Further Investigations of the Young Child's Conception of Number

Rochel Gelman and Marsha F. Tucker

University of Pennsylvania

GELMAN, ROCHEL, and TUCKER, MARSHA F. *Further Investigations of the Young Child's Conception of Number.* CHILD DEVELOPMENT, 1975, 46, 167–175. 3 experiments on the nature of the young child's use and understanding of small numbers are reported. The first considers the conditions that might lead Ss to count when making absolute estimates of small set sizes (2–5) and what effect heterogeneity has on accuracy. 144 3-, 4-, and 5-year-old children were tested with either homogeneous or heterogeneous arrays; all Ss were tested on each number at each of 3 time exposures. 48 3- and 4-year-old children were run in the remaining experiments designed to assess Ss' reactions to the effect of a surreptitious substitution of a differently colored item or a different kind of object in an expected array of 3 items. The results of the absolute-judgment experiment are consistent with the assumption that preschoolers will often count when estimating small numbers. They also provide empirical constraints for theoretical accounts of the estimator processes employed by preschoolers, for example, that they can ignore heterogeneity. The subsequent experiments show that young children treat color and identity substitutions as number-irrelevant transformations.

Our previous work on the development of number concepts distinguished between estimators and operators (Gelman 1972b). Generally, the distinction was between processes involved in the ability to (*a*) assign quantitative values (absolute or relative) to sets of objects and (*b*) interpret the effects of transforming a set. Processes which yield a quantitative representation of a set are estimators. Operator processes are used to reason about numerosity. For example, operators classify addition and subtraction as operations that alter number and displacement and classify item substitution as transformations that do not alter numbers.

A variety of processes might be used separately or in combination to arrive at a quantitative representation of a set, that is, to estimate numerosity. These include counting[1] and subitizing (the perceptual recognition of a number of objects) as well as processes which derive quantitative judgments from a variety of perceptual cues, for example, length, density, arrangement of the array. Some of these processes are more likely than others to yield accurate estimates of the actual number

The research reported in this paper and the preparation of the manuscript were supported by PHS grant no. HD-04598 and a Guggenheim Foundation fellowship held by R. Gelman. We thank the children, parents, teachers, and administrative staffs of Brookline Elementary School, Greene Towne Montessori School, Chestnut House, Har Zion-Wynnefield Nursery School, S. Paley Day Care Center, Saint Mary's Cooperative School, Tabernacle Nursery School, YW-YMHA Broad Street Nursery School—all in or near the Greater Philadelphia area—and the Barn Child Care Center, Irvine, California. We also thank Janet May for her assistance in running some of the experiments and data analyses, and C. R. Gallistel for his critical reading of an earlier version of this manuscript. Requests for reprints should be sent to R. Gelman, Psychology Department, University of Pennsylvania, 3813–15 Walnut Street, Philadelphia, Pennsylvania 19174.

[1] We use counting to refer to the ability to coordinate the number name series with items in an array and not simply the ability to recite the number series. Note that we treat counting as an estimator, whereas Klahr and Wallace (1973) treat it as an operator. Our reason for doing so is that counting is generally not an operation that is performed on quantitative representations. Instead, it usually serves as a procedure for translating sets of objects into a quantitative representation. However, if, by definition, counting involves an assumption of iteration (repeated addition), then counting can be classified as both an operator and an estimator.

of items in a set (or sets). The process used to estimate may affect the results of experiments designed to test the child's operator scheme for number (Gelman 1972b). By taking precautions to make sure the children were dealing with numbers they could accurately estimate, we have shown that preschool children categorize the transformations that can be performed on small sets (two to four) as number-relevant and number-irrelevant (Gelman 1972a, 1972b). When children in a "magic" experiment encounter arrays whose spatial properties (e.g., length, density) have been surreptitiously altered, they do not think that the numerosity has been affected. In contrast, when they encounter arrays whose numerosity has been surreptitiously changed, they postulate subtraction and addition transformations to explain the unexpected changes in number.

In this paper, we investigate first, the nature of the processes by which preschoolers estimate small numbers, and second, the generality of the number-relevant versus number-irrelevant categorization scheme in the child's operative thinking about small numbers.

