INFANCY, *1*(4), 403–428

Object Permanence in Five-and-a-Half-Month-Old Infants?

Richard S. Bogartz and Jeanne L. Shinskey
Department of Psychology
University of Massachusetts, Amherst

Thomas H. Schilling
Department of Behavioral Sciences
Fitchburg State College

Event Set × Event Set designs were used to study the rotating screen paradigm introduced by Baillargeon, Spelke, and Wasserman (1985). In Experiment 1, 36 5½-month-old infants were habituated to a screen rotating 180° with no block, a screen rotating 120° up to a block, or a screen rotating 180° up to and seemingly through a block. All infants were then tested on the same 3 events and also a screen rotating 120° with no block. The results indicate that infants are using novelty and familiarity preference to determine their looking times. To confirm this, in Experiment 2, 52 5½-month-old infants were familiarized on either 3 or 7 trials to a screen rotating 180° with no block or a screen rotating 120° with no block. All infants were then tested on the same test events as in Experiment 1. Infants with fewer familiarization trials were more likely to prefer the familiar rotation event. The results of these 2 experiments indicate that infants did not use the possibility or impossibility of events but instead used familiarity or novelty relations between the habituation events and the test events to determine their looking times, and suggest that the Baillargeon et al. study should not be interpreted as indicating object permanence or solidity knowledge in young infants.

In the decade and a half since its publication, the study of object permanence in 5-month-old infants by Baillargeon, Spelke, and Wasserman (1985) has been frequently cited in research and theoretical articles and plays a supportive role in vari-

Requests for reprints should be sent to Richard S. Bogartz, Department of Psychology, University of Massachusetts, Amherst, MA 01003. E-mail: bogartz@psych.umass.edu

ous formulations of the young infant's cognitive capacities. In that study, 5-month-old infants were familiarized with a screen moving back and forth 180° in the manner of a drawbridge. The infants were then shown two events. In one event, the screen repeatedly started out flat on the table pointing toward the infant, rising up away from the infant and approaching a box, occluding the box, stopping at the box, and returning, disclosing that the box is still there. In the other event, the screen repeatedly started out flat on the table pointing toward the infant, rising up away from the infant and approaching a box, occluding the box, and then apparently performing the impossible by continuing on downward away from the infant, traversing 180°, and seemingly passing through the space occupied by the box until flat on the table, and then returning, disclosing that the box is still there. The infants looked longer at the apparently impossible event. The longer looking usually has been interpreted in terms of the infants expecting the screen to stop at the box, being surprised or puzzled at its not stopping, and therefore looking longer. This interpretation, in turn, has supported the conclusion that these infants must know that the box continues to exist after it is occluded and must know that one solid object cannot pass through another or there would be no surprise.

Piaget (1954) concluded that young infants do not have object permanence. He showed that infants this age and even some months older will not displace an occluder that is concealing an object that the infant would otherwise reach for. There is a discrepancy between this often replicated finding of Piaget and the apparent knowledge of object permanence in the drawbridge experiment and other looking-time experiments that appear to demonstrate object permanence in young infants (Baillargeon, 1993; Baillargeon & Graber, 1987; Kellman & Spelke, 1983; Spelke, Breinlinger, Macomber, & Jacobson, 1992). The usual response to this discrepancy is to argue that in young infants looking time is a more sensitive measure of the infant's knowledge and that the infant does not displace the occluder to get the occluded object because he or she cannot coordinate the displacement action with the reaching action: He or she cannot coordinate the means with the ends. However, recent work suggests that the means–ends explanation probably will not be sufficient to account for failures to reach for occluded objects (Munakata, McClelland, Johnson, & Siegler, 1997; Shinskey, 1999).

The claim that young infants have object permanence has not been received with universal enthusiasm (Fischer & Bidell, 1991). Haith (1988) suggested concern for artifacts and short-lived afterimages as possible explanations for some of the results indicating object permanence. Meltzoff and Moore (1998) argued that these experiments may show "representation of past events" but not "object permanence" (p. 215). That is, they may be based on infants detecting a relation between the familiarized material and the test material without necessitating object permanence. Along similar lines, Bogartz, Shinskey, and Speaker (1997) and Bogartz and Shinskey (1998) suggested that studies that appear to reveal object permanence and a sophisticated knowledge of what is possible and what is impossible may perhaps more parsimoniously be understood in terms of well-known and well-documented preferences for novel events under some circumstances and for familiar events under others (for documentation of familiarity preferences in infants, see Hunter & Ames, 1988; Hunter, Ross, & Ames, 1982; Rose, Gottfried, Melloy-Carminar, & Bridger, 1982). In the typical looking-time experiment, the infant is habituated or familiarized to one stimulus event and then is tested on two events. Elsewhere (Bogartz & Shinskey, 1998; Bogartz et al., 1997), we have detailed the problems with this design that tend to confound possibility–impossiblity preference with novelty–familiarity preference. We have also shown how using the Event Set × Event Set design together with a new class of models that we have introduced, permits the explicit statistical test of whether the impossibility of events is having an effect on looking time. This new approach allows for the possibility that looking time differences result from the novelty or familiarity relations between the familiarization events and the test events rather than as a result of surprise on the part of the infant to so-called impossible events.

In the Event Set × Event Set design, instead of infants being familiarized to one event and then tested on two others, all of the events being used are candidates for both the familiarization or habituation stage and the test stage of the experiment. This permits a study of the effect of the relation between the familiarized–habituated event and the test event on test event looking times. We can illustrate how the drawbridge experiment of Baillargeon et al. (1985) can be recast as an Event Set × Event Set design. Let the familiarization event and the two test events in that design be denoted by 180, 120 + B, and 180 + B. In this notation, the numbers 180 or 120 indicate the degrees through which the occluding screen rotates away from the infant before reversing direction, and the +B appended to the number indicates that a block was placed beyond the screen's axis of rotation such that a 120° rotation would bring the screen up to the location of the block and a 180° rotation would cause the block to pass completely through the space occupied by the block. Thus, the block is occluded when the screen rotates away from the infant and then disoccluded when it rotates back toward the infant. (Of course, on 180 + B trials the experimenter removes the block after it is occluded so that the impossible event can occur and then replaces the block before the disocclusion occurs.) Each of these events would serve as the familiarization event for one third of the participants, forming a between-participant factor of the design. Following familiarization, each infant would be tested on all of the events in random order, forming a within-participants factor of the design. In the design that we actually used in Experiment 1, an additional event denoted by 120 and consisting of the screen repeatedly rotating 120° and back, but with no block present and then returning to its starting position, was added as a fourth level of the within-participants factor. Thus the design is as shown in Figure 1.

Figure 1 also shows the predictions from a general model that can be reduced to two theoretically relevant special cases. In the general model, the level of looking (L) is increased or decreased by amounts specific to two kinds of effects. The first

Habituation Trials (Between)	Test Trials (Within) 180	120+B	180+B	120
180	L-CB-CS-P	L+CB+CS-P	L+CB-CS+P	L-CB+CS-P
120+B	L+CB+CS-P	L-CB-CS-P	L-CB+CS+P	L+CB-CS-P
180+B	L+CB-CS-P	L-CB+CS-P	L-CB-CS+P	L+CB+CS-P

FIGURE 1 The Event Set × Event Set design for Experiment 1, using the Baillargeon et al. (1985) stimuli. *Note.* L = the general level of looking; CS = the effect of a change in the screen rotation condition from 180 on habituation to 120 on test or vice versa; –CS = the effect of no change; CB = the effect of a change in the block condition from presence of block during habituation to absence of block during test or vice versa; –CB = the effect of no change; P = the effect of an impossible event on the test trial; –P = the effect of a possible event on the test trial.

kind results from the comparison of the current event with the stored information about the event to which the infant was familiarized or habituated. One component of this first kind of effect has to do with the relation of the perceived screen to the screen in the habituation event; specifically, it is the effect of seeing a change in the rotation of the screen (CS). The other component has to do with the relation of the perceived block to the block in the habituation event; it is the effect of seeing a change in the presence of a block (CB). For each of these two parameters, a positive value reflects a novelty preference and a negative value reflects a familiarity preference. A particular virtue of this modeling approach is that it leaves the question open as to which type of preference exists for a given individual and regards the question as one of parameter estimation rather than assuming automatically that infants always prefer the novel stimulus.

