Child Development, November/December 2000, Volume 71, Number 6, Pages 1525–1534

Can Young Infants Add and Subtract?

Ann Wakeley, Susan Rivera, and Jonas Langer

Three experiments (N = 68), using Wynn's procedure, tested 5-month-old infants' looking time reactions to correct and incorrect results of simple addition and subtraction transformations. The aim was to investigate both the robustness and the parameters of infants' arithmetic competence. Experiments 1 and 2 (N = 44) were replications of Wynn's first two experiments in which infants were shown addition (1 + 1 = 1 or 2) and subtraction (2 − 1 = 1 or 2) requiring imprecise calculation. Experiment 3 (N = 24) was a subtraction counterpart (3 − 1 = 1 or 2) to Wynn's third experiment requiring precise calculation of addition (1 + 1 = 2 or 3). Unlike Wynn, we found no systematic evidence of either imprecise or precise adding and subtracting in young infants. Our results, together with the mix of both positive and negative findings from other studies of infant arithmetic, suggest that infants' reactions to displays of adding and subtracting are variable and, therefore, that infants' numerical competencies are not robust. This conclusion is consistent with previous findings indicating that simple adding and subtracting develops gradually and continuously throughout infancy and early childhood.

INTRODUCTION

Young infants seem sensitive to the number of elements in perceptual displays, as long as the set sizes are less than four elements. For example, they look longer at novel displays that differ in numerosity from familiar displays (one more or one less). Young infants discriminate between very small numerosities in both static (e.g., Antell & Keating, 1983; Starkey & Cooper, 1980; Strauss & Curtis, 1981) and moving displays (van Loosbroek & Smitsman, 1990) and in sets of items presented simultaneously (e.g., Starkey & Cooper, 1980) or sequentially (Canfield & Smith, 1996; Wynn, 1996). However, even 6- to 8-month-olds still may not discriminate between small numbers of objects when their contour length or their area is controlled (Clearfield & Mix, 1999a, 1999b).

Perhaps the most advanced numerical ability in young infants has been reported by Wynn (1992; see also Simon, Hespos, & Rochat, 1995, and partial replications by Koechlin, Dehaene, & Mehler, 1997, and Moore, 1997; see Langer, Rivera, Schlesinger, & Wakeley, in press, and Wynn, 1998, for reviews). Wynn (1992) tested 5-month-old infants on a series of trials in which they saw a simple arithmetic transformation performed on a hidden display of dolls (1 + 1 or 2 − 1) in two experiments. On alternating trials, the infants were then shown either a correct (i.e., 1 + 1 = 2; 2 − 1 = 1) or an incorrect (i.e., 1 + 1 = 1; 2 − 1 = 2) result of the transformation. Infants in the subtraction condition differed from infants in the addition condition in their preference for looking at two dolls. Infants tended to look longer at two when it was the incorrect result and tended to look longer at one when it was the incorrect result, although the latter finding was not reliable across experiments. Additionally, in a third experiment, infants looked longer at three dolls when shown an addition transformation in which both the correct and incorrect results were in the correct direction (1 + 1 = 2 or 3). Wynn concluded that young infants calculate the exact result, as well as the ordinal direction, of simple arithmetic transformations performed on small sets of items, and that evidence of these abilities in early infancy suggests that they are innate.

However, the addition and subtraction problems presented in Wynn's Experiments 1 and 2 (see also Moore, 1997; Simon et al., 1995; Uller, Carey, Huntley-Fenner, & Klatt, 1999) could be solved by *imprecise* adding and subtracting. They can be solved by calculating the ordinal direction of the addition or subtraction transformation or, as suggested by Wynn (1992), by simply expecting the arithmetic transformation to result in a different number (see also Simon, 1997, 1998). Wynn (1992) concluded that the results of her Experiments 1 and 2 were ". . . consistent with two distinct hypotheses: (1) that infants are able to compute the precise results of simple additions and subtractions and (2) that infants expect an arithmetical operation to result in a numerical change, but have no expectation about either the size or the direction of the change" (p. 749). The only evidence of *precise* calculation comes from Wynn's Experiment 3 (1 + 1 = 2 or 3) in which both the correct and incorrect result are

 0009-3920/2000/7106-0004

in the correct ordinal direction. However, while infants' longer looking at the incorrect result in this experiment could be based on a precise numerical expectation, it could also be due to a simple perceptual preference for looking at more dolls, as found, for example, by Koechlin et al. (1997) and Xu and Carey (1996).

Evidence for young infants' ability to add and subtract is surprising given previous developmental findings on older infants and toddlers. These studies paint quite a different picture of young children's numerical competence. For example, Cooper (1984) found that infants do not even discriminate between very small number ordinal relations (i.e., less-than or greater-than relationships between very small sets) until age 14 to 16 months. Moreover, Houdé (1997), Huttenlocher, Jordan, and Levine (1994), Starkey (1992), and Vilette (1996; Vilette & Mazouz, 1998) found that children do not add and subtract very small numbers of objects correctly until their third or fourth year.

