Cognition 175 (2018) 1–10

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Five-month-old infants have expectations for the accumulation of nonsolid substances

Erin M. Anderson*, Susan J. Hespos, Lance J. Rips

Psychology Department, Northwestern University, United States

ARTICLE INFO

Keywords:
Cognitive development
Infant cognition
Substance concepts
Representation of nonsolids

ABSTRACT

Infants fail to represent quantities of non-cohesive substances in paradigms where they succeed with solid objects. Some investigators have interpreted these results as evidence that infants do not yet have representations for substances. More recent research, however, shows that 5-month-old infants expect objects and substances to behave and interact in different ways. In the present experiments, we test whether infants have expectations for substances when the outcomes are not simply the opposite of those for objects. In Experiment 1, we find that 5-month-old infants expect that when a cup of sand pours behind a screen, it will accumulate in just one pile rather than two. Similarly, infants expect that when two cups of sand pour in separate streams, two distinct piles will accumulate rather than one. Infants look significantly longer at outcomes with an inconsistent number of piles, providing evidence that infants have expectations for how sand accumulates. To test whether the number of cups or the number of pours guided expectations about accumulation, Experiment 2 placed these cues in conflict. This resulted in chance performance, suggesting that, for infants to build expectations about these outcomes, they need both cues (cup and pour) to converge. These findings offer insight into the nature of infants' representations for non-cohesive substances like sand.

1. Introduction: Knowledge of substances and its source

Our theories of how infants conceive entities in their environment have changed dramatically. Early theories suggested that infants perceive a sensory flux, aware of features like color and shape but not of individual objects (James, 1890). Subsequent evidence countered this view, showing that infants—even in the first weeks of life—have sophisticated knowledge about how objects behave and interact (Spelke, Breinlinger, Macomber, & Jacobson, 1992). Two-month-old infants have expectations about the naïve physics of occlusion and containment, and they look significantly longer at events that violate these expectations than at events that conform to them (Aguiar & Baillargeon, 1999; Baillargeon et al., 2012; Hespos & Baillargeon, 2001). These studies provide evidence that infants possess core principles about the solidity, continuity, and persistence of objects.

Although studies of early object concepts demonstrated that infants perceive objects in much the way adults do, the world is not made of objects alone. Nonsolid substances like water, milk, and soil are intrinsic parts of human experience, and our interactions with these substances differ markedly from our interactions with objects. We expect to be able to push a toy car across the floor but not a puddle of water. Our reactions in these situations reflect our awareness that objects and substances have distinct physical properties and so behave in distinct but predictable ways. However, current evidence about infants' knowledge of substances is less detailed than evidence about their knowledge of objects. Do infants have substance-specific expectations? Or, do beliefs about objects arise from a privileged domain of knowledge, while our beliefs about substances derive from how they differ from objects?

Current evidence is unclear on whether infants have principled expectations for non-cohesive substances, with early results suggesting gaps in their knowledge (Cheries, Mitroff, Wynn, & Scholl, 2008; Chiang & Wynn, 2000; Huntley-Fenner, Carey, & Solimando, 2002; Rosenberg & Carey, 2009). Huntley-Fenner et al. (2002) showed 8-month-old infants a pile of sand, then concealed the pile behind a screen, and, poured a second pile of sand behind a nearby but spatially-separated screen. In this situation, adults would expect to see two piles of sand if the screens were taken away. However, when the two screens were removed, the infants spent no more time looking at a display containing just one pile than they did at a display containing two. But with similar-looking (but solid) sand-pile shaped objects, infants performed as expected in this paradigm, looking longer at the one-object display than at the two-object display. Together, these findings suggest that infants may have so little knowledge of (or so little ability to

* Corresponding author at: Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, IL 60208, United States.
E-mail address: erinanderson2014@northwestern.edu (E.M. Anderson).

https://doi.org/10.1016/j.cognition.2018.02.009
Received 29 March 2016; Received in revised form 5 February 2018; Accepted 8 February 2018

process) substances that they are unable to predict a substance's continued presence when it is briefly out of sight. Spelke and Kinzler (2007) took this interpretation a step further: They proposed that infants have principled expectations for objects but not for substances.

Infants' difficulties with non-cohesive substances extend to collections of small solid objects. In a paradigm similar to Huntley-Fenner et al. (2002), Chiang and Wynn (2000) showed 8-month-old infants a noncohesive pile of Lego pieces or a cohesive Lego pyramid. In the noncohesive case, after concealing a first pile behind a screen, the experimenter brought out a second pile of Legos and placed it behind a separate screen. When the screens were removed, only one pile of Legos appeared. Looking time at this magical disappearance was compared to a second condition consisting of an expected disappearance from two piles on the stage. In this second condition, after screens concealed both piles, infants then saw one pile removed from behind a screen. When the screens were removed, only one pile remained. The infants spent no more time looking at the magical disappearance than at the expected disappearance when noncohesive piles of Legos were used. However, when the Legos were pressed together to form cohesive pyramids and were lowered behind the screens, infants performed as expected, looking significantly longer at the $2–0 = 1$ display than at the $2–1 = 1$ display. Chiang and Wynn (2000) concluded that infants track cohesive objects as discrete individuals but cannot track non-cohesive entities.

In contrast to these findings, more recent studies suggest that basic reasoning about non-cohesive substances appears in the first months of life, provided infants are tested on principles appropriate to them. Hespos, Ferry, and Rips (2009) found that infants are sensitive to motion cues for a liquid and that these cues then guide their expectations about how the liquid will behave. In the habituation trials, infants viewed a clear cup containing a liquid whose surface deformed and shifted as liquids ordinarily do when the cup was tilted and rotated. This cue led them to look longer when the cup was upturned and the contents tumbled out (as if solid) compared with when the cup was upturned and the contents poured out (as if liquid). In a second experiment, infants who saw the same motion cues in the tilting cup expected that the contents were permeable—for example, that a solid cylinder would pass through the top surface and not remain on top. It is possible that infants' expectations are specific to liquids, emerging from their early experience drinking and bathing. Yet new research reveals that expectations about the non-cohesive qualities of liquids generalize to other substances (Hespos, Ferry, Anderson, Hollenbeck, & Rips, 2016). When infants viewed a clear cup containing sand that was tilted and rotated, the motion cues caused them to expect that the contents would pass through a grid. In addition, when the contents of the clear cup were tiny glass balls that tilted back and forth in the cup, infants expected that a solid object would pass through the top surface and not remain on top. In each of these experiments, an object condition showed that the converse was true, too: Infants looked longer at substance behaviors when they expected an object and at object behaviors when they expected a substance. Together, these findings provide evidence that principles for non-cohesive substances emerge around the same time as object principles in development and apply to unfamiliar as well as to familiar materials.

