to 0.05, 0.9, and 0.001. This modeled the fact that necessarily younger infants have had less real-world experience than older infants as well as the fact that the information-processing capacities of younger infants—which were captured by the learning rate, weight decay, and number of hidden units—is less robust than older infants. Together, these parameters implemented a very simple model of age and development that is consistent previous connectionist simulation studies[17-19]. Future research should examine parametrically the effect of these changes on networks' performance.

Results

Figure 3 shows the networks' mean "looking times" to the Contact and No Contact test events across the People and Inanimate Object conditions. Given evidence of non-normality in the model's response to the test trials across conditions based on a Shapiro-Wilk test of normality, $W = 0.67$, $p < .001$, data were analyzed using non-parametric permutation tests with 10,000 replications. As can be seen in the Figure 3, networks assigned to the Inanimate Object condition "looked" equally long at (i.e., produced equivalent sum-squared error to) the Contact ($M = 443.35$) and No Contact ($M = 443.35$) test events, $p = .85$. Likewise, networks assigned to the People condition "looked" equally long at the Contact ($M = 443.28$) and No Contact ($M = 443.29$) test events, $p = 1$.

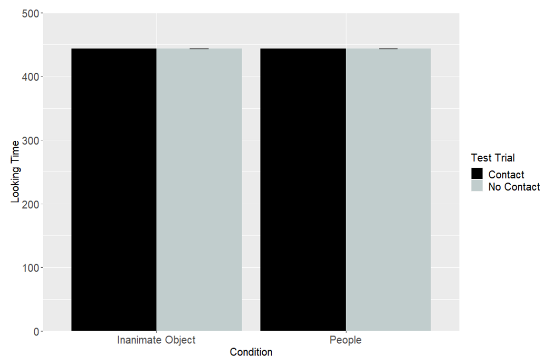

Figure 3. Networks' mean "looking time" (i.e., SSE) to the Contact and No Contact test events across conditions.

Discussion

Simulation 2 revealed that regardless of the condition to which networks were assigned they "looked" equally long at the No Contact and Contact test events. This result was due to the fact that networks in Simulation 2 possessed less robust information-processing capacities than models in Simulation 1. In addition, models in Simulation 2 received less pretraining experience—which corresponds roughly to "real-world" experience—than models in Simulation 1. This finding is important because it makes a testable prediction, which should be explored in future behavioral research: Infants younger than 7 months of age (though it is not clear *how* much younger given that the parameters used to model age here may not map linearly onto age in the real-world) should look equally long at the Collision and No Collision test events across conditions. This simulation also makes a second testable prediction: Younger infants should show greater looking overall compared to older infants. This prediction is based on the fact that, averaged across condition and test trial, younger networks ($M = 443.32$, $SD = 0.03$) showed longer looking (i.e., produced greater sum-squared error) overall than older networks ($M = 56.76$, $SD = 8.44$), $t(62) = 1464.9$, $p < .0001$. Importantly, this finding is unlikely to be an artifact of this simulation and accords well with previous developmental findings. This research has indicated that younger infants do tend to look longer at stimuli compared to older infants, which is a finding that may reflect older infants' greater information-processing capacities compared to that of younger infants[15-16]

An open question that remains concerns what predictions the network (and by extension, the theory) makes for modified stimuli such as people with object parts (e.g., a person whose legs have been replaced with the bottom half of an object) and objects with people parts (e.g., an object with human legs). If the associative-learning account is correct that infants' knowledge about human and object causal action reflects learned associations between particular surface features (e.g., legs) and different kinds of causal action (e.g., contact and action-at-a-distance causality), then infants assigned to the Inanimate Object condition should show equivalent looking towards the Contact and No Contact test events. In contrast, infants assigned to the People condition should look longer at the No Collision test event than the Collision test event. This is because infants may have learned an association between the bottom-half of objects and contact causality only. However, this prediction remains speculative, and it is unclear whether a connectionist model—which implements this associative account—would make this prediction.

Simulation 3

Simulation 3 explored how a connectionist system "treats" modified people and object stimuli.

Method, Training, and Testing

Simulation 3 was similar to Simulation 1 except that networks assigned to the People condition were habituated to people without legs, whereas networks assigned to the Inanimate Object condition were habituated to objects that possessed legs. Networks' pretraining experienced was identical to the first two simulations.

Results

Figure 4 shows the networks' mean "looking times" to the Contact and No Contact test events across the People and Inanimate Object conditions. The data revealed that the difference in "looking time" (i.e., difference in SSE) to the No Contact and Contact test events was greater for models assigned to the People with Legs condition (N = 64) ($M_{\text{No-Collision}}$ = 50.86, $SD_{\text{No-Collision}}$ = 4.84; $M_{\text{Collision}}$ = 63.55, $SD_{\text{No-Collision}}$ = 10.52) than models assigned to the Objects with Legs condition (N = 64) ($M_{\text{No-Collision}}$ = 52.32, $SD_{\text{No-Collision}}$ = 5.33; $M_{\text{Collision}}$ = 53.55, $SD_{\text{No-Collision}}$ = 4.94), $t(62) = -54.91$, $p < .0001$.