Small-Number Estimation

There is now considerable evidence that 3- and 4-year-old children can accurately estimate the numerosity of set sizes of one to four and sometimes five (Beckmann 1924; Descoeudres 1921; Gelman 1972b; Lawson, Baron, & Siegel 1974; Smither, Smiley, & Rees 1974). The first experiment addresses two questions about estimators. (1) To what extent does the estimator used by the young child ignore number-irrelevant properties of the set? (2) To what extent do young children count—even when estimating small numbers? The first question is addressed by studying the effect of stimulus heterogeneity on estimation accuracy; the second, by studying the effect of stimulus presentation time or estimation accuracy.

There are some studies of the effect of stimulus heterogeneity on the preschool child's ability to estimate number, but the available data are inconsistent. Von Gast (1957) reported that preschoolers who accurately assign cardinal numbers to small sets (one to five) of homogeneous items fail to do so when confronted with sets whose items vary in color, size, shape, and identity. However, Beckmann (1924) found that 3- and 4-year-olds consistently estimated small numbers (one to four) accurately, even when the sets were comprised of heterogeneous materials. It was our guess that Von Gast's children were set to establish an expectancy for homogeneous arrays (which were presented first), an expectancy that may have masked their ability to estimate heterogeneous arrays accurately. In the present experiment, Ss were assigned independently to homogeneous and heterogeneous conditions. This was done to control for the possibility that Ss would establish an expectancy for one type of stimulus—an expectancy that could interfere with their performance on the other type.

It is generally assumed that the young child relies primarily on perceptual processes, for example, subitizing, to estimate small numbers (Klahr & Wallace 1973; Piaget 1952; Pufall, Shaw, & Syrdal-Lasky 1973). Some time ago Beckmann (1924) drew attention to the possibility that a counting process is also available to the young child. He pointed out that very young children (2 and 3 years of age) were likely to count aloud when estimating set sizes as small as two and three. And Gelman (1972b) noted a tendency for 3- and 4-year-old children to count aloud in order to determine whether unexpected alterations in small sets (e.g., increases in length and number) corresponded to an actual change in number. Such counting-aloud data do not allow one to conclude that the child who counts fails to use perceptual processes when estimating small numbers (see Klahr & Wallace 1973). They simply highlight the fact that there are conditions under which a child can and will count in order to estimate. Indeed, as Schaeffer, Eggleston, and Scott (1974) argue, perceptual and counting processes may be used in concert. Here we focus on investigating the extent to which young children invoke a counting mechanism in order to arrive at a representation of the numerosity of small sets. The decision to vary stimulus exposure time in order to assess the young child's tendency to utilize a counting process when estimating derived from the following considerations. Counting involves a serial tagging process—a process that takes time. For the child who might be inclined to count, the longer he has to process a given set size, the more opportunity he has to count. Accordingly, increases in stimulus-exposure time might be expected to increase the child's tendency to count. And if, as Beckmann suggests (1924), the young child prefers to use a counting process when estimating, increases in stimulus-exposure time should result in increases in estimation accuracy.

The Estimator Experiment

Method

Design.—The design of the experiment involved four main variables: (*a*) age (3-, 4-, or 5-year-old groups); (*b*) array type (homogeneous or heterogeneous); (*c*) set size (2, 3, 4, 5, 7, 11, and 19);[2] and (*d*) exposure time (1 sec, 5 sec, 1 min). Age and type of array were between-group variables. Set size and exposure time were within-group variables.

Half of the Ss in each age group were assigned haphazardly to either the homogeneous or heterogeneous condition. An effort was made to assign comparable numbers of girls and boys to each independent group.

Each S was tested on each set size at each exposure time. The sets of stimuli used at each exposure time included one linear array for each of the numbers. The items forming an array of a given number were noticeably different from one exposure time to the next. At a given exposure time, the order in which a child encountered the various numerosities was randomized.

All arrays for a given exposure time (1 sec, 5 sec, 1 min) were shown before S was tested at the next exposure time. The order of times was varied as follows: the 1-min exposure was alway run last. Whether a S first encountered the 1-sec or 5-sec series was counterbalanced. The choice of 1 sec as the shortest exposure time reflected our conjecture that the Ss could perceive the stimulus array within 1 sec.

Subjects.—The Ss were 144 children (76 boys, 68 girls) attending nursery schools and kindergartens in the Philadelphia area. They were predominantly middle class. There were three age groups with 48 children in each. The median ages of the 3-, 4-, and 5-year-old groups were 3.7, 4.5, and 5.3, respectively.