The evidence from these paradigms successfully counters the claim that infants cannot represent non-cohesive substances. However, these studies leave unclear the source of infants' substance knowledge. On the one hand, infants may gain this knowledge directly from the substances themselves, mastering principles that are specific to substances (see Rips & Hespos, 2015, for a discussion of these principles). On the other hand, their substance knowledge may arise in a derivative way from simple contrasts with object knowledge. According to the latter alternative, they may identify substances as non-objects and predict the substances' behavior as the opposite of objects'. If this is true, expectations for substances might still be built on the expectations for objects that appear in the first three months of life (Spelke et al., 1992). Past studies of substances like Hespos et al. (2009) cannot resolve this issue because test trials in these experiments always contrasted object and substance outcomes: If a solid stops another solid at its surface, a substance should allow the solid to go through. If a solid maintains its shape, a substance should deform. If a solid cannot pass through a small grid, a substance should.

In the current study, we wanted to test infant expectations for non-cohesive substances when the probable outcome was not simply the opposite of an object outcome. To do this, we looked at pouring sand into piles. Non-cohesive substances like sand can merge into a single pile or divide into separate piles, depending, in principled ways, on the situation. For example, consider sand poured from a cup onto a table. Pouring the sand can create a specific number of piles, depending on the location of the pours. Two pours in distinct locations normally produce two piles, whereas two pours in the same location normally produce one. Solid objects, on the other hand, do not merge or divide under the same conditions. Dropping solid objects does not typically change the number of objects in a way that depends lawfully on location (e.g., dropping two apples in one vs. two locations does not change the number of apples). For this reason, an infant who knows only (a) the behavior of solid objects under these transformations and (b) that nonsolids behave in ways opposite that of solids, would not be able to predict how the number of piles of nonsolids varies with the location of pours. If, however, they are able to reason about the sand's behavior independent of object rules, then they should expect that the number of piles will match the number of pours into distinct locations.

2. Experiment 1: Do infants know that the number of pours determines the number of piles?

In the current study, we test infants' expectations about the naïve physics of sand. If infants see a single cup of sand poured behind a screen, would it violate their expectations to reveal two distinct piles?

2.1. Method

We tested whether 5-month-old infants have principled beliefs about the non-cohesive nature of sand during pouring events. Infants were randomly assigned to either the single-pour or double-pour condition. Fig. 1 illustrates these events. In the single-pour habituation condition, infants saw a single cup filled with sand poured onto a tray. A screen hid the portion of the tray where the sand accumulated. In the double-pour habituation condition, infants saw simultaneous pours from two cups of sand emptied in separate streams onto a tray. Again, a screen hid the portion of the tray where the sand accumulated. After each pour, the tray was removed from the stage and emptied. The pouring/emptying cycle was repeated continuously until the trial ended. After habituation trials were over, the infants saw an alternation between two types of test trials. The only difference between habituation and test events was that the screen was removed after each pour, revealing sand piles on the tray. When the screen was removed, on half the trials it revealed a single pile of sand, and on alternate test trials, two separate piles. If infants have expectations about sand that go beyond a mere contrast with object rules, then infants in the single-pour condition should look longer at the two-pile test trials compared to the one-pile test trials. For infants in the double-pour condition, we would expect the opposite pattern of results.

2.1.1. Participants

The participants were 34 healthy, full-term infants (16 female, 18 male), ranging in age from 4 months 15 days to 6 months 16 days ($M = 5$ months 12 days). Half the infants were assigned to the single-pour condition ($M = 5$ months 13 days, 9 female) and the other half were assigned to the double-pour condition ($M = 5$ months 10 days, 7 female). Eleven additional infants were tested but eliminated from the final analysis: one because of fussiness (defined as more than four test trials in which the infant was coded as crying or fussy by two

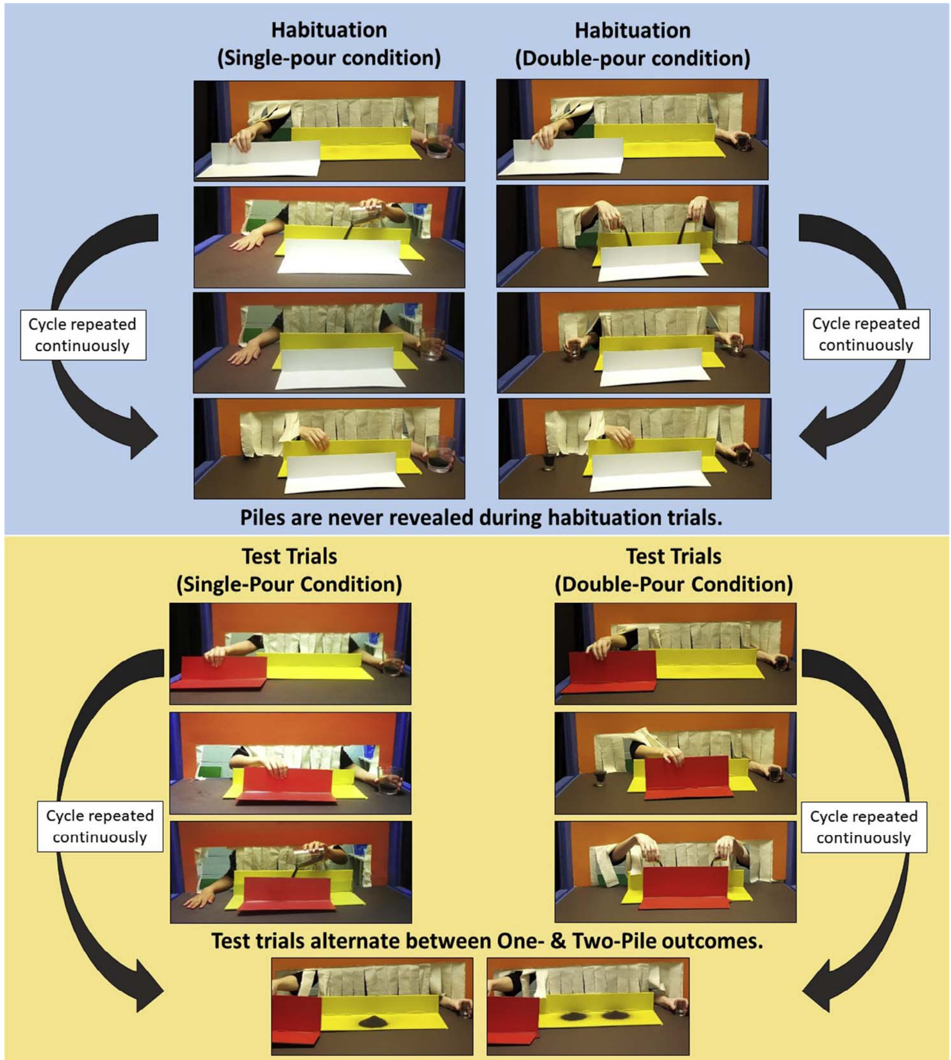

Fig. 1. Illustration of the habituation and test trials in Experiment 1. Infants were randomly assigned to either the single-pour or double-pour condition. Test trials alternated between one-pile and two-pile outcomes regardless of condition.

independent coders), one for falling asleep (defined as more than four test trials in which the infant was coded as sleepy by two independent coders), six because they had bowel movements during the experiment, and three because they looked the maximum amount on five or more of the six test trials. Before beginning data collection, we calculated that we would need a total sample size of at least 32 infants based on the goal of obtaining a power of 0.80 with an alpha level of $p < .05$, and on the effect sizes (0.3–0.5) we found in previous studies on substance knowledge using this looking-time paradigm (Faul, Erdfelder, Lang, & Buchner, 2007). After data collection and exclusions were complete, we had 17 infants in each condition.