Given this disparity in the reported ontogenetic data, the present set of experiments sought to investigate both the robustness of infants' ability to add and subtract and the parameters of this phenomenon. Accordingly, two experiments were replications of Wynn's (1992) Experiments 1 and 2. A third experiment was a subtraction counterpart to Wynn's addition problem (Wynn, 1992, Experiment 3) in which we showed the infants $3 - 1 = 1$ or 2. This is a further test of infants' ability to calculate precisely the results of simple arithmetic transformations because both the correct and incorrect result are in the expected ordinal direction. In our experiment, however, the incorrect result is a smaller number of dolls, thus eliminating the potential confound of the incorrect result also being the greater number of items (as in Wynn's Experiment 3).

EXPERIMENTS 1 AND 2: IMPRECISE ADDING AND SUBTRACTING

These data were collected in two phases: an initial replication of Wynn's (1992) Experiment 1 ($N = 32$; mean age 5 months, 16 days) and a second attempt to replicate Wynn's Experiment 2 ($N = 12$; mean age 5 months, 17 days). Half of the participants in each experiment were female infants, half were male infants. All procedures in the two experiments were the same. Statistical analyses of data from the individual experiments yielded the same results as those from the combined data. Therefore, only the combined data are presented here.

Method[1]

Participants. Participants were 44 full-term infants (22 male, 22 female) ranging in age from 5 months, 2 days to 5 months, 29 days (mean age 5 months, 16 days). Another 18 infants were excluded from the experiment: eight infants were excluded because they were fussy on three of four trials or because they became extremely upset at any point in the study; three infants were excluded because they failed to look at the initial set of dolls or the arithmetic transformation on one or more test trials; two infants were excluded because the mother made a procedural error by speaking to or touching her infant during a trial; five infants were excluded for experimenter error or equipment malfunction. All infants were from the greater San Francisco Bay Area. Parents were contacted first by letter and then by telephone; they were not monetarily compensated for their participation.

Apparatus. The apparatus was modeled after that used by Wynn (1992). It consisted of a stage-like enclosure with a small opening in the right-side panel and a trap door in the back, set into an opening in a room partition. The stage was 43 cm high, 88 cm wide, and 74 cm deep. The interior walls were decorated with pink and blue stripes, and the floor was painted blue. A mechanical, rotating gray cardboard screen (30 cm high and 28 cm wide) was mounted in the center of the stage, with the motor housed beneath the apparatus. A curtain was mounted in front of the opening of the stage. It could be lowered after each trial by an experimenter who was completely hidden behind the apparatus. The stimuli used for numerical transformations were plastic, squeezable 5-inch-high Mickey Mouse dolls. The testing room was dark except for a low illumination above the infant and a light mounted above and projecting down onto the stage.

A video camera, mounted in the partition above the apparatus, was focused on the infant. The image was recorded onto videotape and also projected online to monitors in two separate rooms, where observers recorded the infants' looking patterns and state during the session.

Procedure. Before the experiment began, we allowed the infants to play with the Mickey Mouse dolls for a few moments in order to familiarize them with the dolls' properties (e.g., that the dolls squeaked

[1] The authors thank Karen Wynn for providing them with the details of her method and procedure, which were followed for Experiments 1 and 2. She included specific details that were not described in Wynn (1992), such as placement of the dolls on the stage, timing of the transformations, and squeaking the dolls to attract infants' attention to the initial displays and the arithmetic transformations.

when squeezed). During the experiment, infants sat in a "Sassy Seat" attached to a low table situated in front of the stage-like enclosure. Infants sat approximately 1 m from the center of the stage. Parents sat in a child-sized chair to the left and slightly behind their infants. The parents were instructed not to interact with their infants during the experiment.

Two experimenters situated behind the apparatus performed the manipulations of the Mickey Mouse dolls, and raised and lowered the curtain at the start and end of trials. These two experimenters could not see the infant, nor could the infant or parent see them. An observer in a separate room, "blind" to experimental condition and trials, watched the infant on a monitor and pressed a button when the infant looked at the display. A fourth experimenter, in yet another adjacent room, executed the computer program, recorded the session on videotape, and monitored the infant's state during the experiment.

The procedure was mechanized. A computer program sent a signal in the form of a light-emitting diode (LED) to the experimenters behind the apparatus, indicating the start of trials and thus when to raise the curtain. The computer program also kept track of button presses from the on-line observer and activated the LED to indicate when to lower the curtain. The program also sent a signal to the motor to raise and lower the screen.

All trials ended when the infants looked away from the display for 2 consecutive seconds after having looked for at least 2 cumulative seconds, or after the infants had looked at the display for 30 cumulative seconds. The interval between pretest trials was 10 s. If infants failed to look at the display for 2 cumulative seconds within 20 s, the trial ended. This occurred on one pretest trial during these experiments.