The second kind of effect is an increase in the amount of looking due to surprise at the occurrence of an impossible event (P). In Figure 1, a positive value of P denotes preference for looking at the impossible event. Under the general model, looking-time effects can be due to both surprise at impossibility and to the novelty–familiarity relations between the familiarization or habituation events and the test trial events. Thus, for example, the model asserts that looking at a 120 + B test trial following habituation to a 180 trial would be described by the equation

$$\text{Looking time} = L + CS + CB - P$$

because there was a change in the screen from 120 to 180, a change from block present to block absent, and there was nothing impossible about the screen rotating 120° up to but not through the block.

If all that determines the differences in looking time is whether the event being looked at is possible or impossible, then, $CB = CS = 0$ and $P > 0$. This may be the case if the infant represents the occluded block, knows that the rotating screen cannot pass through the block, is surprised when it appears to do so on the 180 + B trials, and therefore looks longer. We call this the *possibility model* and take it to represent the assumptions held by Baillargeon et al. (1985). On the other hand, suppose that the infant has no knowledge concerning what is possible and what is impossible, so that P = 0. Instead, looking time is determined by the relation between the test event and the familiarization–habituation event. In this case, a quantity CB is added to general level of looking, L, when there is a change from block present in the familiarization–habituation event to block absent in the test event, or vice versa. When there is no change in block presence from the familiarization–habituation event to the test event, then that quantity CB is subtracted. Similarly, a quantity CS is added to the general level of looking, L, when there is a change in the amount of rotation of the screen from the familiarization–habituation event to the test event, but when there is no change in the amount of screen rotation, then that quantity CS is subtracted. The sign of the CB and CS parameters are not necessarily the same for all infants. We call this a *novelty–familiarity comparison model.*

Under the possibility model, the matrix of predicted values becomes that shown in Figure 2A. Under the novelty–familiarity comparison model, the matrix of pre-

(A) The possibility model

Habituation Trials (Between)	Test Trials (Within) 180	120+B	180+B	120
180	L-P	L-P	L+P	L-P
120+B	L-P	L-P	L+P	L-P
180+B	L-P	L-P	L+P	L-P

(B) The perceptual processing model

Habituation Trials (Between)	Test Trials (Within) 180	120+B	180+B	120
180	L-CB-CS	L+CB+CS	L+CB-CS	L-CB+CS
120+B	L+CB+CS	L-CB-CS	L-CB+CS	L+CB-CS
180+B	L+CB-CS	L-CB+CS	L-CB-CS	L+CB+CS

FIGURE 2 The matrix of predicted values under (A) the possibility model and (B) the novelty–familiarity comparison model. *Note.* L = the general level of looking; CS = the effect of a change in the screen rotation condition from 180 on habituation to 120 on test or vice versa; –CS = the effect of no change; CB = the effect of a change in the block condition from presence of block during habituation to absence of block during test or vice versa; –CB = the effect of no change; P = the effect of an impossible event on the test trial; –P = the effect of a possible event on the test trial.

dicted values becomes that shown in Figure 2B. The two matrices of predicted values provide clearly distinct predictions that enable an empirical test to determine which model better fits the data. A third case, of course, is one in which both the possibility–impossiblity effects and the novelty–familiarity preference effects are playing a role in determining looking time. This case is represented in Figure 1.

Inspection of the second and third cells in Row 1 of the matrix in Figure 2B shows that the two cells constitute a replication of the Baillargeon et al. (1985) test trials on the possible and impossible events following habituation to the 180° screen without a block. Consider the values in these two cells that are predicted by the novelty–familiarity comparison model. If the value of CS is negative, then the infant will look longer at the 180 + B event than at the 120 + B event. A negative value of CS is interpreted as a preference for the familiar screen rotation rather than for the novel one. From our perspective, this familiarity preference occurs because the infant is not finished forming a stable representation of the 180° rotation.

The model reveals that the familiarity preference effect for the screen rotation can produce the type of looking-time difference found by Baillargeon et al. (1985). Under this analysis, the difference occurs not because the infant knows what is possible and what is impossible and not because the infant knows that the block continues to exist behind the occluding screen, but instead, because of a familiarity preference effect, which, in the Baillargeon et al. study, is confounded with the possibility–impossibility variable. Thus, we see that the data from the Baillargeon et al. study cannot decide between the two models because both models can predict the obtained result. The purpose of Experiment 1 is to use the Event Set × Event Set design in Figure 1 to determine which of these competing models better describes infant looking.

EXPERIMENT 1

Method

Participants. Thirty-six, full-term 5½-month-old infants (16 girls, 20 boys) from the Amherst, Massachusetts, area participated in the study. The infants ranged in age from 5 months 3 days to 6 months 6 days, with a mean of 5 months 19 days. Nine additional infants were tested but were not included in the sample, 7 because of fussiness, 1 because of experimenter error, and 1 because of inattention. Participants were recruited from state birth records. Parents were contacted by mail and a subsequent telephone call. Participation was voluntary.

Apparatus. Testing took place in a brightly lit experimental room (300 cm × 225 cm). A large wooden stage (203 cm × 141 cm × 70 cm) was located in the center of the room. The displays were presented within a wooden, three-sided box (87 cm

× 89.5 cm × 75.5 cm), painted black and set inside the stage. The screen apparatus was placed on a wooden platform such that the infant's eye level was approximately at the center of the screen. The screen apparatus consisted of a yellow tissue-covered foam core board (20 cm × 16 cm) attached to a rotating rod mounted at the center of a black painted wooden box (22.5 cm × 35.5 cm × 9.5 cm) below it. The screen was rotated by a motor inside the black box and was operated by the experimenter using a keyboard interface to an IBM computer. The screen could travel from 0° to 120° and back in 8 sec, or from 0° to 180° and back in 12 sec. The block consisted of a foam core board (17 cm × 11 cm) covered with red tissue on which gold stars were pasted. The back of the block was painted black so that it could not be seen in its resting position flat on the surface of the black box below it. The block was raised on a rod to its 90° standing position with clear fishing line pulled by the experimenter through a hole in the back wall of the stage. The fishing line was attached to a small handle on the side of the black box, and the handle was attached to the rotating rod. A black cloth curtain was lowered over the stage opening between trials. The curtain had pictures of infants' faces attached to it, in an attempt to prevent the infant from becoming fussy during intertrial intervals.

A video camera focused on the infant's face was centered behind the back wall of the box, just above the floor of the box, so that only the lens protruding through a hole in the back wall was visible to the infant. A second camera on a tripod above and behind the infant focused on the displays on the stage. The video monitor for the camera focusing on the infant's face was used by the observer in a separate adjacent observation room. Both the displays and the infant's looking behavior were recorded and served to check on the accuracy of looking times and the reliability of interobserver ratings. Looking times were recorded online using a keyboard interface to a Macintosh SE/30 computer operated by the observer in a separate room. This observer was blind to the experimental events manipulated by the experimenter in the adjacent room and could see only the infant's face on the video monitor in the observation room.