Stimuli.—The homogeneous stimuli consisted of three sets of 14 × 35.5-cm white posterboard cards. Each card displayed a row of identical items. Each set of cards contained one card for each set size, that is, numerosity. All Ss saw an array of a given numerosity three times—once during each of the three exposure-time conditions—but the color (red or blue) and shape (star or circle) of the items comprising a given numerosity were varied across exposure times. The items on each card were approximately 0.6 cm in diameter. Within each stimulus condition half the Ss were shown displays where density was fixed (1.5 cm between the center of items) and length varied, and half were shown displays where length was fixed (at 31.5 cm) and density varied. The cards were displayed at a distance of approximately 70 cm. Thus, in the constant-length condition the visual angle subtended by the array was approximately 25.4° regardless of numerosity. In the constant-density condition the visual angle for the set sizes considered in this paper (two to five) ranged from approximately 1.7° for a set size of two to 5.4° for a set size of five.[3]

Procedure.—Each S's participation took 2 days. On day 1, *E* played with S using commercially bought toys. Day 2 consisted of the experiment proper.

To start the experiment, *E* told S that he would be shown some cards with stars and/or circles on them and that he was to tell *E* how many things he saw on each card. The *E* then showed S a sample card. During the testing, *E* repeated the basic question: "How many things are there on this card?" just before presenting each display. The *E* presented the cards by holding them upright at S's eye level. The child was allowed to answer when he was ready. The time variable was controlled by *E*. For the 1-sec and 5-sec conditions she counted to herself. For the 1-min condition she held up the card until 1 min had elapsed or Ss gave their response, whichever occurred first. We recognize that this method of varying time exposure is crude. We chose it for its simplicity and with the age of our Ss in mind. In our experience, the fewer props and instructions, the greater the chances of gaining a young S's

[2] The larger arrays were included as part of a project on the young child's notion of large numbers.

[3] The apparent discrepancy in the range of values of visual angle for the constant-density and constant-length conditions is due to the fact that the data reported in this paper are only for set sizes of one to five.

attention and cooperation in this kind of experiment.

Within each time condition, there was at least a 30-sec break between items. The *E* used this time to record S's answer and tendency to count overtly.

Results

Since this paper addresses questions about the young child's small-number estimators, results for arrays larger than five are not presented. The accuracy scores on these larger set sizes were comparable to those reported in the literature reviewed by Gelman (1972b). The conclusions reached below are not altered by considerations of the large-set results.

Chi squares comparing the performance of Ss tested first at 1 sec or first at 5 sec revealed no significant effects. Similarly, the fixed density–fixed length variable showed no systematic effects. Accordingly, subsequent analyses ignored these variables.

To determine the effect of heterogeneity, we computed χ^2 (or Fisher Exact probabilities, when expectancies were small) for each number at each exposure time and age. Of the 45 comparisons, only one yielded significance at $p < .05$. In view of the number of comparisons, we attach no significance to the fact that one of them exceeded the conventional confidence level for rejection of the null hypothesis. Thus, heterogeneity did not significantly affect accurate estimation at any age. Accordingly, we combined the data from the heterogeneous and homogeneous conditions in analyses of the effect of set size and exposure time.

Set-size and exposure-time effects.—The number of Ss who gave accurate number answers for each set size at each exposure time is presented in table 1. There is an overall effect of exposure time on accuracy, with Ss performing best in the 1-min condition. When Ss were given a long time to answer the question, the results were comparable to those of other investigators (Beckmann 1924; Descoeudres 1921). Almost all Ss at all ages were able to assign the correct answer to two-item arrays. A very considerable proportion of Ss were accurate with three-item arrays. The accuracy of the 3-year-old group drops off with four-item arrays. Each successive age group is successively better able to deal with the numbers three to five.

In table 1, there also appears to be an age × set size × exposure time interaction. As exposure time increases, so does the 3-year-olds' accuracy on every number, including the number two. In contrast, time did not affect the 4- and 5-year-olds' performance on the number two. But as the number to be estimated increased, so did the tendency for these Ss to be adversely affected by reduced exposure time.