Design. Infants were randomly assigned to one of two conditions: addition or subtraction. Half ($n = 22$) of the infants were assigned to each condition. Boys and girls were equally represented in each condition.

Pretest trials. The session began with two pretest trials in which the infants saw either one or two Mickey Mouse dolls standing on the stage, with the number shown first counterbalanced across infants.

Test trials. The sequence of events presented in the addition condition ($1 + 1 = 1$ or 2) was as follows. (1) A hand holding one doll entered the stage-like enclosure through the side-panel opening in the stage, squeaked the doll once to attract the infants' attention, then placed it on the stage. (2) The mechanized screen rotated up to conceal the doll. (3) A hand holding a second Mickey Mouse doll entered the stage through the side-panel opening, squeaked it once to ensure that the infants attended to the transformation, then placed the doll behind the screen and withdrew from the stage. (4) The screen rotated down, revealing either two dolls (correct result) or one doll (incorrect result). On the incorrect result trials, one doll was surreptitiously removed through a trap door at the back of the apparatus.

The sequence of events displayed in the subtraction condition ($2 - 1 = 1$ or 2) was as follows. (1) A hand entered the stage-like enclosure holding two Mickey Mouse dolls, squeaked each doll once to attract the infants' attention and placed the dolls on the stage. (2) The screen rotated up to conceal the dolls. (3) A hand again entered the display through the side-panel opening, reached behind the screen and grasped a doll, squeaked it once to ensure that the infants attended to the transformation, and removed the doll from the stage. (4) The screen rotated down revealing either one doll (correct result) or two dolls (incorrect result). On the incorrect result trials, one doll was surreptitiously added through a trap door at the back of the apparatus.

Infants were presented with six test trials. On alternating trials, the infants saw either a correct or incorrect result of the numerical transformation. Half of the infants in each condition saw the correct result first, and half saw the incorrect result first. Looking time was calculated from the point at which the screen was lowered.

Coding and reliability. Looking-time data were derived from the video recordings of each experimental session. The coders were blind to experimental condition and trial type (incorrect or correct result). This is unlike Wynn's (1992) study, where the analyzed looking-time data came from the on-line observers. We assessed reliability by comparing a second coder's results on 23% of the infants' sessions. We calculated the number of seconds the coders agreed on where the infant was looking. Agreement between coders was 96.3%.

Results

Following Wynn's (1992) analysis, we calculated a derived variable, infants' preference for looking at two dolls over one doll.[2] We then used this derived variable to compare infants' preference for looking at two dolls over one doll in the addition versus subtrac-

[2] We also conducted a factorial analysis of looking times for all three experiments in this study to determine whether there were any systematic differences in the data that could not be discerned by the overall comparison of derived scores. The results were consistent with the *t*-test comparisons. Therefore, they are not reported here, but are available upon request.

tion conditions using between-group *t* tests. These comparisons were conducted for both the pretest and test trials. Like Wynn, we used a two-tailed test of significance for the pretest trials and a one-tailed test for the test trials. Unlike Wynn, we also report a two-tailed test for the test trials because significantly longer looking at the correct result would be informative.

Consistent with Wynn's (1992) results, the infants' pretest preference for looking at two dolls over one doll in the addition condition versus the subtraction condition was not significantly different, $t(42) = -.28$, $p > .05$, two-tailed test (see Table 1). In contrast to Wynn (1992), no significant difference was found in the test trials between infants' preference for looking at two dolls over one doll when two dolls was the incorrect result (subtraction) versus when two dolls was the correct result (addition), $t(42) = -.04$, $p > .05$, one- or two-tailed test. Infants looked equally long at one and two dolls in both conditions whether the results were correct or incorrect (see Table 1).

Wynn (1992) excluded infants who had a pretest preference of 10 s or more for either one or two dolls and infants with a test preference of greater than two and one half standard deviations from the mean for their group. Therefore, we repeated the above analyses excluding six infants who had a pretest preference of 10 s or more. None of the infants met Wynn's second exclusion criterion, which demonstrates that our results were not due to outlying values in the direction opposite the hypothesis. Again, we found no significant differences in preference for two dolls over one doll between the two conditions (addition versus subtraction) in either the pretest or test trials.

Binomial tests revealed no significant differences because half of the infants ($n = 22$) looked longer at the incorrect result and half looked longer at the correct result ($n = 22$). These proportions were the same without the six infants who met the exclusion criteria.

Additionally, we conducted separate *t* test comparisons for each of the three trial pairs. These comparisons yielded a statistically significant difference between the addition and subtraction conditions in preference for looking at two dolls over one doll for the first pair of trials, $t(42) = 1.69$, $p < .05$, one-tailed test; but $p > .05$, two-tailed test (see Table 1). Looking time preferences were not significantly different on the second and third trial pairs, however, although it should be noted that we found an equally large difference in the opposite direction on the second pair of trials. Moreover, when the data of six infants who met the exclusion criteria were eliminated from the analyses, none of the comparisons was significant, even with one-tailed tests.