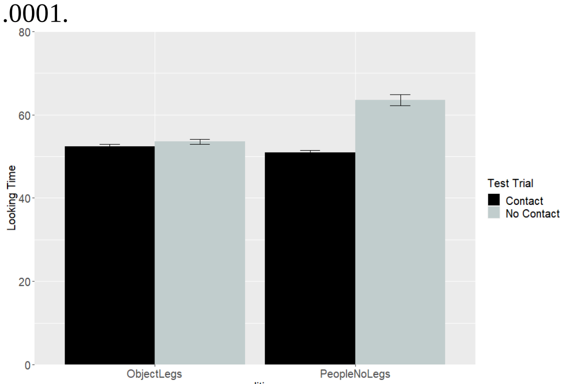

Figure 4. Networks' mean "looking time" (i.e., SSE) to the Contact and No Contact test events across conditions.

Discussion

The results of Simulation 3 indicated that networks assigned to the People with No Legs condition looked longer at the No Collision test event than at the Collision test event whereas networks assigned to the Objects with Legs condition looked equally long at both test events. Below I discuss implications of this finding.

General Discussion

The aims of the present series of simulations were twofold. First, Simulation 1 was designed to determine whether an associative learning mechanism—implemented in an artificial neural network—is sufficient to account for infants' looking behavior in Spelke et al. (1995). Second, Simulation 2 was designed to determine whether this perspective predicts a developmental progression in infants' knowledge about people and object causal action. Third, Simulation 3 tested what predictions this associative learning perspective makes for objects that possessed legs and people that did not possess legs. The data from the present series of simulations are important because they indicate not only that an associative learning mechanism can explain infants' looking behavior in Spelke et al. (1995) but that this mechanism predicts that infants' knowledge about people and object causal action should undergo a developmental progression.

These results are also important because they make a second testable prediction: Infants assigned to the People with No Legs condition should look longer at the No Collision test event than at the Collision test event, whereas infants assigned to the Object with Legs condition should look equally long at both test events. Note that the Core Knowledge perspective would not make these predictions. This is because proponents of this perspective maintain that infants' causal knowledge about people and objects is present from birth and is insensitive to the particular surface features of an object or person—infants' causal knowledge about people and objects is assumed, instead, to be abstract and extends beyond the immediate perceptual input. Thus, it should be possible to determine whether core knowledge or an associative learning mechanism underlies infants' causal knowledge about people and object by testing infants younger than 7 months of age as well as exposing infants to modified people and object stimuli. Future research would benefit from testing these competing predictions. Behavioral research designed to address these issues is in the beginning stages in my lab. Although we do not extend the present proposal to other studies, there is no reason to think that the present account could not be extended to explain studies such as that by Muentener and Carey (2010)—they showed that 8-month-olds understand that human hands, but not inanimate objects, can cause state changes in other things. Finally, it is worth mentioning that the present series of simulations assume that infants' responses to the test events was due to learned associations between legs and different kinds of causality, but it is an open question whether legs (or some other feature) is involved in this association. The point of the present series of simulations was simply to provide a proof of concept that associative learning is sufficient to explain infants' developing knowledge about people and object causal action; one need not invoke innate knowledge or core systems to explain this knowledge.

References

1. Bélanger, N. D., & Desrochers, S. (2001). Can 6-month-old infants process causality in different types of causal events?. British Journal of Developmental Psychology, 19(1), 11-21.
2. Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. Cognitive Development, 5(2), 193-207.
3. Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality?. Cognition, 25(3), 265-288.
4. Rakison, D. H., & Krogh, L. (2012). Does causal action facilitate causal perception in infants younger than 6 months of age?. Developmental science, 15(1), 43-53.
5. Oakes, L. M. (1994). Development of infants' use of continuity cues in their perception of causality. Developmental Psychology, 30(6), 869.
6. Kominsky, J. F., Strickland, B., Wertz, A. E., Elsner, C., Wynn, K., & Keil, F. C. (2017). Categories and constraints in causal perception. Psychological Science, 28(11), 1649-1662.
7. Cohen, L. B., Rundell, L. J., Spellman, B. A., & Cashon, C. H. (1999). Infants' perception of causal chains. Psychological Science, 10(5), 412-418.
8. Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. Developmental science, 10(1), 89-96.
9. Spelke, E. S., & Kinzler, K. D. (2009). Innateness, learning, and rationality. Child Development Perspectives, 3(2), 96-98.
10. Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action.
11. Elman, J. L. (1990). Finding structure in time. Cognitive science, 14(2), 179-211.
12. Mareschal, D., French, R. M., & Quinn, P. C. (2000). A connectionist account of asymmetric category learning in early infancy. Developmental psychology, 36(5), 635.
13. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. nature, 323(6088), 533-536.
14. Sirois, S., & Mareschal, D. (2002). Models of habituation in infancy. Trends in Cognitive Sciences, 6(7), 293-298.
15. Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. Cognition, 83(2), B35-B42.
16. Slone, L. K., & Johnson, S. P. (2015). Infants’ statistical learning: 2-and 5-month-olds’ segmentation of continuous visual sequences. Journal of experimental child psychology, 133, 47-56.
17. Benton, D. T., Rakison, D. H., & Sobel, D. M. (2021). When correlation equals causation: A behavioral and computational account of second-order correlation learning in children. Journal of Experimental Child Psychology, 202, 105008.
18. Benton, D. T., & Lapan, C. (2022). Moral masters or moral apprentices? A connectionist account of sociomoral evaluation in preverbal infants. Cognitive Development, 62, 101164.
19. Rakison, D. H., & Lupyan, G. (2008). Developing object concepts in infancy: An associative learning perspective. Monographs of the Society for Research in Child Development, i-127.
20. Muentener, P., & Carey, S. (2010). Infants’ causal representations of state change events. Cognitive psychology, 61(2), 63-86.