Parents provided informed consent (IRB protocol #STU00010996) before the experiment and were compensated $20. The race/ethnicity of participants in this and the experiment that follows was 81% non-Hispanic, 16% Hispanic, and 3% who chose not to answer. The racial makeup was 75% white, 3% Asian, 3% African American, and 15% multiracial. The remaining 4% chose not to answer. The mothers included 93% whose highest level of education was a college degree or higher, 3% with some college, 2% with a high school diploma, and 2% who chose not to answer. The fathers included 89% whose highest level of education was a college degree or higher, 6% with some college, 1% with high school only, 1% who did not complete high school, and 3% who chose not to answer.

2.1.2. *Apparatus*

Parents sat in a chair with the infant on their lap facing a wooden puppet stage that displayed all stimulus objects 100 cm in front of the infant. The stage measured 243.5 cm high, 128 cm wide, and 61 cm deep. The opening in the front of the stage that displayed the objects was 93 cm above the floor, 61 cm high, and 106 cm wide. The back wall had one rectangular opening with cloth fringe over the opening that allowed the experimenter to manipulate the stimuli on stage. A curtain that covered the infants' view of the stage was raised and lowered between trials (see Fig. 1).

The substance was black sand (Scenic Sand, Activa Product, Marshal, TX, USA). Stimuli in the single-pour condition consisted of a 355 ml clear plastic cup on a stage, filled with 130 g of sand, which was raised and then poured behind an occluding screen. In the double-pour condition, two clear 42 ml shot glasses – each filled with 65 g of sand – were raised and poured approximately 15 cm from each other. In the two-pile outcome, the closest borders of the piles were typically 3 to 5 cm apart. The sand landed on a yellow tray 50.8 by 18.5 cm. The occluding screen used during habituation trials was L-shaped cardboard 12.7 cm high, 33 cm wide and 10.2 cm deep that was covered in white contact paper. The occluding screen used in test trials had red contact paper and similar dimensions. The single-pour to two-pile trick was achieved by hiding a 90-degree metal divider behind the test screen to divert the sand into separate piles in the single-pour condition for the two-pile outcome. The double-pour to single pile trick was achieved by hiding a wide funnel (19.8 cm at its mouth) behind the test screen that funneled the sand into the center of the tray.

2.1.3. Events

Infants viewed a series of 6 to 9 habituation trials immediately followed by 6 test trials alternating between the one- and two-pile outcomes. Trial order (probable or improbable test event first) was counterbalanced across participants.

2.1.3.1. Single-pour condition. For the single-pour condition, when the curtain was raised at the start of a habituation trial, one clear cup filled with sand was visible on the stage beside an empty yellow tray (see Fig. 1). The experimenter's right hand moved the white occluding screen in front of the tray before returning the hand to rest on the floor of the stage (2 s). This occluding screen never moved during habituation following its initial placement in front of the tray while the following 16-to-18-s cycle repeated continuously until the trial ended. The left hand, already grasping the cup, lifted the cup, moved it above the tray, tilted it to pour sand behind the screen (4 s), and then returned the cup to rest on the stage (2 s). The left hand then removed the empty cup from the stage and returned with a filled cup (4 s). Finally, the right hand removed the tray to empty it backstage and then returned it, all with the screen still in place (6–8 s). (During emptying, the tray was angled and pulled back behind the stage, so the infant could not see the resulting piles on the tray.)

In test trials, the experimenter's right hand moved the red screen in front of the tray before returning the hand to rest on the floor of the stage (2 s). Then the left hand, already grasping the cup, raised it, moved it above the tray, tilted it to pour sand behind the screen (4 s), and then returned the cup to rest on the stage (2 s). The right hand then removed the screen to reveal one or two piles of sand (2 s). The left hand then removed the empty cup from the stage and returned with a full one (4 s). Finally, the right hand removed the tray to empty it backstage and then returned it empty (6–8 s). This 20-to-22-s cycle was repeated continuously until the trial ended.

2.1.3.2. Double-pour condition. For infants in the double-pour condition, the two cups were poured in unison behind the screen. The only difference in the timing was that the sequence was 4 s longer, due to the experimenter replacing the second empty cup.

2.1.4. Procedure

During the experiment, infants sat on their parent's lap in front of the apparatus. Parents were asked to refrain from interacting with the infants during the experiment, and to close their eyes during the test trials. Each trial began as the experimenter raised the screen and said "Go" which signaled the coders to begin tracking infant looks. Each trial ended when the infant either looked away for 2 consecutive seconds after having looked at the event for at least 2 s or looked at the event for 60 cumulative seconds without looking away for 2 consecutive seconds. A computer determined the end of the trial and beeped to signal the experimenter to lower the screen. The habituation criterion was a decline of at least 50% in total looking duration from the first three to the last three habituation trials (or a maximum of nine trials). The average number of trials to reach criterion was seven.

Each infant viewed six test trials alternating between the one-pile and two-pile outcomes. The type of test event shown first was counterbalanced across infants. The criteria for ending a test trial were the same as those for the habituation trials. If a test trial ended before the number of piles was revealed for the first time, we reran that test trial type at the end of the experiment. (This happened on 4% of the test trials.) The infants we tested saw an average of two "reveals" per test trial.

To control for potential experimenter bias, five research assistants and two authors alternated as the presenting experimenter. An analysis of the effect of experimenter on looking time between test events (when piles matched the number of pours vs. when they did not) shows no interaction between these, $F(6, 27) < 1$, $p = .97$.

2.1.5. Coding

There was a small hole in the front face of the stage containing a camera that captured a video image of the infant's face. Two research assistants in a separate room viewed this image and coded infants' visual fixations online watching separate monitors and using separate keyboards. Each researcher depressed a computer button when the infant attended to the events on stage and released the button when the infant looked away (Chang & Wang, 2014).

These live coders were blind to the condition in which the infant was being run and to the outcomes. If agreement between coders was at or above 90%, data was entered based on the looking times indicated by Coder 1. If agreement was lower than 90%, trials were recoded off-line from the videotape. As with presenting experimenters, the role of each coder was randomly assigned between several research assistants. Interobserver agreement for looking durations of all infants averaged 92%. We also compared the total looking time on test trials, as recorded by coder 1, with the total looking time as recorded by coder 2. The intraclass correlation coefficient was 0.996 with a 95% confidence interval from 0.991 to 0.998, $F(1, 29) = 235.926$, $p < .001$.