Discussion

Our findings are inconsistent with the results reported by Wynn. With one exception, the analyses of infants' looking times using Wynn's (1992) derived variable did not reveal any significant differences between infants in the addition and subtraction conditions in their preference for looking at two dolls over

Table 1 Mean Looking Times and Standard Deviations to Correct and Incorrect Results of Addition and Subtraction Transformations for Experiments 1 and 2 Combined ($N = 44$)

	Two Dolls		One Doll		Preference for Two Dolls over One Doll	
Looking Time	*M* (seconds)	*SD*	*M* (seconds)	*SD*	*M* (seconds)	*SD*
Addition						
Pretest trials	9.22	7.25	9.75	6.68	−.53	6.11
Test trials (1 + 1)			*Incorrect Result*			
Pair 1	9.17	7.70	10.60	8.09	−1.43	6.38
Pair 2	12.69	9.09	8.66	6.31	4.04	8.76
Pair 3	7.41	7.38	7.31	6.97	.10	8.83
Total pairs	9.76	5.78	8.85	5.45	.91	4.14
Subtraction						
Pretest trials	7.13	3.99	8.23	6.28	−1.10	7.40
Test trials (2 − 1)	*Incorrect Result*					
Pair 1	12.60	8.24	10.28	6.19	2.31	8.19
Pair 2	9.16	6.61	9.27	6.88	−.19	8.62
Pair 3	7.35	6.20	7.00	6.01	.35	7.88
Total pairs	9.70	5.18	8.85	5.01	.85	5.58

one doll. The sole exception was on the first pair of trials when Wynn's exclusion criteria were not applied and a one-tailed test used. Factorial analyses found no systematic preference for incorrect results of arithmetic transformations in particular subgroups of infants or particular experimental conditions (see Footnote 2).

EXPERIMENT 3: PRECISE SUBTRACTING

Because we had anticipated replicating Wynn's (1992) results, a concurrent third experiment was conducted. It sought to determine the parameters of infants' cognition about simple subtraction. Accordingly, we investigated a subtraction counterpart to Wynn's Experiment 3 (i.e., $1 + 1 = 2$ or 3). In our experiment, we tested infants' looking-time reactions to $3 - 1 = 2$ or 1.

Method

Participants. Participants were 24 full-term infants (12 male, 12 female) ranging in age from 5 months, 5 days to 5 months, 28 days (mean age 5 months, 16 days). Another eight infants were excluded from the experiment, two due to fussiness, one because of inattentiveness to the display, two because of experimenter error, and three because of a procedural error by the mother. Criteria for excluding infants were the same as in Experiments 1 and 2. All infants were recruited from the greater San Francisco Bay Area in the same manner as before.

Apparatus and procedure. The apparatus and procedure were identical to those used in Experiments 1 and 2.

Design. Half of the infants saw the correct result first, and half saw the incorrect result first in the test trials. In each order, half the participants were female and half were male.

Pretest trials. The pretest trials were identical to those described in Experiments 1 and 2.

Test trials. The sequence of events in this subtraction condition ($3 - 1 = 1$ or 2) was as follows. (1) Three dolls were sequentially squeaked and placed on the stage one at a time. (2) The mechanized screen rotated up to conceal the dolls. (3) A hand again entered the display through the side-panel opening, reached behind the screen and grasped a doll, squeaked it once to ensure that the infants attended to the transformation, and removed the doll from the stage. (4) The screen rotated down, revealing either two (correct result) or one (incorrect result) doll(s). On six alternating trials, the infants saw either a correct or incorrect result of the numerical transformation. Half of the infants in each condition saw the correct result first, and half saw the incorrect result first. Looking time was calculated in the same way as in Experiments 1 and 2. Once during this experiment, a test trial ended because the infant failed to look at the display for 2 cumulative seconds.

Coding and reliability. Coding was conducted in the same way as in Experiments 1 and 2, as was the calculation of reliability based on 25% of the sessions. Agreement between coders was 97.7%.

Results

Using a derived variable, we conducted a one-sample t test to compare infants' preference for looking at one doll over two dolls in the pretest and test trials against the null hypothesis of no preference. Analysis of the pretest trials revealed no significant preference for one doll over two dolls, $t(23) = -1.20$, $p > .05$, two-tailed test (see Table 2). On the test trials also, the infants showed no preference for the incorrect result, $t(23) = .74$, $p > .05$, one- or two-

Table 2 Mean Looking Times and Standard Deviations to Correct and Incorrect Results of Subtraction Transformations for Experiment 3 ($N = 24$)

	One Doll		Two Dolls		Preference for One Doll over Two Dolls	
Looking Time	*M* (seconds)	*SD*	*M* (seconds)	*SD*	*M* (seconds)	*SD*
Subtraction						
Pretest trials	7.35	4.76	9.10	4.76	−1.75	7.12
Test trials (3 − 1)	*Incorrect Result*					
Pair 1	10.28	8.89	10.30	8.27	−.02	9.76
Pair 2	8.61	8.32	8.34	7.28	.27	9.71
Pair 3	9.61	9.43	7.52	7.50	2.09	8.10
Total pairs	9.50	5.98	8.72	5.68	.78	5.17

tailed test (see Table 2). These analyses were repeated excluding three infants by Wynn's (1992) criteria, but the results did not change. We also conducted separate *t* tests for each of the three trial pairs. None of these comparisons was statistically significant (see Table 2).