Displays. Four displays were presented to the infants. Each display began with the screen in its resting position at 0°, flat on the surface of the black box below it. The 180 display consisted of the screen moving by itself, back and forth in a 180° arc, toward and away from the infant, so that it took 12 sec to complete the rotation from 0° to 180° and back again. In the 180 + B display, the impossible event, the red block was visible for the first 1 sec of the rotation; the screen moved all the way to 180° as if the block were not there, and then the screen moved back toward 0°, with the block becoming visible again for the last 1 sec of the rotation. This was accomplished by having the experimenter drop the block to its resting position flat against the surface of the black box below it during the middle portion of the 12-sec rotation. In the 120 + B display, the red block was visible for the first 1 sec of the rotation; the screen rotated up to the position of the block at 120° and then reversed its direction, with the red block becoming visible again for the last 1 sec of the rotation. In the 120 display, the block was not present, and the screen rotated 120° away from the infant and then back to 0°. The 180, 180 + B, and 120 + B displays served as between-participant habituation events. The 120 display was added to these three to make four within-participants test events.

Design. The design was a 3 (habituation events, between-participants) × 4 (test events, within-participants) factorial design. Twelve participants were habituated to the 180 display, 12 to the 180 + B impossible display, and 12 to the 120 + B display. Following habituation, each infant was presented with four test trials, one for each of the four test displays. The event sequences on the four test trials were constrained such that if the infant was habituated to a screen motion of 180°, then the first test trial would use a screen motion of 120°, and vice versa. Within each habituation group, this constraint allowed for 12 possible test trial event sequences. Within each of the three groups of 12 infants, each infant received a different one of these test trial sequences of the four test displays.

Procedure. The infant was placed in an infant seat attached to a table so that the infant's face was approximately 100 cm from the stimuli, and the infant's eye level was approximately 40 cm above the floor of the box. The parent sat to the infant's right with his or her back to the stage. The parent was asked not to look at the displays, so as not to influence the infant's behavior, and not to interact with the infant unless the infant became fussy.

Each trial began with the raising of the curtain. In the two displays involving the block, the block was presented first until the infant had been judged to look at it, and then the screen began moving. Looking time started with the first look at the block after the curtain rose. In the two displays not involving the block, the screen did not begin moving until the infant was judged to be looking toward the apparatus on the stage. Looking time and the trial ended when the infant was judged to look away from the display for 2 consecutive sec after looking for at least 4 cumulative sec, or to look for a total of 120 sec. Habituation trials continued until 14 trials were presented, or until the habituation criterion was met: total looking time on three consecutive trials that was less than 50% of the sum of the looking times on the first three trials exceeding 12 sec. The end of each trial was signaled by a beep from the computer and the experimenter lowered the curtain on signal from the computer. The intertrial interval for habituation trials was approximately 15 sec.

When the computer signaled the end of the last habituation trial, the curtain dropped and two pretest trials were presented. The pretest trials are included because they were included in the experiment by Baillargeon et al. (1985). In the pretest trials, the curtain was raised to reveal the red block standing alone. The screen lay flat against the floor of the alley with the block clearly visible behind it. The curtain was dropped when the infant had been judged to look at the block for 3 cumulative sec. At the end of the second pretest trial, the curtain was dropped and the test trials began. The intertrial interval for test trials was approximately 10.5 sec (based on the measurement of 25% of the intertrial intervals). The infant received four test trials so that each test display was seen once. As in the habituation phase, trials ended when the infant was judged to look away for 2 sec after looking for at least 4 cumulative sec, or to look for a total of 120 sec.

Results

Interobserver reliability. All of the 36 infants except 1 were scored by two observers. Interobserver reliability was assessed by determining the correlation between the judged total looking time on each of the first four test trials for one observer (Jeanne L. Shinskey) and those for the other (Richard S. Bogartz). The correlation was .992 and compares favorably with previous such values for our lab that have been around .98. The slope of the regression line was .996, indicating that, except for an additive constant, the two observers were judging the looking times equivalently.

Habituation. The mean number of trials to criterion for the 180, 180 + B, and 120 + B habituation groups were 7.50, 6.75, and 7.58, respectively. These means did not differ significantly, $F(2, 33) = .80$, $p = .46$. The mean total looking times during habituation for the 180, 180 + B, and 120 + B habituation groups were 244.34 sec, 169.69 sec, and 243.45 sec, respectively. The difference between the means was marginally significant, $F(2, 33) = 3.067$, $p = .06$. The difference results from there being less total looking to 180 + B than to the other two events. It is interesting to note that the 180 + B condition involves habituation to the impossible event of the screen appearing to pass right through the block. If such an impossible event is surprising to the infant, one would expect that habituation would take longer, but the data do not lend any support to that prediction. This is consistent with other previous failures to find slower habituation to the impossible event (Bogartz & Shinskey, 1998; Bogartz et al., 1997). The trials to criterion were essentially the same for both the 180 and the 180 + B conditions and the difference in total looking time, if real, is in the wrong direction.

Rivera, Wakeley, and Langer (1999) claimed that infants prefer to look at screen rotations that have more motion (180°) rather than at those with less motion (120°). The habituation trial looking times obtained in this experiment do not support that claim.

Test trials. Table 1 shows the results of fitting the model in Figure 2B to the test trial data. This was done using multiple regression including an effect for individual participants and a trials effect. Bogartz et al. (1997) provided a tutorial on the

TABLE 1
Observed and Predicted Mean Looking Times For the Event Set × Event Set Design in Figure 2B

Group	*180*	*120 + B*	*180 + B*	*120*
180				
Observed	11.04	18.71	13.89	14.21
Predicted	11.24	17.53	12.68	16.39
120 + B				
Observed	20.35	10.39	18.86	16.44
Predicted	20.27	12.92	18.96	13.90
180 + B				
Observed	13.33	17.60	12.09	20.11
Predicted	12.16	19.41	12.16	19.41

procedures but they did not show how to include a trials effect. There are four test trials. The trial effect for Test Trial *i* (*i* = 1–4) is defined as the mean looking time for Test Trial *i* minus the overall mean looking time. Thus, the test trial effects sum to zero as in a fixed-effects factor in the analysis of variance (ANOVA); including the test trial effects in the regression analysis extracts from error the variability due to the responses occurring on different trials. The results in Table 1 show that the model is fitting the data closely and that there is substantial variability from cell to cell that is predicted by the model parameters, CB and CS. The mean absolute deviation of observed means from predicted means is 1.53 sec. The regression analysis yielded an R^2 of .624 which is significant, $F(39, 105) = 12.18$, $p = .00$.

To test the effect of the theoretical variables CB and CS, the model containing only L, participants, and trials was fit and the difference between the model including CB and CS and the model excluding CB and CS was tested. The difference was statistically significant, $F(2, 105) = 13.69$, $p = .00$, supporting the model in Figure 2B.

There are two ways to test whether the possibility–impossibility variable is playing a part. The first is to compare the full model in Figure 1 plus participants and trials effects against the model in Figure 2B plus participants and trials effects. Because the model in Figure 2B is a special case of the full model, this test is done using the difference between the sum of squares for regression under the full model and the sum of squares for regression under Figure 2B. This difference, divided by the difference in the number of parameters, one, is a mean square that can be divided by the mean square for error under the full model to obtain an *F* ratio. When this was done, the obtained $F(1, 104) = .0097$, $p = .92$ was not significant, indicating that the P parameter accounted for no additional variance beyond that accounted for by the model in Figure 2B.