2.2. Results

We calculated the mean looking time for the single-pour and the double-pour conditions depending on habituation condition. Fig. 2 presents these mean looking times for the habituation and test trials. Infants looked longer at test events when the number of piles revealed behind the screen conflicted with the number of pours. In the single-pour condition, infants looked an average of 35.70 s ($SD = 13.02$) for the two-pile outcome and an average of 26.68 s ($SD = 10.86$) for the one-pile outcome. In the double-pour condition, infants looked an average of 29.17 s ($SD = 12.82$) for the one-pile outcome and an average of 22.96 s ($SD = 13.32$) for the two-pile outcome. Across conditions, 24 of 34 infants looked longer at the test trials where the number of pours conflicted with the number of piles than at the test trials where the number of pours and piles were consistent ($p = .024$, binomial comparison). Our data in Experiment 1 significantly deviates from a normal distribution per the Shapiro-Wilks test. Therefore, we performed the following parametric tests on log-transformed data, following recommendations outlined by Csibra, Hernik, Maxcaro, and Tatone (2016).

A 2 × 2 × 2 × 2 repeated measures analysis of variance (ANOVA) with the between-subject factors of habituation condition (single pour vs. double pour), order (conflicting outcome first vs. second in test pairs), sex (male vs. female) and the within-subject factor of test event outcome (one-pile vs. two-pile) revealed a significant interaction between condition and outcome, $F(1, 26) = 8.35$, $p = .008$, $\eta_p^2 = 0.243$ confirming that infants looked longer at the test events where pours and piles conflicted. There was also a three-way interaction with habituation condition (single vs. double pour), sex, and test event outcome (one-pile vs. two-pile), $F(1, 30) = 4.407$, $p = .046$, $\eta_p^2 = 0.145$. That is, male infants looked longer at the conflicting number of piles for both habituation conditions, whereas female infants showed this pattern for the single-pour habituation condition but not for the double-pour habituation condition. There was no main effect of outcome ($F(1, 26) < 1$, $p = .762$) nor condition ($F(1, 26) = 2.038$, $p = .165$).

To determine whether the looking patterns were consistent throughout the test period, a 2 × 2 × 3 repeated measures ANOVA examined whether this looking pattern between outcomes (one pile vs. two) and condition (one pour vs. two) differed with the test trial pair (1st, 2nd, or 3rd). The analysis revealed an effect of test trial pair, as looking decreased over the course of test, $F(2, 23) = 5.636$, $p = .010$, $\eta_p^2 = 0.329$. Additionally, as in the omnibus, there was a significant

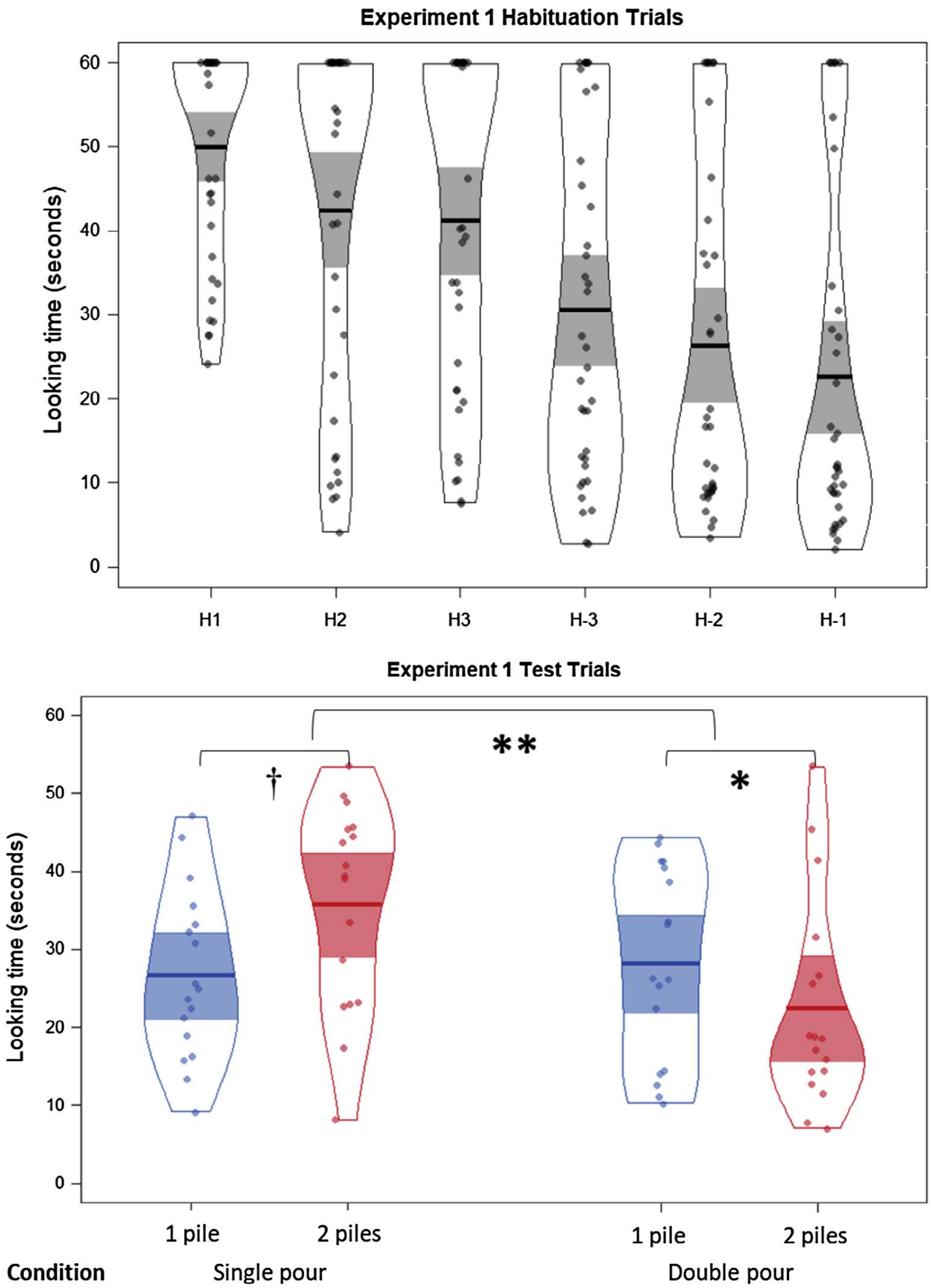

Fig. 2. Bean plots showing mean looking times during the habituation and test trials. For the habituation trials, we collapsed across the single- and double-pour conditions. The graphs include results for the first three trials (H1, H2, H3) and the last three trials before the habituation criterion was met (H-3, H-2, H-1). The central line in each box is the mean, and the upper and lower shaded portions represent the 95% Confidence Intervals (CIs) for this mean (i.e., there is a 95% probability that the true population mean falls within this interval). Dots indicate the raw data points. The width of the bean indicates the density of the data distribution at a looking time value. For the test trials, we separated the data by habituation to the single- and double-pour conditions. The * indicates a p-value less than 0.05, and the † indicates a p-value less than 0.10. The ** indicates an interaction between condition and outcome, with p-value less than 0.01.

interaction between outcome and condition, $F(1, 24) = 11.66$, $p = .002$, $\eta_p^2 = 0.327$. However, there was no interaction between trial pair, outcome, and condition, $F(2, 23) = 1.189$, $p = .323$, suggesting that infants discriminated between events across test trials.