A binomial test (Siegel, 1956) showed that 14 infants looked longer at the incorrect result and 10 looked longer at the correct result, a statistically nonsignificant difference ($N = 24$, $p > .05$). Without the 3 infants who met the exclusion criteria, 13 infants looked longer at the incorrect result and 8 looked longer at the correct result ($N = 21$, $p > .05$, binomial test).

Discussion

This experiment, a subtraction counterpart of Wynn's (1992) Experiment 3, was designed to further investigate infants' ability to calculate *precisely* the results of simple arithmetic transformations. In Wynn's addition problem ($1 + 1 = 2$ or 3) and in our subtraction variant ($3 - 1 = 1$ or 2), both the incorrect and correct results of the arithmetic transformation differed in numerosity from the initial untransformed set and were in the correct *direction* of the transformation. In our subtraction problem, the incorrect result was a smaller number of dolls than the correct result. In contrast to Wynn, whose incorrect addition result was a larger number of dolls, we found no evidence that young infants calculate precisely the results of simple subtraction. As in Experiments 1 and 2, the infants in this experiment showed no systematic preference for the incorrect versus the correct result of the arithmetic transformations.

One possible account for the divergence between our results and those of Wynn (1992) is the difference in working memory load posed by the initial sets in the two problems (three objects in our experiment versus one object in Wynn's experiment). Although it is possible that remembering three objects may be more taxing on working memory than remembering one (or even two), it does not seem likely that this would account for our results in the $3 - 1$ condition. First, the infants in our experiments performed similarly whether they saw one, two, or three dolls in the initial set. Second, studies of number discrimination find no differences in young infants' ability to discriminate between very small sets when the order of presentation is 3 to 2 or 2 to 3 (e.g., Starkey & Cooper, 1980; van Loosbroek & Smitsman, 1990). Third, Wynn (1995) reported that infants looked longer at the incorrect result in a condition in which they saw $3 - 1 = 3$ versus $2 + 1 = 3$, but not when they saw $2 + 1 = 2$ versus $3 - 1 = 2$. If working memory load presented by the initial set accounted for infants' success, one would expect the $3 - 1 = 3$ problem to be more difficult than the $2 + 1 = 2$ problem.

GENERAL DISCUSSION

The present experiments failed to replicate reports of young infants' discrimination between incorrect and correct addition and subtraction of very small numbers (Simon et al., 1995; Wynn, 1992; and partially by Koechlin et al., 1997, and Moore, 1997). Young infants did not discriminate when only ordinal judgments or simple expectations of change in the number of items were required (in Experiments 1 and 2 and their combined data). Given these results, it is not surprising that young infants also did not discriminate when precise subtraction was required (Experiment 3).

One possible reason that our results diverge from Wynn's (1992) is procedural. Our experimental methods were both mechanized and more controlled in two ways. (1) Our procedure was computer driven. A signal from a live observer, monitoring infants' looking from a separate room, was sent to a computer program which in turn drove the course of the test trials. (2) The looking-time data analyzed in our experiments were calculated from the videotapes of the sessions, allowing a more precise determination of infant looking. Statistical comparisons of the on-line looking times, however, yielded the same results as the comparisons of the looking times calculated from videotapes of the sessions. Since divergent findings may be due to procedures specific to different laboratories, it should be emphasized that the procedures used in our laboratory have yielded positive results in other experiments with young infants (Rivera, 1998, 1999; Rivera, Wakeley, & Langer, 1999; Schlesinger & Langer, 1999).

Another possible reason for the divergent findings is substantive. If infants' discrimination of incorrect from correct addition and subtraction results is not consistent or robust, then we should expect considerable variation across studies in infants' reactions to incorrect results of arithmetic transformations. For example, in Wynn's (1992) Experiment 1, the data presented in Table 1 (p. 750) indicate that infants in the addition condition looked equally long at the incorrect and correct results. There was only a .5 second difference between infants' looking times at two dolls (correct result) versus one doll (incorrect result). So, too, Koechlin et al. (1997) found that young infants look equally long at incorrect and correct addition results, but look longer at incorrect subtraction results.