To protect against the possibility that the CB and CS effects are correlated with the P effects and are thus soaking up variance that would otherwise be accounted for by the P variable, the second way to test whether the P variable is playing a part is to exclude the variables CB and CS from the analysis and test whether the model including P, participants, and trial effects does significantly better than the model including only participants and trials effects. When this was done, the obtained $F(1, 106) = .03$, $p = .86$, was not significant, indicating that the P parameter accounted for no additional variance beyond that accounted for using only the participant effects and the trials effects.

Inspection of Table 1 reveals some interesting aspects of the data. As would be expected from the model in Figure 2B, the longest looking times are observed when there is a combined change from one screen rotation to the other and from one block condition to the other. Thus, for example, going from 120 + B to 180 or from 180 + B to 120 produces the longest looking times. This suggests that there is both a screen novelty and a block novelty preference in these infants, at least on average.

Because the first row of Table 1 includes a replication of the Baillargeon et al. (1985) conditions, namely testing on 120 + B and 180 + B following habituation on 180, it is of interest to observe the results in the second and third cells of the first row of Table 1. Following habituation to the 180 screen, the infants on average look at the 120 + B event for 18.71 sec but look at the 180 + B for only 13.89 sec. We see that there is a suggestion of a screen novelty effect here too, but this difference is not significant. We attribute the location of these means to the averaging of the means for two different subgroups (see Figure 3) in which there were more in-

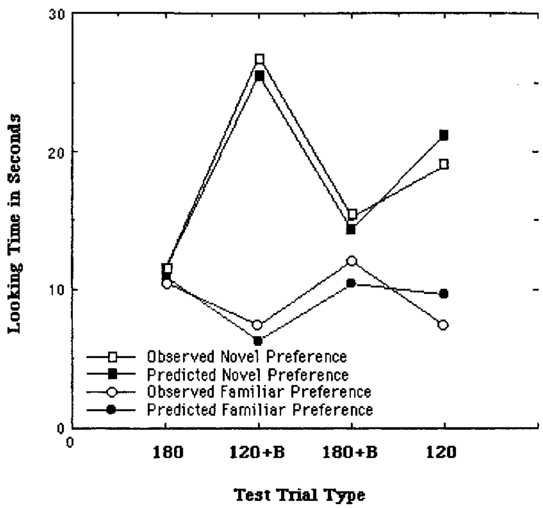

FIGURE 3 Mean observed and predicted looking times on the four test trials following habituation to the 180 screen alone plotted separately for infants showing a screen novelty preference ($n = 7$) or a screen familiarity preference ($n = 5$).

fants with a novelty preference and fewer with a familiarity preference. This is clarified in the following discussion of infant sorting.

Because we found it useful in a previous investigation (Bogartz & Shinskey, 1998) to sort infants with respect to whether they had a novelty preference or a familiarity preference, we decided to do the same type of sorting here. Infants habituated to 180 or 180 + B were defined to have a screen novelty preference if their average look to the test trials 120 + B and 120 was longer than their average look to the test trials 180 + B and 180. Infants habituated to 120 + B were defined to have a screen novelty preference if their average look to the test trials 180 + B and 180 was longer than their average look to the test trials 120 + B and 120. Infants habituated to 180 were defined to have a block novelty preference if their average look to the test trials 120 + B and 180 + B was longer than their average look to the test trials 180 and 120. Infants habituated to 120 + B or 180 + B were defined to have a block novelty preference if their average look to the test trials 180 and 120 was longer than their average look to the test trials 180 + B and 120 + B. The regression analysis was performed using coding that reflected these preferences. Thus, for a given infant, the coding for the CS and CB parameters was either 1 and 1, 1 and –1, –1 and 1, or –1 and –1. For example, an infant who preferred novelty for the screen and familiarity for the block would have the coding 1 and –1. In all, then, the regression analysis included an overall level L, a trial effect, an individual participant level for each participant, and the CS and CB parameters coded for novelty or familiarity preference. (When the possibility parameter was included in the analysis, all of the previously mentioned effects remained and the one possibility parameter P was added to the regression analysis.)

The predicted values shown in Table 1 are based on the analysis described previously with the exclusion of the possibility parameter. The observed and predicted values from this analysis for only the infants habituated to 180 are presented separately in Figure 3 for the infants who had a screen novelty preference and the infants who had a screen familiarity preference. (Note that although the two curves are sorted by whether the infants preferred screen novelty or screen familiarity, the curves are obtained from the regression analysis that used coding for both the screen preferences and the block preferences.) For infants with a screen familiarity preference, the Baillargeon et al. (1985) result is replicated: longer looking at the impossible 180 + B event than at the possible 120 + B event. For the infants showing a screen novelty preference just the reverse effect is shown.

To summarize, it is our interpretation that the familiarity preference subgroup replicates the Baillargeon et al. (1985) results, and the novelty preference subgroup gives the reverse finding because infants presented with the same events, and even run to the same habituation criterion, can individually differ in the amount of stimulus information they have processed. Those who have processed the events less will prefer to look more at the familiar events. Those who have processed the events more will prefer to look more at the novel events. In the

Baillargeon et al. study there was a predominance of infants who preferred the familiar screen event. We believe that the overall means can obscure the individual infant preferences and that individual parameter estimation with sorting into subgroups can help to avoid such obscuring.

As Bogartz and Shinskey (1998) noted, the separation of the infants according to the sign of their parameter estimates requires some justification. The variation in the sign of the parameter values may be simply variation around a mean value at or close to zero. If this were the case, selection of the infants by the sign of their parameter estimate could be viewed as merely using deviations due to error to artificially create two groups that in fact differ only due to error. Bogartz and Shinskey introduced a statistical test for deciding whether the variation of the parameter estimates was greater than may be expected merely as a result of chance variation. They determined what the theoretical variation of the parameter estimates should be under the null hypothesis of chance variation and created an F test based on this determination. Basically, the measure they used was the mean squared variation of the parameter estimates from their own mean. The test they proposed is adequate as long as the parameter values do not cluster into groups of relatively homogeneous values that differ between clusters but not within. If this is the case, the method they introduced assigns too many degrees of freedom to the variation of the parameter estimates. In this context, we have applied an alternative method for approaching the problem.

To show that the variation in the parameter estimates of the block and screen effects was greater than that due to chance, it is sufficient to show that each of these sets of estimates covaries significantly over infants with the estimated total amount of looking time on the test trials. This is so because the estimated total looking time is the mean of the looking times on the test trials, whereas the block and screen effects are estimated as contrasts of the looking times on the test trials. Thus, for each infant, the estimators themselves are uncorrelated. For example, for an infant habituated on the 180 screen with no block, let the four test trial scores be denoted X_{180}, X_{120B}, X_{180B}, and X_{120}. Then, the estimators for the infant's level CS and CB values are obtained by $(X_{180} + X_{120B} + X_{180B} + X_{120})/4$, $(X_{120} + X_{120B} - X_{180} - X_{180B})/4$, and $(X_{120B} + X_{180B} - X_{180} - X_{120})/4$. To see that these three estimators are orthogonal, it is sufficient to note that the coefficients of X_{180}, X_{120B}, X_{180B}, and X_{120} for the three estimators are ¼, ¼, ¼, ¼ for the first, –¼, ¼, –¼, ¼ for the second, and –¼, ¼, ¼, – for the third. Pairwise orthogonality is guaranteed by the fact that the sum of the products of the coefficients is zero for any two of the three sets of coefficients. Consequently, any correlation of the parameter estimates over infants cannot be due to the form of the parameter estimators themselves but must be due to real variation in the parameters resulting from their relation to some psychological variable.