Further analyses revealed that this pattern of looking longer at the conflicting outcome was evident within each condition. In the double-pour condition, infants looked significantly longer at the one-pile outcome, $F(1, 16) = 6.079$, $p = .025$, $\eta_p^2 = 0.275$. In the single-pour condition, looking at the two-pile outcome was on the brink of significance, $F(1, 16) = 4.279$, $p = .055$, $\eta_p^2 = 0.211$.

Next, we compared the infants' looking durations during habituation trials between the conditions. There was no difference in the number of trials to meet the habituation criterion for a single versus double pour ($M = 7.53$; $M = 7.39$ respectively, range = 6 to 9, $F(1, 32) < 1$, $p = .90$), nor in the amount of looking time that infants accumulated during habituation ($M = 284.68$ s and $M = 276.20$ s, respectively, $F(1, 32) < 1$, $p = .805$), nor in the decline of looking across habituation trials, $F(5, 28) < 1$, $p = .819$. Similarly, at test, infants in Experiment 1 showed equal looking at one-pile and two-pile outcomes when we collapsed across conditions, $M_{\text{one-pile}} = 27.43$ s ($SD = 11.49$) and $M_{\text{two-piles}} = 29.07$ ($SD = 14.65$). These findings suggest that the infants did not have a bias to look longer at the one-pile over the two-pile outcome or vice versa.

2.3. Discussion

Our findings show that infants look longer at events in which a single-pour of sand produces two sand piles and a double-pour produces a single pile. This demonstrates that infants have expectations about sand that are not simply the opposite of those for object outcomes. This result is unlikely to be a low-level perceptual bias, because infants reacted in opposite ways to the same test outcomes (i.e., one pile or two piles) depending on whether they saw a single- or double-pour.

We know from earlier studies that infants are aware of situations in which solid objects and nonsolid substances display different behavior. For example, five-month-old infants understand that substances, unlike objects, are permeable (Hespos et al., 2009) and non-cohesive (Hespos et al., 2016). However, the result of Experiment 1 is unique because the expectations infants are displaying for sand's accumulation here cannot be based on contrasts with object rules.

It is possible, though, that in challenging or unfamiliar situations, infants might default to using object rules rather than using the contrasts derived from those rules. In this case, they might expect a one-to-one mapping between pouring events and piles, just as dropping objects creates a one–one mapping between the dropping events and the dropped objects. In addition, the pouring event for Experiment 1 provided two visual cues that infants may have used to form correct expectations about the number of piles. One of these was the number of pours, but the other was the number of cups. Because these cues covaried – each pour always came from a separate cup – it is impossible to distinguish the independent contribution of these two cues in Experiment 1. This raises the question of whether infants are inferring the number of piles from the number of cups, from the number of pours, or from both. In Experiment 2, we separate these cues by testing infants' expectations for one or two piles of sand when they see either two consecutive pours of sand from one cup or one merged pour from two cups of sand.

If infants' expectations for number of piles are based on the number of pours, then infants in a two-pour, one-cup condition should expect two piles, while infants in a one-pour, two-cup condition should expect one pile. If, however, their expectations are based on the number of cups, then infants in a two-pour, one-cup condition should expect one pile and infants in a one-pour, two-cup condition should expect two piles. If infants are relying on both cups and number of pours to direct their expectations, then looking times at these events are likely to return mixed results since the cues are in conflict. Finally, if infants are defaulting to object rules, then they should expect to find two piles in both conditions: Even if two objects were dropped into one area, two objects should still result.

3. Experiment 2: Do the number of pours or the number of cups determine the number of piles?

In Experiment 2, we tested whether 5-month-old infants anticipate the resulting number of sand piles based on the number of pours or cups of sand. Infants were randomly assigned to either a two-pour, one-cup condition or a one-pour, two-cup condition. Fig. 3 illustrates these events.

3.1. Method

In the two-pour, one-cup habituation condition, infants saw a clear cup pour half its contents onto the left side of the tray, then pour the remaining sand onto the right side of the tray. In the one-pour, two-cup habituation condition, infants saw two cups of sand pour simultaneously in one merged stream of sand onto a tray. Again, a white screen hid the portion of the tray where the sand accumulated. As in Experiment 1, the tray was removed from the stage and emptied after each pour. Similarly, the pouring/emptying cycle was repeated continuously until the trial ended. After habituation trials were over, the infants saw an alternation between two types of test trials. Like Experiment 1, the only differences between habituation and test events was that a red screen was used instead of a white one, and that the screen was moved to the side after each pouring event, revealing sand piles on the tray. On half the trials, one pile of sand was revealed, and on the alternating trials, two piles of sand were revealed.

3.1.1. Participants

The participants were 32 healthy, full-term infants (19 female, 13 male), ranging in age from 4 months 1 day to 6 months 28 days (M = 5 months 11 days). Half the infants were assigned to the two-pour, one-cup condition (M = 5 months 8 days, 9 female) and the other half were assigned to the one-pour, two-cup condition (M = 5 months 14 days, 10 female). Nine additional infants were tested but eliminated from the final analysis: three because of fussiness (same criterion as Experiment 1), one because they had a bowel movement during the experiment, one because they looked the maximum amount on five or more of the six trials, and four because they were inattentive (defined as the infant looking away before the first reveal on three or more test trials).

3.1.2. Events

Infants viewed a series of 6 to 9 habituation trials immediately followed by 6 test trials alternating between the one- and two-pile outcomes. Trial order was counterbalanced across participants.

3.1.2.1. Two-pour, one-cup condition. For the two-pour, one-cup condition, when the curtain was raised at the start of a habituation trial, a single sand-filled cup was to the left of the empty cardboard tray. The experimenter's right hand moved the white screen in front of the tray before returning the hand to the stage floor (2 s). Then the following 19–22-s cycle was repeated until the habituation trial ended. The left hand, already grasping the cup, lifted it, moved it above the left side of the tray, tilted it to pour half the sand behind the screen (4 s), moved the cup above the right side of the tray, tilted it to pour the remainder of the sand behind the screen (3–4 s) and then returned the cups to rest on the stage (2 s). The left hand then removed the empty cup from the stage and returned with a full cup (4 s). Finally, the right hand removed the tray to empty it backstage and then returned it, all with the screen still in place (6–8 s).

In test trials, all timing was the same, except that the experimenter removed the screen before refilling the cup (2 s) so that the number of accumulated piles was visible to the infant. The 21–24-s test was repeated continuously until the trials ended.