Even the subtraction effect, however, disappeared when they included all of their 56 subjects and did not exclude 27 subjects who met Wynn's (1992) exclusion criteria. Additionally, Moore (1997), testing Wynn's (1992) $1 + 1 = 1$ or 2 and $2 - 1 = 1$ or 2 conditions, found longer looking at the incorrect results only for the female infants in his sample, not for the male infants.

Other studies by Wynn (Wynn, 1995; Wynn & Chiang, 1998) have also shown that infants look significantly longer at incorrect results of simple arithmetic transformations in some conditions but not in others. Using a variation of Wynn's original addition and subtraction procedure, Wynn (1995) compared 5-month-old infants' looking time reactions to $3 - 1 = 3$ versus $2 + 1 = 3$ in one condition and to $2 + 1 = 2$ vs. $3 - 1 = 2$ in a second condition. Infants looked significantly longer at the result of the incorrect subtraction problem ($3 - 1 = 3$), but not at that of the incorrect addition problem ($2 + 1 = 2$).

As a final example, Wynn and Chiang (1998) tested 8-month-old infants' looking times at correct and incorrect results of simple addition and subtraction problems in which either the initial set or the result was zero (i.e., an empty display). Infants in one condition were tested with a "magical disappearance" versus an "expected disappearance" ($0 + 1 = 0$ versus $1 - 1 = 0$), and in another condition, infants saw a "magical appearance" versus an "expected appearance" ($1 - 1 = 1$ versus $0 + 1 = 1$). Infants looked significantly longer at the incorrect result ($0 + 1 = 0$) only in the "disappearance" condition. Critically, infants did not look longer at the incorrect result ($1 - 1 = 1$) in the "appearance" condition. These findings appear counter to claims that infants are able to calculate precisely the results of simple arithmetic transformations. Furthermore, they are inconsistent with prior results suggesting that subtraction may be easier than addition for young infants (Koechlin et al., 1997; Wynn, 1992, 1995).

Overall, then, the studies of infant arithmetic provide mixed results regarding young infants' numerical competencies. In some experiments, infants appear to discriminate between incorrect and correct results of arithmetic transformations, while in others they do not. So, too, we found great variability between our test trial results in Experiments 1 and 2; the differences between the derived scores across the three trial pairs were 3.74, −4.23, and .25.

In addition to the variability across studies or trials, the statistical variability across infants is also high. This is reflected in the confidence intervals for magnitude of tested differences. We calculated the limits of the 95% confidence interval for the derived variable, that is the preference for looking at two dolls in the subtraction versus addition conditions.[3] For our combined Experiments 1 and 2, the confidence interval for the difference between group means (−.04) had a lower limit of −3.0 s and an upper limit of 2.9 s. This confidence interval includes the possibility of a significant difference in either direction. For the combined data of Wynn's Experiments 1 and 2, we found a confidence interval with a lower limit of 1.1 s and an upper limit of 7.3 s, around a mean of 4.2. Again, the range is quite large, and although it spans only positive values, it includes the possibility of obtaining a small nonsignificant difference. The ranges of these confidence intervals reflect considerable imprecision in estimating the effects in such studies.

Although Wynn (1995; Wynn & Chiang, 1998) proposes that various inconsistencies in the data point to specific limitations in infants' early numerical competence, the current evidence is not sufficient to determine that this is indeed the case. These inconsistent findings across and within experiments may, in fact, simply show that infants' reactions to numerical operations are variable and fragile. The results of the present three experiments support such a conclusion. This conclusion and the more radical one by Simon (1997, 1998)—that the findings by Wynn (1992) and his own replication (Simon et al., 1995) can be accounted for by young infants' non-numeric spatiotemporal discrimination—are consistent with the theory that arithmetic cognition develops later and perhaps gradually in ontogeny. While our findings are counter to reports of adding and subtracting by young infants, they are consistent with previous ontogenetic findings on the development of ordinal judgments of very small number and simple arithmetic. These ontogenetic findings, which we will now review, indicate gradual and continuous progress in these abilities with age.

Cooper (1984), using an habituation procedure, found that infants begin to discriminate ordinal relations between numbers around ages 14 to 16 months; that is, they could discriminate between less-than and greater-than relations between very small sets. Prior to that, at ages 10 to 12 months, infants discriminated between equality and inequality relations only.

Furthermore, a number of studies have investigated the development of addition and subtraction in toddlers and young preschoolers using dependent

[3] Wynn (1992) does not report variance statistics. Therefore, we used her reported t values and group means to calculate the standard error of the difference between the subtraction and addition conditions in infants' preference for two items. Based on a t value of 2.73 and a difference between means of 4.20 s, we found that the standard error of the difference was 1.54 in Wynn (1992). We used this standard error to determine her confidence intervals.

measures such as reaching, object manipulation, and verbal response (e.g., Houdé, 1997; Huttenlocher et al., 1994; Sophian & Adams, 1987; Starkey, 1992; Vilette, 1996; Vilette & Mazouz, 1998). Young children's performance on simple addition and subtraction tasks with very small sets of concrete objects indicate that by their second year children know the ordinal effect of these operations, that addition yields more and subtraction yields less (e.g., Sophian & Adams, 1987; Starkey, 1992). The ability to calculate the exact results of simple addition and subtraction problems, however, develops gradually during early childhood (e.g., Huttenlocher et al., 1994; Starkey, 1992).