Regression analysis revealed that the total looking time on the four test trials was significantly correlated with the CS estimate ($R = .714$, $F = 35.38$, $p = .00$) and was also significantly correlated with the CB estimate ($R = .368$, $F = 5.33$, $p = .03$). Showing that the screen and block parameters significantly covary with total looking time reveals that variation in those parameter estimates is due to something more than chance variation and justifies sorting the infants on the basis of their parameter estimates. The CS and CB estimates were uncorrelated ($R = .105$, $F = .38$, $p = .54$).

Discussion

These results strongly support the conclusion that the impossibility–possibility variable is not at work in the screen rotation experiment. The results also suggest that previous findings of novelty and familiarity preference varying from infant to infant need to be taken into account in analysis of looking-time data. Treating the individual preferences as free parameters estimated from the data is appropriate. Bogartz and Shinskey (1998, p. 155) provided the reasoning, based on analyses of the criterion method by Bogartz (1965), indicating why preference variation is to be expected generally when a habituation criterion is used. Still, it seemed desirable that confirmation of such preference variation be obtained by experimental manipulation. Our approach in Experiment 2 was to try to bias some of the infants toward familiarity screen preference by familiarizing them for less trials and to bias the other infants toward novelty screen preference by familiarizing them for more trials, with the expectation that results for those infants will be different on the test trials.

EXPERIMENT 2

Our approach in this experiment was to familiarize approximately one half of the infants for only three trials and the remainder for seven trials. Approximately one half of each of these groups were familiarized on the 120 event and the remainder on the 180 event. Figure 4 shows the model for these two screen familiarization conditions. We expected that whether the infants were familiarized on the 180 or the 120 event, three familiarization trials would be more likely to yield negative values of CS and seven familiarization trials would be more likely to yield positive values. It was not so clear what to expect in the case of CB. On one line of thought, no training with the block stimulus is the same whether there are three such trials or seven in that if the block stimulus has never been presented, it will in either case be just as novel when presented on the test trials. Alternatively, it may be that with more familiarization trials, the existing representation to which the test trial event is being compared may be stronger in some sense, in which case the effect of the presentation of the block on the 180 + B and 120 + B test trials may have a greater effect on looking time.

Familiarization (Between)		Test Trials (Within)			
Trials	Rotation	180	120+B	180+B	120
3	180	L-CB-CS	L+CB+CS	L+CB-CS	L-CB+CS
3	120	L-CB+CS	L+CB-CS	L+CB+CS	L-CB-CS
7	180	L-CB-CS	L+CB+CS	L+CB-CS	L-CB+CS
7	120	L-CB+CS	L+CB-CS	L+CB+CS	L-CB-CS

FIGURE 4 The model for Experiment 2, ignoring the expected variation in CS depending on how many familiarization trials are given. *Note.* L = the general level of looking; CS = the effect of a change in the screen rotation condition from 180 on habituation to 120 on test or vice versa; –CS = the effect of no change; CB = the effect of a change in the block condition from presence of block during habituation to absence of block during test or vice versa; –CB = the effect of no change.

Method

Although we regard the methods employed in Experiments 1 and 2 as equivalent for our purposes, we present them in detail because Experiment 1 was conducted in Bogartz and Shinskey's lab and Experiment 2 was conducted in Schilling's lab.

Participants. Fifty-two infants aged 5½ months plus or minus 14 days participated in the study. All infants were reported by parents to have been born full-term, with a birthweight greater than 2,500 g and at the time to be in good health. Nine additional infants were dropped because of fussiness or failure to complete all of the testing trials. One of these was due to a camera malfunction. All infants were tested at the Infant Studies Laboratory at Franklin and Marshall College in Lancaster, Pennsylvania. Names of parents and infants were obtained from public birth announcements listed in a local newspaper. The parents of potential participants were mailed a letter that explained the general nature of the procedure. Parents were asked to mail an enclosed postcard indicating a willingness to become members of a pool of potential participants. During a lab visit, parents signed a consent form that explained the nature of the procedure and advised the parents of their right to terminate their infants' participation at any time without any questions asked. After a testing session, all parents were debriefed as to the purpose of the study, given an opportunity to ask questions, and the infants were given a T-shirt and certificate for their participation.

Stimuli and apparatus. The moving screen was a white, foam core screen, 20.3 cm long × 18.4 cm high, situated on a black box 38.1 cm long × 10.8 cm high × 22.9 cm wide. The block (18 cm long × 10 cm high × 1 cm thick) had black and white checkerboard stripes and was large enough so that the screen would hit it if the screen progressed any further than 120°. In addition, movement of the block was controlled by an experimenter who used a lever to raise and lower the block. To raise the block, the experimenter pulled a string that lifted the block from the rear base of the apparatus. To lower the block, the experimenter released the string and the block then rested flush with the apparatus. This experimenter was able to raise and lower the block while positioned behind the rear center panel (from the infant's point of view) of a four-sided screen that sat on the same table as the apparatus. There was a small hole in the lower right corner of the rear center screen through which an experimenter could see the screen movement. Attached to either side of the center panel were two blinders that prevented infants from seeing the experimenters and the rest of the lab room. In front of the infants was a curtain that could be opened and closed by the experimenter behind the rear center screen. While the curtain was closed, infants could no longer see the apparatus.

The movement of the screen was controlled by a Cybermation SCK–2000 controller kit. The controller card was mounted underneath the black base of the moving screen. The moving screen was run by a 12-V AC Center Tap, 4 amp transformer that powered a 12-V stepper motor. Software for an IBM computer allowed the parameters of speed and movement and degrees of arc of the screen to be set prior to the start of a trial.

The length of trials and looking data were recorded by a computer. The software cued the experimenter when to begin a trial, when to end a trial, and when a specific number of trials had been reached. All testing sessions were videotaped by a Sony Camcorder video camera that recorded a direct view of the infants' gazes to the screen. All data were copied from the 8 mm camcorder tape onto VHS VCR tapes with a timer image superimposed on every frame in .01 sec.

Procedure. The infants were seated in a car seat during the entire procedure. The parents were seated next to the infants to the right of the right blinder. This arrangement allowed parents and infants to see each other without permitting parents to view the apparatus. Before familiarization trials started, parents were reminded not to interact with the infants while the curtain was open. After a trial ended and the curtains were closed, parents were permitted to gently pat or talk to a fussy infant during the intertrial period.

Three experimenters participated in the procedure. In an adjacent control room, Experimenter 1 scored looking time from a video monitor and Experimenter 2 initiated and terminated the screen movement at the beginning and end of a trial. Experimenter 3 opened and closed the curtains at the beginning and end of a trial. Experimenter 3 was cued when to open and close the curtains by Experimenter 1 via headphones. Finally, Experimenter 3 raised and lowered the block attached to the apparatus. This block served as the obstacle in the path of the moving screen during the test trials. A trial started with a single button press on the keyboard by

Experimenter 2, which in turn activated the moving screen to begin its arc. Once a trial began, Experimenter 1 began scoring by hitting a key on a keyboard that was connected to the Macintosh computer.

Infants were randomly placed into a group that received either three 180° (n = 11), three 120° (n = 14), seven 180° (n = 14), or seven 120° (n = 13) familiarization trials. Infants in all groups received four test trials: 180, 120, 180 + B, and 120 + B. The four test trials were presented in random order determined individually before the start of the procedure. A trial began with a .1-sec fixation to the screen and continued until the infants accumulated 5 consecutive sec of looking time at the screen. If an infant looked away for 2 consecutive sec after the accumulation criterion was met, then the trial ended. Otherwise, a trial ended after the infant accumulated 60 sec of looking time at the screen. During all trials, the screen moved through its arc (i.e., 180 or 120, block or no block) at a rate of 45° per sec.