3.1.2.2. One-pour, two-cup condition. For the one pour, two-cup condition, when the curtain was raised at the start of a habituation trial, two sand-filled cups were on either side of an empty cardboard tray. The experimenter's right hand moved the white screen in front of the tray before returning the hand to grip on the right cup (2 s). Then the following 20–22-s cycle was repeated until the habituation trial ended. Each hand lifted a cup, moved them together above the center of the tray, tilted them to pour sand behind the screen in a single stream (4 s), and then returned the cups to rest on the stage (2 s). The left hand then removed one empty cup from the stage and returned with a full cup (4 s), then the right hand replaced the other cup with a full one (4 s). Finally, the right hand removed the tray to empty it backstage and then returned it, all with the screen still in place (6–8 s). The occluding screen never moved during habituation after its initial placement in front of the tray at the start of the trial. During emptying, the tray was angled and pulled back behind the stage, so the infant could not see the resulting piles on the tray.

In test trials, all timing was the same, except that the experimenter removed the screen before refilling the cup (2 s) so that the number of accumulated piles was visible to the infant. This 24–26-s test cycle was repeated continuously until the trials ended.

3.1.3. Procedure, apparatus, and coding

The procedure, apparatus and coding were the same as in Experiment 1. If a test trial ended before the accumulated pile was revealed, we reran that test trial type at the end of the experiment. (This happened on 8% of the test trials). Inter-observer agreement for looking durations of all infants averaged 92%. We also compared the total

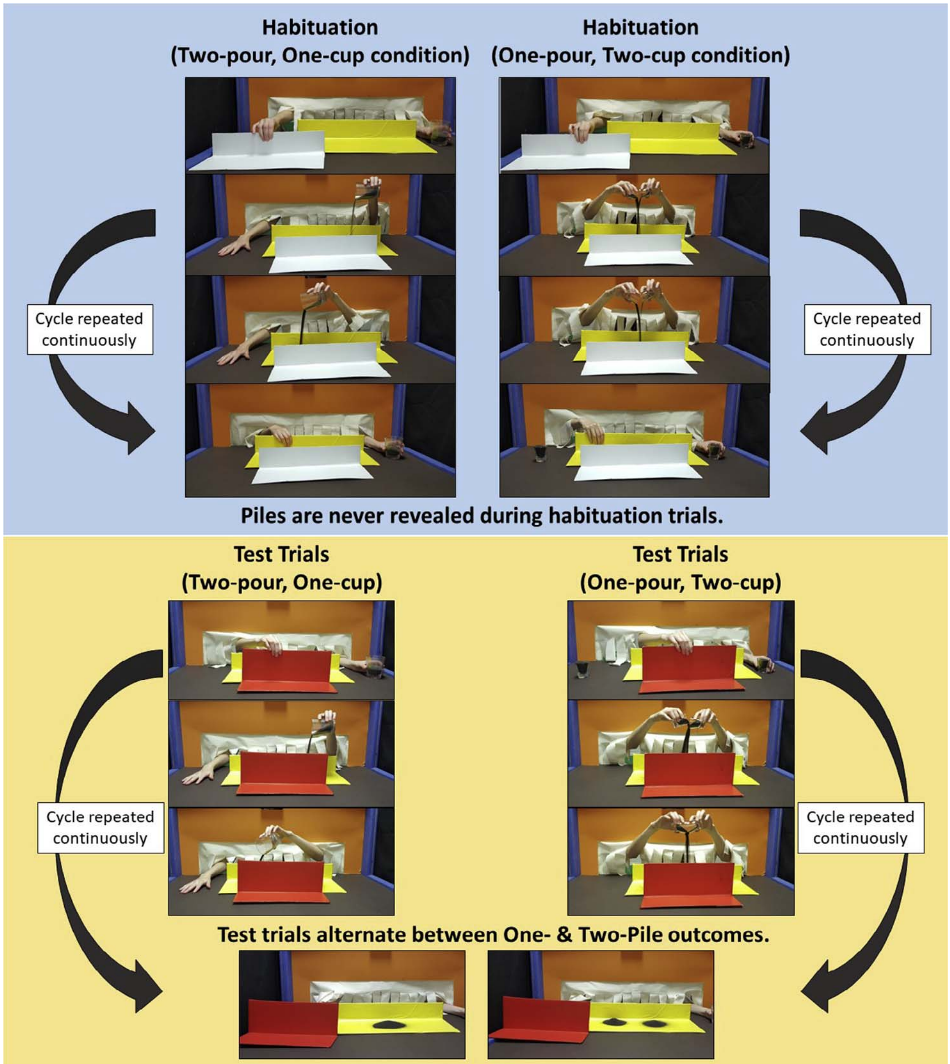

Fig. 3. Illustration of the habituation and test trials in Experiment 2. Infants were randomly assigned to either the two-pour, one-cup or one-pour, two-cup condition. Test trials alternated between one-pile and two-pile outcomes regardless of condition.

looking time on test trials, as recorded by coder 1, with the total time as recorded by coder 2. The intra-class correlation coefficient was 0.998 with a 95% confidence interval from 0.993 to 0.998, $F(1, 31) = 284.74$, $p < .001$.

To control for potential experimenter bias, 5 research assistants and 2 authors alternated as the presenting experimenter. An analysis of the effect of experimenter on looking time between test events that matched pours vs. cups shows no interaction between these, $F(6, 25) = 0.957$, $p = .473$.

3.2. Results

We calculated the mean looking times for habituation and test trials, separately for each condition, and Fig. 4 presents these means. Infants in Experiment 2 did not show that same looking patterns as in Experiment 1. Looking times were split evenly: Of the 32 infants, 15 looked longer when the number of piles matched the number of pours, and 17 looked longer when the number of piles matched the number of cups, $p = .86$. In the two-pour, one-cup condition, the average looking time for the one-pile outcome was 27.97 s ($SD = 11.53$) and the two-pile outcome was 28.60 s ($SD = 9.76$). In the one-pour, two-cup condition, the average looking time for the two-pile outcome was 24.90 s ($SD = 12.47$) and the one-pile outcome was 20.68 s ($SD = 6.75$). Our data in Experiment 2 did not significantly deviate from the normal distribution per the Shapiro-Wilks test. For consistency with Experiment 1, and following recommendations by Csibra et al. (2016), we performed the parametric statistics on the log transformed data.

3.2.1. Within experiment analysis

A 2 × 2 × 2 × 2 repeated measures analysis of variance (ANOVA) with the between-subject factors of habituation condition (two-pours, one-cup vs. one-pour, two-cups), test order (one- or two-pile outcomes first), and sex, and the within-subject factor of test outcome (one-pile outcome vs. two-pile outcome) revealed a significant three-way interaction of condition, outcome, and order, $F(1, 24) = 5.122$, $p = .033$, $\eta_p^2 = 0.176$. Infants on average looked slightly longer when the outcomes matched the number of cups compared to the number of pours, and this was more pronounced when the first test trial shown during test had the outcome matching number of pours. Critically, though, there was no interaction between the habituation condition and the outcome, $F(1, 24) < 1$, $p = .709$, and neither the one-pour, two-cup condition nor the two-pour, one-cup condition resulted in significant looking differences between the number of piles, $F(1, 15) < 1$, $p = .479$, and $F(1, 15) < 1$, $p = .596$, respectively. A 2 × 2 × 3 repeated measures ANOVA also tested whether the looking pattern between outcome (one pile vs. two) and condition (1 pour vs. 1 cup) depended on the test trial pair (1st, 2nd, or 3rd). The analysis did not yield an effect of test trial pair, $F(2, 28) = 1.708$, $p = .20$, nor an interaction between test trial and outcome