Perhaps the studies with young children that bear most directly on the divergence between Wynn's (1992) results and ours are Vilette (1996; Vilette & Mazouz, 1998) and Houdé (1997). They used a Wynn-type procedure with 2- and 3-year-old French children. The procedures used in both studies were similar to Wynn's except that the children were asked to verbally indicate whether the results they saw were "bon" ("right") or "pas bon" ("not right"). Vilette (1996; Vilette & Mazouz, 1998) presented 2- and 3-year-old children with the following problems: $2 + 1 = 2$ or 3; $3 - 1 = 2$ or 3; $2 + 1 - 1 = 2$ or 3. The 2-year-olds' performance was at chance level on all three problems. The 3-year-olds were successful on the addition problem, but their performance was still at chance on the subtraction and the addition/subtraction problems. Houdé (1997) presented only addition problems with smaller set sizes to children of 2½ and 3½ years of age. The children were presented with $1 + 1 = 1$ or 2 (the same as Wynn's, Experiments 1 and 2 addition condition), and $1 + 1 = 2$ or 3 (the same as Wynn's Experiment 3). The 2½-year-olds responded correctly on the $1 + 1 = 1$ or 2 problem, but not on the $1 + 1 = 2$ or 3 problem. In contrast, the 3½-year-olds successfully solved both problems. Thus, 2½-year-olds were only successful when the incorrect solution was in the *wrong ordinal direction* of the transformation.

These findings by Vilette (1996; Vilette & Mazouz, 1998) and Houdé (1997) are consistent with results by Huttenlocher et al. (1994) and Starkey (1992) showing that, although young children know the ordinal effect of arithmetic transformations, the ability to calculate precisely simple addition and subtraction problems develops gradually throughout the toddler and preschool years. All these findings on the early development of simple adding and subtracting very small numbers are also consistent with the present findings that young infants do not add or subtract (even imprecisely). They are more difficult to reconcile with Wynn's (1992) findings that young infants add (even precisely) and subtract (but see Houdé, 1997, for a speculative reconciliation). Of course, the dependent measures are different for young infants and young children.

It is perhaps worth noting briefly that another advanced numerical skill, intermodal auditory/visual matching of very small numbers, that has also been attributed to young infants (Starkey, Spelke, & Gelman, 1990), has been followed by failures to replicate (Mix, Levine, & Huttenlocher, 1997; Moore, Benenson, Reznick, Peterson, & Kagan, 1987). There, too, even 3-year-olds do not correctly match very small numbers intermodally (see Mix, Huttenlocher, & Levine, 1996). Again, the dependent measures are different for young infants and young children: visual attention and pointing, respectively. However, this set of studies on matching of numbers by young infants and preschoolers forms a compelling parallel to the kinds of questions raised by the present set of experiments.

It is not possible to conclude from our results that infants do not have the numerical competence to discriminate between incorrect and correct results of addition and subtraction. However, in our three experiments, we found no systematic evidence of such competence. Furthermore, the review of the literature reveals inconsistent results. Therefore, whatever arithmetic competence young infants have must be fragile and subject to disruption. At this point the most comprehensive answer to whether young infants can add and subtract would seem to be "Not proved." Moreover, the developmental studies with older infants and toddlers have shown that young children do not recognize the ordinal effect of addition and subtraction until their second year and do not calculate exactly until their third or fourth year. This state of empirical research findings leads us to hypothesize that adding and subtracting are not innate but rather are arithmetic competencies that develop over the course of several years.

What seems certain from this brief review of the current body of findings, including those from the present study, is the need for much further research on the origins and early development of arithmetic cognition. Two efforts would seem to be most immediately valuable. The first would be to apply meta-analytic statistical techniques to the data of the various studies that have used the Wynn (1992) paradigm with young infants. The second would be to study the development of addition and subtraction from young infancy to early childhood, a direction that has already been fruitful (e.g., Houdé, 1997; Starkey, 1992; Vilette, 1996; Vilette & Mazouz, 1998; Wynn & Chiang, 1998) and is being pursued in our laboratory.

ACKNOWLEDGMENTS

The authors extend special thanks to Alice Klein, Matthew Schlesinger, and Prentice Starkey for their helpful insights in designing this project. They are especially grateful for the hard work and dedication of research assistants Azure Kacura, Sulki Kim, Nancy Lee, Kyra Lyons, Jaana Lehtinen, Shanti Prasad, and Lynna Tsou. The authors also gratefully acknowledge their computer programmer, Ephram Cohen, and the Institute of Human Development at the University of California, Berkeley, for supporting this research.