Scoring. All trials from all participants were scored several times. Initially, all infants were scored live during the testing session. Online scoring was necessary because the computer controlled trial length based on accumulated looking time. One experimenter watched a video monitor that recorded the infants' gazes and the apparatus. This experimenter (who was not blind to condition) could hit keys on a keyboard that recorded the looking time of each gaze. All participants' data were also scored from videotapes. Two independent scorers used a VCR that has a shuttle control to score the videotapes. This shuttle enabled the scorer to stop the tape at precisely the point at which a look began and to advance the tape to the point at which the look had terminated. Each scorer was blind to the type of trial being presented. This was possible because for each rotation, only the first 45° of the screen movement as the screen went through the first part of its rise and last 45° of the screen movement as the screen returned to its initial position were videotaped. Because the event distinguishing the trial type, 120 versus 180 and block versus no block, occurred after the screen had risen 45° and before it descended the final 45°, the scorer could not know what type of trial was occurring.

The scorers began each trial to make sure the infants had accumulated 5 consecutive sec of looking time at the apparatus at the start of each trial. Each scorer marked the time that an infant began a directed gaze at the apparatus, then advanced the tape, and marked the time when the infants stopped looking at the apparatus. The scorers then advanced the tape through the nonlooking interval to make sure that 2 consecutive sec of nonlooking had not elapsed. This scoring procedure continued for each looking and nonlooking period on every trial for each infant.

If a scorer found that an infant had not accumulated 5 consecutive sec at the start of a trial, the infant's data were dropped. Also, an infant's data were eliminated if a scorer found that an infant had actually looked away from the apparatus for more than 2 consecutive sec and the experimenter failed to terminate the trial. The data from the two scorers were compared using Pearson's correlation coefficient.

Results

Interobserver reliability. Ten of the infants were selected at random. Interobserver reliability was assessed by determining the correlation between the judgments of total looking time on each training trial and each test trial by two judges. The first observer was Thomas S. Schilling and the second observer was an undergraduate assistant who was blind to the hypotheses of the experiment and to the trial type. The correlation was .998, the slope of the regression line was 1.009, and the intercept was –0.041, indicating that the two observers were judging the looking times virtually identically.

Participant parameter tests. The primary purpose of this experiment was to use the model predictions to show that the preference effects could be experimentally affected and that the model parameters would properly reflect this variation. To do this, an estimate of CS and CB was obtained for each individual participant. Let LT be the looking-time function such that, for example, LT(120 + B) is the looking time to test trial 120 + B. Then, Figure 4 reveals that for infants familiarized to the 180 event, CS can be estimated by [LT(120 + B) + LT(120) – LT(180 + B) – LT(180)]/4, and CB can be estimated by [LT(180 + B) + LT(120 + B) – LT(180) – LT(120)]/4. Similarly, for infants familiarized to the 120 event, CS can be estimated by [LT(180 + B) + LT(180) – LT (120 + B) – LT(120)]/4, and CB can be estimated by [LT(180 + B) + LT(120 + B) – LT(180) – LT(120)]/4. In each case, negative values of the parameter reflect a preference for the familiar event and positive values reflect a preference for the novel event.

The estimates of CS and of CB were analyzed separately using in each case a one-factor between-participant design with the factor having the four familiarization conditions 3–180, 3–120, 7–180, and 7–120 as its four levels. For each parameter, three planned orthogonal contrasts were performed. The first contrast tested the difference between the mean parameter values at the two levels of number of familiarization trials, three versus seven. The second contrast tested the difference between the means for 3–180 and 3–120, and the third contrast tested the difference between the means for 7–180 and 7–120.

For the CS variable, the means for 3–180, 3–120, 7–180, and 7–120 were –3.44, –4.27, .61, and 3.83. These means were significantly different, $F(3, 48) = 10.45$, $p = .00$. The difference between the mean CS values for infants receiving three familiarization trials versus infants receiving seven familiarization trials was significant, $F(1, 48) = 26.99$, $p = .00$. The difference between the mean CS values for 3–180 and 3–120 was not significant, $F(1, 48) = .62$ and the difference between the mean CS values for 7–180 and 7–120 just missed being significant, $F(1, 48) = 3.97$, $p = .052$. The results clearly indicate that we successfully experimentally manipulated the CS parameter value by manipulating the number of familiarization trials. This result strongly supports the argument that looking time on the test trials is determined by familiarity or novelty preferences for the screen rotation event resulting from the relation between the representation of the familiarization event and the perception of the test event.

For the CB variable, the means for 3–180, 3–120, 7–180, and 7–120 were –1.96, –.54, .38, and 1.12. These means were not significantly different, $F(3, 48) = .38$ and none of the contrasts were significant. The overall mean value of CB was –.176, suggesting that on average, familiarization without the block stimulus produced neither a familiarity nor a novelty preference with respect to the block stimulus. However, the mean square for error for the analysis was 54.61, suggesting quite a bit of variability in the CB values. We are inclined to interpret this as suggesting that for some infants the addition of the novel block stimulus produced greater looking, in that those infants were further along in processing the rotating screen component, but that other infants were still more involved in processing the screen rotation, resulting in less attention to the block stimulus. Some support for this conjecture comes from the fact that although the four group means for the CB estimates did not differ from each other, they were nevertheless correlated with the four group means for the CS estimates, $r = .85$, suggesting that the more the rotation event was processed the more the block event became of interest.

Model fitting. Following the model fitting approach for Experiment 1, the appropriate row of Figure 4 was used together with participant effects to model the test trial data for each of the four groups. The models fit well in all cases. For groups 3–180, 3–120, 7–180, and 7–120, respectively, the proportion of variance accounted for by the models was .797, .612, .706, and .561.

Graphs of the test trial observed and predicted values for the four groups are shown in Figure 5. Inspection of these graphs and the results of a 2 × 2 × 2 × 2 mixed model ANOVA in which the between-participant factors were number of familiarization trials (three vs. seven) and amount of screen rotation (180 vs. 120) and the within-participants factors were amount of screen rotation (180 vs. 120) and block (present vs. absent) reveals several interesting aspects of the data. When the infants are familiarized for three trials, they look longer at the familiar screen rotation than at the novel one regardless of which screen rotation they were familiarized to, $F(1, 48) = 14.44$, $p = .04$. With seven familarization trials, this screen familiarity preference vanishes and the infants instead look longer at the novel screen, $F(1, 48) = 5.68$, $p = .02$. When familiarization was seven trials to the 120 rotation event, the infants look longer to the novel 180 rotation events, $F(1, 48) = 8.15$, $p = .01$. When familiarization was seven trials to the 180 rotation event, there is only a nonsignificant, slightly longer looking to the novel 120 events, $F(1, 48) = .22$. One interpretation of this lack of a strong novelty preference is that the 120 rotation is included in the 180 rotation and so is not so novel to the infant as is the 180 rotation following familiarization to the 120 rotation because the 180 rotation is not included in the 120 rotation.