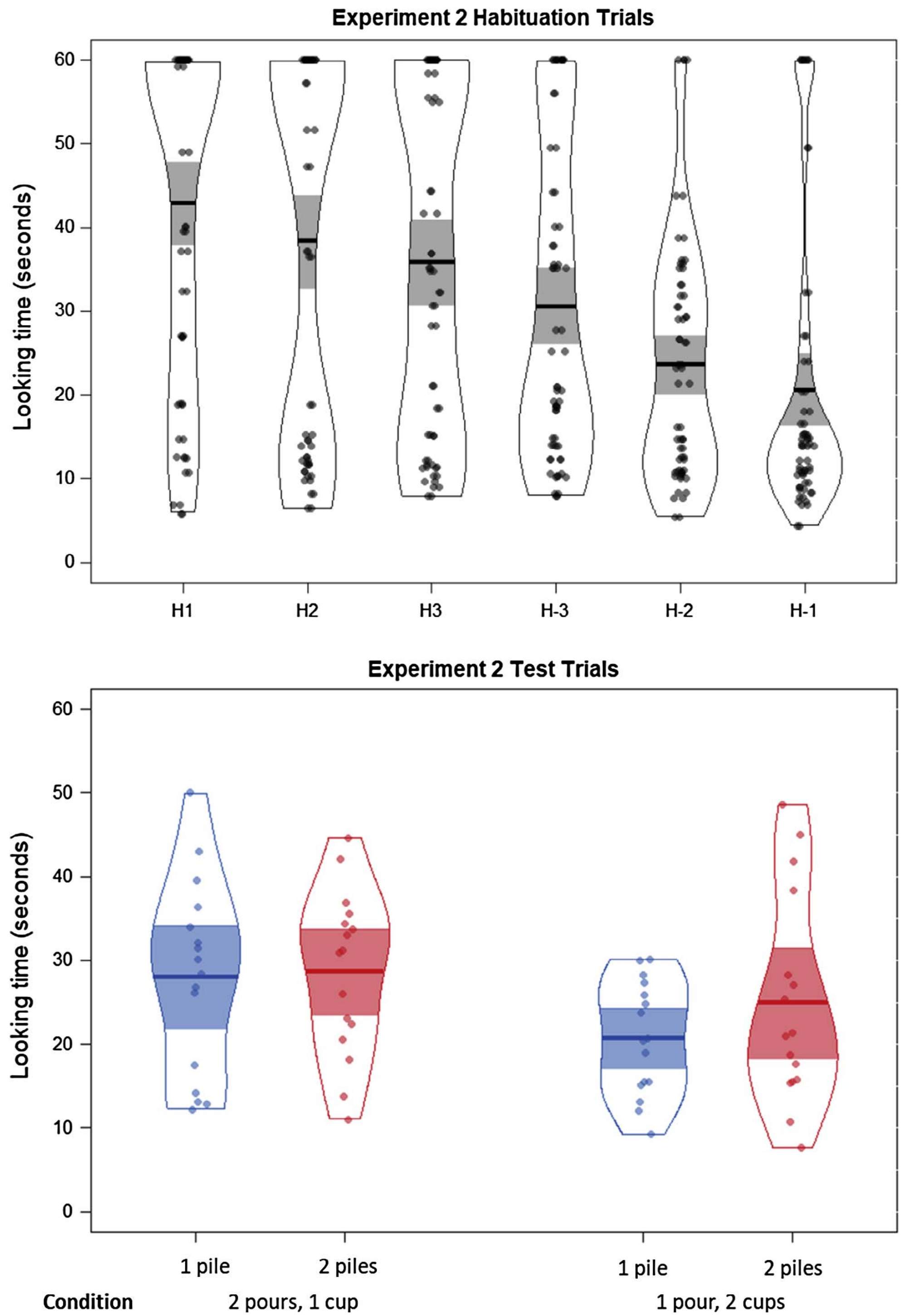

Fig. 4. Bean plots showing mean looking times during the habituation and test trials. For the habituation trials, we collapsed across the two habituation conditions. The graphs include results for the first three trials (H1, H2, H3) and the last three trials before the habituation criterion was met (H-3, H-2, H-1). The central line in each box is the mean, and the upper and lower portions of each box represent the 95% Confidence Intervals (CIs) for this mean. Dots indicate the raw data points. The width of the bean indicates the density of the data distribution at a looking time value. For the test trials, we separate the data by habituation condition.

and condition, $F(2, 28) < 1$, $p = .941$.

As in Experiment 1, we also compared the habituation patterns of the infants who were habituated to one-pour, two-cups to those habituated to two-pours, one-cup. Habituation condition did not predict the number of trials to meet the habituation criterion ($M = 7.94$; $M = 7.88$ respectively, range = 6 to 9, $F(1, 30) < 1$, $p = .89$), nor the amount of looking time that infants accumulated during habituation ($M = 253.18$ s and $M = 255.32$ s, respectively, $F(1, 30) < 1$, $p = .551$) nor the decline of looking across habituation, $F(5, 26) < 1$, $p = .865$.

3.2.2. Cross-experiment analyses

Next, we ran a series of analyses looking at the combined data from Experiments 1 and 2. A 2 × 2 × 2 repeated measures ANOVA found an interaction between experiment (1 vs. 2), outcome (1 vs. 2 piles), and condition (1 vs. 2 cups), $F(1, 62) = 5.93$, $p = .018$, $\eta_p^2 = 0.087$. Infants in the two-pour, one-cup condition showed equivalent looking between the number of piles, but in the other conditions across Experiments 1 and 2, infants differentiated these outcomes. Across experiments, the interaction between outcome and condition was trending but not significant, $F(1, 62) = 3.56$, $p = .064$. There was also an effect of the number of cups across the experiments, such that, on average, infants looked for shorter periods of time at two-cup events, regardless of the number of piles that resulted, $F(1, 62) = 6.036$, $p = .017$, $\eta_p{}^2 = 0.089$.

Despite the interaction in looking at test outcomes, there was no effect of habituation decline by experiment, $F(5, 58) < 1$, $p = .456$, nor an interaction between habituation decline, experiment, and condition, $F(5, 58) < 1$, $p = .606$. Additionally, participation in Experiment 1 or 2 did not predict the number of trials needed to reach habituation ($F(1, 60) < 1$, $p = .34$) nor the total amount of looking acquired during habituation ($F(1, 60) < 1$, $p = .335$).

3.3. Discussion

In Experiment 2, we contrasted the number of pours with the number of cups to determine whether infants were using cups or pours to guide their expectations for sand's accumulation. We found that infants did not distinguish between one- and two-pile outcomes under these conditions, suggesting that, for infants to form expectations at five months of age, having both cues point towards the same outcome is important.