The statistical significance of the effect of time on accuracy for a given number was assessed with Cochran-Q tests. The results of these tests are shown in table 2. It can be seen that the 3-year-olds' performance on all numbers—including the numbers two and

TABLE 1

N Subjects (of 48) Correct in Each Condition

Age and Exposure Time	Set Size				Total Trials Counted by *Ss*
	2	3	4	5	
3 years:					
1 sec	33 (14)[a]	28 (10)	9 (11)	8 (11)	46
5 sec	41 (20)	38 (21)	21 (23)	16 (24)	88
1 min	41 (26)	40 (31)	28 (31)	27 (33)	121
4 years:					
1 sec	44 (3)	37 (4)	23 (5)	17 (8)	20
5 sec	44 (7)	41 (13)	29 (18)	21 (23)	61
1 min	45 (14)	42 (19)	37 (23)	32 (24)	80
5 years:					
1 sec	47 (2)	43 (2)	33 (3)	23 (2)	9
5 sec	44 (3)	44 (3)	37 (7)	26 (12)	25
1 min	47 (3)	46 (6)	42 (11)	38 (24)	44

[a] Number in parentheses indicates number of *Ss* who noticeably counted in each condition.

TABLE 2

Cochran-Q χ^2 Values for Time-variable Effects on Set Sizes of Two to Five

Age	Set Size 2	3	4	5
3 years	14.92*	51.00*	8.30*	19.50*
4 years	0.33	4.20	11.38*	11.50*
5 years	3.60	4.60	9.38*	12.19*

* $p < .05$, $Q(.05, 2) = 5.99$.

three—was significantly affected by time. In contrast, it is only for the large numbers (four and five) that the older Ss' performance was affected by time.

Table 1 includes the data on Ss' tendency to count overtly. As Klahr and Wallace (1973) point out, the reliance on overt counting behavior as an index of the extent to which Ss count is problematic. Such a measure fails to separate Ss who subitize and count from those who only count. We are primarily interested in the question of whether or not a S counts somewhere in the course of an estimation trial and do not take the position that a S who counts does not subitize as well. Therefore, this problem does not concern us here. Klahr and Wallace (1973) also note that a count-aloud measure is insensitive to the possibility that Ss count to themselves. Inasmuch as counting-aloud data are presented in table 1 (these are shown in parentheses), attention is drawn to the fact that this measure could be a conservative index of the extent to which Ss in fact counted.

Inspection of these counting data reveals several trends. First, consider only the counting tendencies for the 1-sec exposure times. It can be seen that 3-year-old children were more inclined to count overtly than were 4-year-olds, who were in turn more inclined to count than were 5-year-olds. Thus, even with a brief stimulus-exposure time, the younger the S, the greater the tendency to count. Second, the tendency for the youngest Ss to count aloud more than the older Ss also holds for the 5-sec and 1-min exposure times. Third, there appears to be an age × set size interaction effect on Ss' tendencies to count aloud: 3-year-olds tended to count aloud regardless of set size. With older children, the larger the set size, the more likely they were to count aloud. Finally, there is an overall effect of time exposure. The longer children had to process arrays, the more they counted aloud.

Discussion

Our results indicate that young children can and do count when estimating small set sizes. Indeed, the younger the child, the greater the tendency to count overtly. The decrease in the overt counting of small sets in older children may reflect either or both of two developments: older children may do more covert counting (cf. Vygotsky 1962); or older children may rely more on estimator processes that do not involve counting, for example, subitizing.

Our data also establish certain constraints on the nature of the processes underlying the young child's estimation of small numerosities. The estimator processes must be such that: (1) given at least 1 sec of processing time, they are indifferent to item heterogeneity, that is, variations in color and shape among items within a set; (2) at stimulus-exposure times of 1 sec or more, they work equally well over visual angles ranging from 1.7° to 25.4°; and (3) increases in exposure time from 1 sec to 1 min increase the accuracy of estimation. For youngest Ss this is so even for set sizes as small as two and three. All of these properties of the estimation process are readily explained by assuming that a counting process is used to estimate numerosity. Of course, number-estimating procedures not based on counting may also satisfy these constraints. These empirical constraints may help promote a more detailed theoretical specification of the processes that are said to yield an estimate of number without relying on counting, for example, intuition (Piaget 1952), or subitizing.