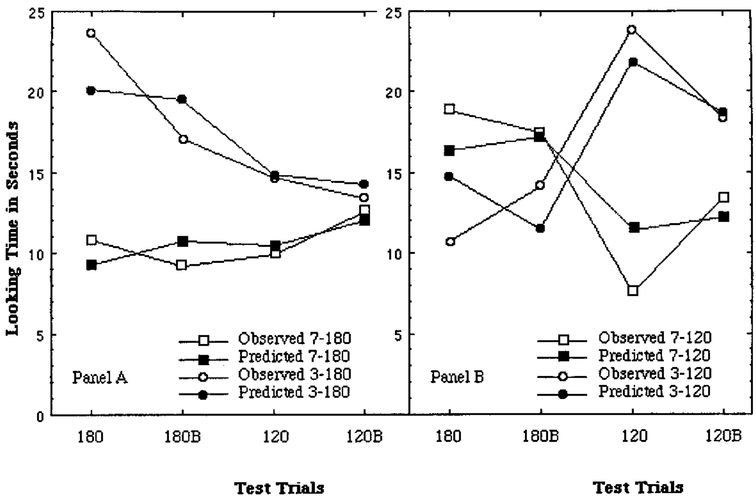

FIGURE 5 Observed and predicted looking times on the four test trials following familiarization for three or seven trials on either the 180 event (Panel A) or the 120 event (Panel B).

Recall that Baillargeon et al. (1985) found that following habituation to the 180 rotation event, infants looked longer to the 180 + B event than to the 120 + B event and that this was also found in our Experiment 1 for the infants who were inferred to have screen familiarity preference. For the infants in our Experiment 1 who were inferred to have screen novelty preference, the reverse finding was obtained. Inspection of Panel A of Figure 5 reveals results that support those inferences. Infants familiarized on rotation event 180 for only three trials look longer at the 180 + B impossible event than at the 120 + B possible event, but infants familiarized on rotation event 180 for seven trials look longer at the 120 + B possible event than at the 180 + B impossible event. This result clearly shows that experimental variation of the familiarity of the 180 screen rotation event results in concomitant variation in which of the two critical test rotation events is looked at longer. This result also coincides nicely with the results of Cohen and Cashon (1998) and the results of Schilling (this issue).

Discussion

Bogartz et al. (1997) and Bogartz and Shinskey (1998) found that infants did not look longer to the impossible event than to the possible event during habituation in their Event Set × Event Set treatments of the Baillargeon and Graber (1987) rabbit study and the Kellman and Spelke (1983) split stick experiment, respectively. This result has been obtained again in our Experiment 1 Event Set × Event Set treatment of the Baillargeon et al. (1985) rotating screen experiment. As these results accrue, it becomes increasingly difficult to accept that infants know that the events are impossible and look longer on the test trials because of it, but show no signs of slowed habituation to the impossible event. If the impossible event is surprising, why is it not surprising when it is first seen? Surprise has always been an inference from longer looking as far as young infants are concerned. There has never been an independent measure of surprise, such as facial expressions, to support that inference. To the extent that the absence of longer looking during habituation undermines the surprise interpretation, it tends to undermine claims based on longer looking during the test trials. The notion that the infants look longer on the test trials because they are surprised becomes less tenable, and therefore, so do the claims that the infants have object permanence and the knowledge of solidity.

The fact that the novelty–familiarity comparison model is fitting well in Experiments 1 and 2 supports our general approach and suggests that great care is needed in interpreting studies that have habituated to only one stimulus event due to the possibility of confounding of effects. Claims that novelty explanations have been circumvented should be viewed with caution because this often is obtained at the price of confounding a possible familiarity preference with the experimental variable of interest. In this context, this observation is exemplified by the fact that Baillargeon et al. (1985) assumed longer looking at the impossible screen motion could not have been due to a novelty effect, because in their study the infants were habituated to the 180° screen motion, which on the impossible test trial was familiar rather than novel. Our Experiment 2 shows experimentally that longer looking to the familiar rather than to the novel is possible. A similar concern arises for Experiment 1 of the Kellman and Spelke (1983) study of infants looking at either a solid stick or a split stick following their looking at a stick with its middle occluded. They assumed that if infants saw the occluded stick as a solid stick during habituation, then longer looking at the split stick would be expected because it would be the novel event. However, if infants preferred to look at the familiar event instead of the novel one, then longer looking at the split stick on the test trials may well be due to a familiarity preference (see Bogartz & Shinskey, 1998). It is this ambiguity of interpretation that the Event Set × Event Set design is intended to help avoid.

As Bogartz and Shinskey (1998) indicated, the Event Set × Event Set design entails a certain kind of symmetry that is psychologically informative. The models we have used have predicted that there should be no difference in test trial looking time between being habituated to Event A and being tested on Event B versus being habituated to Event B and being tested on Event A. In this context, this symmetry can be seen in Figure 2B. For example, compare the predicted value in row 2, column 1 with the predicted value in row 1, column 2. The prediction is that looking time should be the same whether habituated to the 120 + B event and tested on the 180 event or vice versa. This holds throughout the matrix of predicted values. We have found support for this symmetry repeatedly in the past (Bogartz & Shinskey 1998; Bogartz et al., 1997) and have found it in this study (see Table 1). We believe that this symmetry is suggestive of the very type of comparison process that we have hypothesized: namely, that the looking time is based on a perceptually based comparison of the test event with a stored representation of the habituation or familiarization event. We know of no interpretation based on higher level cognitive functions such as reasoning or logical inference that predicts this symmetry.

This study shows again in both experiments the usefulness of recognizing that familiarity or novelty preference can vary from individual to individual, as did the study by Bogartz and Shinskey (1998). Experiment 2 in our study goes beyond previous work in that it demonstrates by experimental manipulation that familiarity preference and novelty preference are real phenomena in this type of experiment and should be considered at the level of the individual.

The finding that the experimentally manipulated familiarity preference and novelty preference for screen rotation produces the expected variation in the CS parameter of the model provides a new level of support for both the novelty–familiarity comparison approach to understanding young infant looking times and for the Event Set × Event Set design approach.

A remark on the use of familiarization rather than habituation in Experiment 2 is in order. We wanted to manipulate amount of exposure to the events in the first stage of the experiment to show that the parameter estimates were sensitive, in the predicted way, to manipulations that varied the novelty of the test trial events. Two options existed. Either we could try to use two different fixed numbers of exposures for the two groups, thereby manipulating number of exposure trials, or we could try to find a second habituation criterion to use. We had no way of knowing with certainty which procedure would be better. It is obvious that the variability in the underlying processes, over a group of infants, can be greater using habituation or can be greater using a fixed number of familiarization trials, depending on the initial distributions of starting values and of rate parameters governing the processes. In the absence of information concerning the correct model and the distribution of initial values, choosing between the two methods amounts to guesswork. Our guess was that the familiarization procedure would do a better job of producing, on average, a difference in the parameter estimates, with less variability. The guess was bolstered by the fact that about two thirds of the infants in Experiment 1 reached criterion in the minimum number of trials possible (six). It seemed that seven trials should do a good job of familiarizing most of the infants and using less than half that many trials should leave the infants with information still unprocessed. The success of the experiment appears to have supported our guess.

This study shows once again that the previously used two-test design that familiarizes on one event and tests on two others is inferior to the Event Set × Event Set design. The two-test design tends to confound familiarity and novelty preferences with the effects of interest, such as possibility–impossibility, whereas the Event Set × Event Set design tends to clarify the determiners of infant looking time. Furthermore, the Event Set × Event Set design provides direct statistical tests that can decide between the different hypotheses at issue and it also provides standard statistical methods for parameter estimation so that the contribution of the different variables to total looking time can be directly estimated.

The results obtained by Cohen and Cashon (1998) are consistent with our findings. They found no support for greater looking to the impossible event. Rather, with their well-habituated older infants, they found a novelty looking preference that is exactly what we would expect, given our results in Experiment 2 and the findings by others that it is older infants with more familiarization trials that show the greatest preference for novel events (Hunter & Ames, 1988; Hunter et al., 1982).