ADDRESSES AND AFFILIATIONS

Corresponding author: Ann Wakeley, School of Education, University of California, Berkeley, and the Department of Special Education, San Francisco State University, San Francisco, CA; mailing address: Institute of Human Development, University of California, 1203 Tolman Hall, Berkeley, CA 94720-1690; e-mail: awakeley@socrates.berkeley.edu. Susan Rivera is at Stanford University School of Medicine, Stanford, CA; and Jonas Langer is at the Department of Psychology, University of California, Berkeley.

REFERENCES

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*, 695–701.

Canfield, R. L., & Smith, E. G. (1996). Number-based expectations and sequential enumeration by 5-month-old infants. *Developmental Psychology, 32*, 269–279.

Clearfield, M. W., & Mix, K. S. (1999a, April). *Infants use contour length — not number — to discriminate small visual sets.* Poster session presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM.

Clearfield, M. W., & Mix, K. S. (1999b). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408–411.

Cooper, R. G. (1984). Early number development: Discovering number space with addition and subtraction. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 157–192). Hillsdale, NJ: Erlbaum.

Houdé, O. (1997). Numerical development: From infant to the child. Wynn's (1992) paradigm in 2- and 3-year olds. *Cognitive Development*, *12*, 373–391.

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology, 123*, 284–296.

Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numerical transformation in five-month-old human infants. *Mathematical Cognition*, *3*, 89–194.

Langer, J., Rivera, S., Schlesinger, M., & Wakeley, A. (in press). Early cognitive development: Ontogeny and phylogeny. In J. Valsiner & K. Connolly (Eds.), *Handbook of developmental psychology.* London: Sage.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (1996). Do preschool children recognize auditory-visual numerical correspondences? *Child Development*, *67*, 1592–1608.

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in human infants: Another look. *Developmental Psychology, 33*, 423–428.

Moore, D. S. (1997, April). *Infant mathematical skills? A conceptual replication.* Poster session presented at the biennial meeting of the Society for Research in Child Development, Washington, DC.

Moore, D., Benenson, J., Reznick, J. S., Peterson, M., & Kagan, J. (1987). Effect of auditory numerical information on infants' looking behavior: Contradictory evidence. *Developmental Psychology, 23*, 665–670.

Rivera, S. M. (1998). *Numerical representation, object knowledge, and word comprehension in infants.* Unpublished doctoral dissertation, University of California at Berkeley.

Rivera, S. M. (1999, April). *Numerical representation as related to an emergent symbolic system in infants.* Poster presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM.

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

Schlesinger, M., & Langer, J. (1999). Infants' developing expectations of possible and impossible tool-use events between the ages of 8 and 12 months. *Developmental Science, 2*, 196–206.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A "non-numerical" account. *Cognitive Development, 12*, 349–372.

Simon, T. J. (1998). Computational evidence for the foundations of numerical competence. *Developmental Science, 1*, 71–78.

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

Sophian, C., & Adams, N. (1987). Infants' understanding of numerical transformations. *British Journal of Developmental Psychology, 5*, 257–264.

Starkey, P. (1992). The early development of numerical reasoning. *Cognition, 43*, 93–126.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210*, 1033–1035.

Starkey, P., Spelke, E., & Gelman, R. (1990). Numerical abstraction in human infants. *Cognition, 36*, 97–128.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development*, *52*, 1146–1152.

Uller, C., Carey, S., Huntley-Fenner, G., & Klatt, L. (1999). What representation might underlie infant numerical knowledge? *Cognitive Development, 14*, 1–36.

van Loosbroek, E., & Smitsman, A. W. (1990). Visual perception of numerosity in infancy. *Developmental Psychology, 26*, 916–922.

Vilette, B. (1996). De la “proto-arithmetiques” aux connaissances additive et soustractives [From “proto-arithmetic” to additive and subtractive knowledge]. *Revue de Psychologie de l’éducation, 3*, 25–43.

Vilette, B., & Mazouz, K. (1998). Les transformations numeriques et spatiales entre deux et quatre ans [Numerical and spacial transformations between two and four years]. *Archives de Psychologie, 66*, 35–47.

Wynn, K. (1992). Addition and subtraction in human infants. *Nature, 358*, 749–750.

Wynn, K. (1995). Origins of numerical knowledge. *Mathematical Cognition, 1*, 35–60.

Wynn, K. (1996). Infants’ individuation and enumeration of actions. *Psychological Science, 7*, 164–169.

Wynn, K. (1998). Psychological foundations of number: Numerical competence in human infants. *Trends in Cognitive Sciences, 2*, 296–303.

Wynn, K., & Chiang, W-C. (1998). Limits to infants’ knowledge of objects: The case of magical appearance. *Psychological Science, 9*, 448–455.

Xu, F., & Carey, S. (1996). Infants’ metaphysics: The case of numerical identity. *Cognitive Psychology, 30*, 111–153.