Operators

A mature conception of number involves, in addition to number estimators, a scheme or set of rules that classify transformations according to the effects they do or do not have on number. There is evidence that children as young as 3 correctly classify addition and subtraction as events which increase and decrease number (Gelman 1972a, 1972b). They also classify alterations in length as irrelevant to number. The findings of the study presented above raise the question of whether young children know that substitution transformations, which alter the color or identity of elements in small sets, are likewise irrelevant to number. To determine whether young children know that a set maintains its numerosity even when items within the set are substituted, the "magic" paradigm was used (Gelman 1972a). This paradigm involves a two-phase experiment: the first phase establishes an expectancy for a given numerosity; the second assesses the child's reaction to variations in the expected set—variations that are produced surreptitiously.

The results of a pilot study and a follow-up study are reported below. In the pilot study children initially saw two linear sets of homogeneous items—a three-mouse and a two-mouse set. Then in phase 2 they confronted the effect of *E*'s having substituted either a differently colored mouse (color-change condition) or a different object (identity-change condition) for one of the mice in the original three-mouse set. In the follow-up study, children initially saw a three-item array containing two green mice and a soldier and an array with two green mice. Then in phase 2 they confronted the effect of *E*'s having replaced the soldier on the three-item array with a green mouse. On the basis of D'Antonio's work (1970), we expected that the children would be surprised by and would describe the covertly introduced variations in color and identity. The question was whether they would decide that these changes were irrelevant to number and indicate that number of objects was the same as expected—despite the changes.

General procedure.—The experiment was run over 2 days. On day 1, *E* played games with each child for approximately 15 min. On day 2, *E* administered both phases of the "magic" task.

The materials used for the magic task were toy mice (green and orange) and a toy soldier, all about the same size (approximately 1.2 × 3.2 cm); two white plates (23.4 cm in diameter); some Velcro cloth to hold the toys in position on the plates; and two white cake covers (8.5 cm high × 25.5 cm in diameter).

The procedure in the magic task was essentially the same as that described in detail elsewhere (Gelman 1972a). Briefly, the procedure involved two phases. The first phase began with the uncovering of two plates—one with a row of three items, the other with a row of two items. The rows were either (*a*) of the same length, and therefore different densities, or (*b*) of the same density, and therefore different lengths. Whether the density or the length difference was redundant to the number difference was counterbalanced. The first time S saw the plates, *E* asked S what was on the plates. Without mentioning the numerosity, *E* then pointed to the three-item plate and told the child that it was the winner plate. Similarly, *E* indicated that the 2-item plate was the loser. Then, *E* told S that S and *E* would take turns hiding the plates under the cans and mixing the cans. After each covering and shuffling, the child was asked to guess which can was hiding the winner plate. He was then told to lift that cover, and he was asked if he had found the winner. If S said he had found the winner and was correct, he was given a prize and a new trial began. If S successfully identified the loser, he was asked where the winner was. In this case, all children pointed to the remaining can, uncovered the plate, and indicated whether it was the winner. If they were correct, they received a prize and the game continued. Whenever a child made an identification error, he was corrected (e.g., "no, that's the loser") and a new trial began. Every lifting of a can counted as a trial. On three of the phase 1 trials, S was asked why he thought a plate was a winner or loser.

After a minimum of 10 or 11 phase trials, in which S responded correctly on at least five of the last six trials, phase 2 began. From the child's point of view the onset of phase 2 was just another trial. To start phase 2, *E* covertly replaced one item on the winner plate with a different item. The position (left, right, or middle) of the substituted item was counterbalanced.

Phase 2 continued until S uncovered the transformed array. When S uncovered the altered array, he was asked if it was the winner. The *E* noted overt surprise reactions and asked the child why the plate won or lost, whether anything had happened, and, if so, what. Children who noticed a change were asked if the change mattered and whether the plate could win even though it had been changed. Then all children were asked how many there were now. Where necessary, the children were asked how many things were on the plate (this was done only near the end of the identity-change condition in the pilot experiment). Finally, Ss were asked where and what the items had been in the original display, who could fix the game, and how it could be fixed. If the child asked for the original object, he was provided with a handful of objects containing two like the missing one, as well as other objects.

The Pilot Study

Thirty-two 3- and 4-year-old children completed both phases of the experiment. Three Ss were dropped when they failed to identify the winner and loser correctly. There were eight children from each age group (median ages 3.8 and 4.7) in each experimental condition.

As indicated above, the initial arrays in this experiment contained homogeneous items—green mice. Thus the phase 1 procedure was identical to our published studies with two- and three-item displays. And, as expected, the phase 1 results were like those previously reported (Gelman 1972a, 1972b). Subjects had little difficulty in learning to identify the "winner" and "loser" plates, and they indicated that their identifications were based on number.