Schilling's (this issue) study was particularly persuasive in that he replicated the Baillargeon et al. (1985) results but also showed that by familiarizing the infants on the screen that moved through the same arc as the possible event, the infants look longer at the possible event than at the impossible event, providing a strong case for a familiarity preference effect for the screen rather than representation of the occluded object.

Meltzoff and Moore (1998), distinguishing between object identity and object permanence, suggested that "younger infants are limited to anticipating appearances based on the place or trajectory of the object when it was last visible," as opposed to older infants who they assume "can represent an invisible object as being in an invisible place or on an invisible trajectory behind an occluder" (p. 212). Meltzoff and Moore attempted to account for the Baillargeon et al. (1985) Experiment 1 result as follows:

> In the violation condition the box is first seen stationary in a place. It was occluded as the screen rotated up, and was absent when the screen lay flat on the table. Over multiple trials, there were repeated disappearances and reappearances. Infants would be expected to set up a representation of the box in place, especially after repeated exposures. If this representation persists over short intervals, infants would expect to see the same box, identified by its place, whenever the place is visible. When the screen is rotated down revealing no box in place, there is a mismatch between perception and representation. This discrepancy yields longer looking. Detecting the discrepancy between the pre- and post-disappearance scenes requires a representation of the past, but object permanence is not necessary. (p. 212)

We find the approach of Meltzoff and Moore (1998) compatible with our general approach. They emphasize, as we do, that young infants represent what they have seen and compare what they are seeing with their representations of what they have seen. Meltzoff and Moore's explanation of the longer looking to the 180° rotation with the block may be partially correct but, given the results of our Experiment 2, it cannot tell the whole story. According to their explanation, the screen rotation during the familiarization stage plays no part in why the infants look longer at one event than at the other, nor does the number of familiarization trials. We have shown that both of these variables have strong effects on looking time to the various events. A. N. Meltzoff (personal communication, September 1, 1998), responding to the previous statement, remarked,

> We don't think that the number of familiarization trials and the screen rotation "play no role." It is a little more accurate to say that our explanation cannot be the whole story. As formulated, Meltzoff and Moore's view does not anticipate that the screen rotation and number of familiarization trials have such a large effect on looking time.

Rivera, Wakeley, and Langer (1999) also suggested that the screen rotation during familiarization is irrelevant to longer looking at the impossible event. They showed that infants would look longer at the impossible event even when no habituation to the 180° screen rotation precedes the test trials. In a second experiment, they showed that longer looking occurred to a 180° rotation than to a 112° rotation even when no block was involved. They concluded that simple perceptual preference for motion accounts for the results and that their findings question Baillargeon's (1987) claim that longer looking to the impossible event is due to infant representational reasoning about physically impossible object permanence events. The results of our Experiment 2 give no support to the hypothesis of a looking preference for the 180° rotation. Rather, it is the type of rotation and the number of trials that determines which screen rotation is looked at longer on the test trials. However, our results are not necessarily in conflict with those of Rivera et al. In our experiments, we have shown a strong effect due to what happens on the familiarization trials. In their experiments, Rivera et al. showed that with no familiarization trials, infants look more at the screen that rotates more. It is plausible that in the absence of any familiarized event, the processes involved in comparison are less involved in looking and so perceptual aspects of the events themselves play the major role in determining looking.

Important new theoretical work (Bogartz & Shinskey, 1998; Leslie, Xu, Tremoulet, & Scholl, 1998; Meltzoff & Moore, 1998) is helping to draw more clearly the distinctions that will be necessary in our attempts to understand the nature of early infant knowledge and the determiners of infant looking and reaching in situations involving occlusion. Some of the classic experiments (Baillargeon & Graber, 1987; Baillargeon et al., 1985; Kellman & Spelke, 1983) that have been widely assumed to demonstrate young infant representation of the occluded object have been recently receiving experimental attention and consideration of alternative interpretations (Bogartz & Shinskey, 1998; Bogartz et al., 1997; Cohen &

Cashon, 1998; Meltzoff & Moore, 1998; Schilling, this issue). This new work effectively calls into question the tenability of the assumption that young infants represent occluded objects in occlusion. The theoretical work by Leslie et al. (1998) assumed such representation and may have to be altered to accommodate these new results.

A body of work that to date remains as a bulwark in support of the assumption of young infant representation of occluded objects is the work on infant arithmetic (Simon, 1998; Simon, Hespos, & Rochat, 1995; Wynn, 1992, 1998). It is interesting to note that Leslie et al. (1998) assumed such representation and included a treatment of the arithmetic work from their perspective, but Meltzoff and Moore (1998) assumed no such representation in young infants and did not refer to the infant arithmetic work. From our standpoint, any approach that assumes that young infants do not represent occluded objects in occlusion will have to explain the infant arithmetic results. We are currently working on this problem.

ACKNOWLEDGMENT

Jeanne L. Shinskey is now at the Department of Psychology, University of Denver.

REFERENCES

Baillargeon, R. (1987). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology, 23,* 655–664.

Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C. E. Granrud (Ed.), *Carnegie-Mellon Symposia on Cognition: Vol. 23. Visual perception and cognition in infancy* (pp. 265–315). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baillargeon, R., & Graber, M. (1987). Where is the rabbit? 5.5-month-old infants' representation of the height of a hidden object. *Cognitive Development, 2,* 375–392.

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in 5-month-old infants. *Cognition, 20,* 191–208.

Bogartz, R. S. (1965). The criterion method: Some analyses and remarks. *Psychological Bulletin, 64,* 1–14.

Bogartz, R. S., & Shinskey, J. L. (1998). On perception of a partially occluded object in 6-month-olds. *Cognitive Development, 13,* 141–163.

Bogartz, R. S., Shinskey, J. L., & Speaker, C. (1997). Interpreting infant looking: The Event Set × Event Set design. *Developmental Psychology, 33,* 408–422.

Cohen, L. B., & Cashon, C. (1998). Eight-month-old infants' perception of possible and impossible events. *Infant Behavior and Development, 21,* 348.

Fischer, K. W., & Bidell, T. (1991). Constraining nativist inferences about cognitive capacities. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 199–235). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Haith, M. M. (1988). 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 multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 5, pp. 69–95). Norwood, NJ: Ablex.

Hunter, M., Ross, M., & Ames, E. (1982). Preferences for familiar and novel toys: Effects of familiarization times in one-year-olds. *Developmental Psychology, 18,* 519–529.

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

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences, 2,* 10–18.

Meltzoff, A. N., & Moore, M. K. (1998). Object representation, identity, and the paradox of early permanence: Steps toward a new framework. *Infant Behavior and Development, 21,* 201–235.

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104,* 686–713.

Piaget, J. (1954). *The construction of reality in the child.* New York: Basic Books.

Rivera, S. M., Wakeley, A., & Langer, J. (1999). The drawbridge phenomenon: Representational reasoning or perceptual preference? *Developmental Psychology, 35,* 427–435.

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.

Shinskey, J. L. (1999). *Why do young infants fail to search for hidden objects?* Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Simon, T. J. (1998, April). *A fully operationalized, demonstrably sufficient, non-numerical account of infant "numerical competence."* Paper presented at the Eleventh International Conference on Infant Studies, Atlanta, GA.

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10,* 253–269.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99,* 605–632.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 6389,* 358.

Wynn, K. (1998, April). *Arguments against non-numerical accounts of infants' numerical competence.* Paper presented at the Eleventh International Conference on Infant Studies, Atlanta, GA.