We interpret these results as evidence that 5-month-old infants have an early concept of how sand accumulates, but that their expectations are clearest when multiple perceptual cues converge: that is, if a single cup produces a single pour, then a single pile is highly probable. In contrast, when two cups produce a single pour or one cup produces two consecutive pours, infants may have conflicting expectations based on each of these cues. The three-way interaction between condition, experiment, and outcome suggested that the one-pour, two-cups condition yielded more similar results to Experiment 1 than the two-pours, one-cup condition. The direction of the effect in the former condition suggests that infants' predictions are beginning to lean more heavily on pours than on cups. If so, this may signal the emergence of the adult perspective in which the number of (locations of) pours rather than the number of cups is the more reliable cue to the number of piles. Future experiments should examine whether working memory or other variables factored into this pattern.

The possibility that multiple cues working in unison lead to more accurate expectations corresponds with a common pattern in depth perception (Granrud, 1993). Similarly, Baillargeon and colleagues have documented that, despite impressive early reasoning, infants may not form the correct expectations for events like containment and occlusion when multiple variables are in conflict (Luo & Baillargeon, 2005; Wang, Baillargeon, & Paterson, 2005; see Baillargeon et al., 2012, for a review).

4. General discussion

In these experiments, we tested infants' expectations for how many piles would result when sand was poured in one stream or in two pours at separate locations. These experiments are unique in testing infants' expectations for substances when the expectations could not be based on contrasts with object behavior. Specifically, because solid objects do not normally fall apart and then reform when dropped, reasoning about accumulation into one or two piles during pouring events cannot be based on principles for objects. The data from Experiment 1 demonstrate that 5-month-old infants encoded the pouring of sand behind a screen, and they anticipated that a single-pour from one cup should accumulate in one pile and that a double-pour from two cups should form two piles. However, when the two cues in a pouring event – the number of cups and the number of pours – conflicted in Experiment 2, infants failed to anticipate whether one or two piles should accumulate. If expectations for test outcomes were based on object behavior, then two cups pouring in the same location should still result in two objects or piles. Infants, though, did not discriminate in this condition. Together, our data show that five-month-old infants' expectations about substances are present under favorable circumstances, and are independent of the principles they possess for objects.

These findings are novel because they counter previous studies showing that infants can easily represent small object quantities but struggle when it comes to substances (Chiang & Wynn, 2000; Huntley-Fenner et al., 2002; Rosenberg & Carey 2009). At the same time, Experiment 2 shows that even when infants can form expectations based on substance principles, unrelated cues might divert infant attention from the most relevant variables needed to predict outcomes. Here, the ease of representing objects compared to substances may have made the conflicting cup cue prominent, drawing infants' attention away from the number of pours. However, the results do not point to infants reasoning based on the number of objects instead of on the number of pours. Both, it seems, must align for infants to succeed at this age.

Future studies should investigate where infants' knowledge of substance behavior depends on general learning processes and where it depends on experiences within the domain of substances. One question is how sensitive infants are to other variables that guide substance behavior. Even though both sand and water share properties like permeability and non-cohesion (Hespos et al., 2016), adults would not typically expect water to accumulate in a bounded pile without a container. What is the course of development as infants identify parameters for viscosities (for example, that of oil vs. honey), grain size (for example, that of sand vs. Cheerios), and other physically important features? Future work could capture other cues such as these that inform infant expectations for separation and accumulation.

4.1. Conclusions

Broadly, these findings help clarify the nature of infants' naïve physics outside of solid objects. Previous work has claimed that reasoning about non-objects lags far behind reasoning about objects (Spelke & Kinzler, 2007; Van de Walle, Rubenstein, & Spelke, 1998). Due to the unrestrained nature of substances like sand, expectations about sand's accumulation may indeed be less robust than expectations for objects when visual cues as to the outcome are in conflict. However, the results from Experiment 1 contribute to the growing body of evidence that infants represent substances in principled ways by 5 months of age. Thus, our experiment shows that expectations about how substances behave and interact appear to emerge around the same time as expectations about objects, suggesting that some of these concepts may develop in parallel.

Acknowledgments

This research was supported by a National Science Foundation grant (BCS-0718513) awarded to Susan Hespos and by the U.S. Department of Education, Institute of Education Sciences, Multidisciplinary Program in Education Sciences Fellowship, Grant Award # R305B140042 to Erin Anderson. We thank the members of the Infant Cognition Lab at Northwestern for their help with data collection and coding. In particular, we thank Robert Chapman for his elegant stimuli design, and Alissa Ferry for introducing us to the Pirate Plot package and bringing our box plots into the future.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2018.02.009.

References

Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology, 39*, 116–157.

Baillargeon, R., Stavans, M., Wu, D., Gertner, R., Setoh, P., Kittredge, A. K., & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative

account. *Language Learning and Development, 8*, 4–46.
Chang, Y.-J., & Wang, S.-W. (2014). *BLT: Experimental program for online coding babies' looking time.* IL, USA: Evanston.
Cheries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2008). Cohesion as a constraint on object persistence in infancy. *Developmental Science, 11*, 427–432.
Chiang, W. C., & Wynn, K. (2000). Infants' tracking of objects and collections. *Cognition, 77*, 169–195.
Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology, 52*(4), 521–536.
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
Granrud, C. (1993). *Visual perception and cognition in infancy.* Hillsdale, NJ.: Lawrence Erlbaum.
Hespos, S. J., & Baillargeon, R. (2001). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science, 12*, 141–147.
Hespos, S. J., Ferry, A., Anderson, E., Hollenbeck, E., & Rips, L. J. (2016). Five-month-old infants have general knowledge of how nonsolid substances behave and interact. *Psychological Science*. http://dx.doi.org/10.1177/0956797615617897.
Hespos, S. J., Ferry, A., & Rips, L. J. (2009). Five-month-old infants have different expectations for solids and liquids. *Psychological Science, 20*, 603–611. http://dx.doi.org/10.1111/j.1467-9280.2009.02331.x.
Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't count: perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition, 85*, 203–221.
James, W. (1890). *The principles of psychology.* New York: H. Holt.
Luo, Y., & Baillargeon, R. (2005). When the ordinary seems unexpected: Evidence for incremental physical knowledge in young infants. *Cognition, 95*, 297–328.
Rips, L. J., & Hespos, S. J. (2015). Divisions of the physical world: Concepts of objects and substances. *Psychological Bulletin, 141*, 786–811.
Rosenberg, R. D., & Carey, S. (2009). Infants' representations of material entities. In B. M. Hood, & L. R. Santos (Eds.). *The origins of object knowledge* (pp. 165–188). London, England: Oxford University Press.
Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605–632.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96. http://dx.doi.org/10.1111/j.1467-7687.2007.00569.x.
Van de Walle, G., Rubenstein, J., & Spelke, E. S. (1998). Infant sensitivity to shadow motions. *Cognitive Development, 13*, 387–419.
Wang, S.-H., Baillargeon, R., & Paterson, S. (2005). Detecting continuity violations in infancy: A new account and new evidence from covering and tube events. *Cognition, 95*, 129–173.