In phase 2 almost all Ss were visibly surprised by the changes. Still, their other reactions indicate that all Ss in the color-change condition judged the color change to be irrelevant to number. Thus, in this condition, Ss said the altered plate still won; justified this position with reference to the correct number, for example, M. M. (3.8) said, "It has one-two-three, three"; and said that there were three mice before and after the transformation. Only five Ss in the identity-change condition behaved like the color-change Ss. The rest tended to suggest that the winner plate no longer won because it only had two mice (see table 3).

The results of the identity-change condition might be taken to indicate that young children are unable to treat changes in identity as irrelevant to number. But before we can accept this conclusion, we consider another possible interpretation of these results. This is that identity-change Ss thought that the winner had to have three items of the same identity class—an idea that could have resulted from the fact that all items on the original winner and loser plates were mice and therefore in the same identity class. Some evidence for this position comes from the fact that all identity-change Ss answered the "how many *things* are there now" correctly. Thus, even though most Ss said the winner plate only had two mice (and therefore was a loser), in response to a probe they were able to say that the same plate had three things.

To check for the possibility that we set these Ss against dealing with the number of things in an array, we ran a follow-up identity-change condition. This involved starting with a heterogeneous array in phase 1 and ending with a homogeneous array in phase 2. If our account of the previous identity-change results is correct, we should find that the children establish expectancies based on both the identity of the objects and the number of objects in the array. The fact that they initially confront a heterogeneous array should provide the child with a clue that the objects need not be homogeneous. If they can take advantage of this clue, then they might be expected to reveal an ability to treat an identity substitution as a number-irrelevant transformation.

Follow-up Study

Method

This experiment involved running the identity-change condition backward. In phase 1, the winner plate had two green mice and one soldier; the loser plate had two green mice. At the end of phase 1, *E* surreptitiously replaced the soldier with a green mouse. Thus, in phase 2, the child encountered three green

TABLE 3

SUMMARY OF PHASE 2 IDENTITY-CHANGE RESULTS

Type of Subject	*N*	Mean Surprise Score[a]	% Who Say They Win	% Who Say "Three" Given They Say Win	% Who Say They Lose	% Who Say "Two" Given They Say Lose
Treat identity change as irrelevant	5	1.00	100	100	0	...
Treat identity change as relevant	11	1.73	31	0	69	75

[a] The surprise scores were based on a 0 (no discernible surprise), 1 (minimal surprise), and 2 (moderate to extreme surprise) rating scale as described in Gelman (1972a).

mice and two green mice. Otherwise, the procedure followed in this experiment was identical to that described above.

The Ss were 17 children (eight boys, nine girls) attending a nursery school in Irvine, California. One of the 3-year-olds was dropped when he refused to continue the game. This left eight Ss in the 3-year-old group (median age 3.5) and eight Ss in the 4-year-old group (median age 4.6). The sample was middle class and ethnically heterogeneous.

Results and Discussion

All Ss learned to identify the winner and loser correctly during phase 1. Thirteen Ss made no identification errors. The remaining Ss (all 3-year-olds) averaged 1.3 errors before reaching criterion.

The responses to phase 1 probe questions provide clear evidence that the children formed expectancies for the kinds of objects that were on the plates; the evidence for number-based expectancies is less clear. Two 4-year-old Ss gave neither type of explanation, saying the winner and loser plates won or lost because they just did. Twelve of the remaining Ss provided explanations that involved highlighting the fact that the winner plate had a soldier whereas the loser plate did not. For example, in response to the probe question S. G. (3.4) said, "mouses and a man" and "mouses," and L. B. (4.0) said, "a boy" and "a mouse." Of these 12 Ss, five also gave accurate number responses; the eight Ss who failed to provide identity accounts gave only number-based explanations. Examples of number accounts come from S. G. (3.4) who also said he won and lost because "one-two-three, three things" and "two things" and A. A. (3.6) who said, "mouses and more; two and three." In all, then, 12 Ss used identity criteria and seven Ss used number criteria.

Thus, if we consider only the phase 1 results, we might conclude that the children were more likely to base their judgments on the identity differences represented in the arrays than the number differences. However, the results of phase 2 go against this conclusion. They instead show that the children formed expectancies for both number and identity.

Table 4 summarizes the phase 2 results. The surprise scores reflect the fact that all but one 4-year-old S showed some surprise when they discovered that there was a plate with three mice and no plate with a soldier on it. All the children who noticed the change explicitly said that somehow a mouse was substituted for a soldier or that a three-mouse plate was substituted for the expected plate. Nevertheless, when asked which of the phase 2 plates won, all the children picked the three-mouse plate. And 14 Ss said it won because it had three or three things. Finally, all but one of the 15 children who noticed the substitution gave some evidence of knowing how to reverse the effect of the substitution. A child was scored as a reverser if: (*a*) he said either he could remove a mouse from the three-mouse plate and put a soldier in its place or he needed a soldier to put on the two-mouse plate; and/or (*b*) the child asked for a soldier and, when given a handful of objects, picked out one soldier and produced a two-mouse plus one-soldier array.

As presented, the phase 2 data failed to capture two other results of interest. First, although the children did end up choosing the three-mouse plate as the winner and said they did so because it had three items, it is not that they did so without hesitation. Consider what D. B. (4.2) said when asked if the altered plate was the winner: "What happened to the drum man?" (Is it the winner plate?) "It is and isn't. It is why it has three mouses. It doesn't have a drum man." Other children would not answer the "Is that the winner?" question until they had looked at both plates, had asked about the missing soldier, and/or had even covered the plates and reshuffled. To us, such behaviors indicate that the children thought it would be preferable to have the original heterogeneous three-item arrays. Inas-

TABLE 4

Summary of Results in Follow-up Magic Experiment

Age	Mean Surprise Score	*N* Noticers	*N* Choosing Three Plate as Winner	*N* Noticers Explaining Adequately	*N* Noticers Reversing
3 years	1.75	8	8	8	7
4 years	1.50	7	8	7	7

PP1 (7).max

much as they could not produce the expected array, they were willing to settle for the homogeneous array on the grounds that it, like the original array, had three items.

The phase 2 results so far discussed suggest that children did form expectancies based on identity and number. Another result provides further support for this conclusion. Twelve of the children were scored as reversers on the basis of what they did when *E* handed them extra objects. The reversal criterion was whether or not they selected one soldier and constructed an array like the winner array in phase 1. Of interest is that eight of the children ended up with one plate that had three mice and another plate that had two mice and a soldier. Their justifications for doing this are telling. All spontaneously said they did this because that meant both plates would have three things and therefore there would be two winner plates. This is illustrated in the following protocol. D. S. (4.7): [*E* asks S how to fix the game.] "You take this [a mouse] off and put on a soldier. Where's a soldier?" [*E* gives S extra object.] "How about two winner ones?" [*S* places the soldier on the two-mouse plate and says:] "This is gonna be a winner plate too. Both have three things." In other words, these children decided that, in the end, number was the better criterion—from their point of view.

The results of both magic experiments together lead to the conclusion that young children can treat substitution as a number-irrelevant operation, at least when they make judgments about small numbers.

Summary

The findings from the estimator experiment indicate that there are conditions which enhance the young child's tendency to count. The conditions that appear to promote reliance on a counting process also are the ones that increase the child's accuracy. The results of this study provide some empirical constraints for the theoretical description of the young child's ability to estimate small numbers accurately. Such constraints are met by a counting-based process; whether they likewise are met by a noncounting-based procedure is a question for further exploration.

The failure to find an effect of heterogeneity on estimation-accuracy scores is congruent with the results of the operator experiments. Together these findings indicate that the young child's estimators and operators can be indifferent to heterogeneity.

References

Beckmann, H. Die Entwicklung der Zahlleistung bei 2–6 jährigen Kindern. *Zeitschrift für Angewandte Psychologie,* 1924, **22**, 1–72.

D'Antonio, M. The development of invariance concepts: color and shape. Unpublished doctoral dissertation, University of Pennsylvania, 1970.

Descoeudres, A. *Le développement de l'enfant de deux à sept ans.* Paris: Delachaux & Niestlé, 1921.

Gelman, R. Logical capacity of very young children. *Child Development,* 1972, **43**, 75–90. (a)

Gelman, R. The nature and development of early number concepts. In H. W. Reese (Ed.), *Advances in child development and behavior.* Vol. **7**. New York: Academic Press, 1972. (b)

Klahr, D., & Wallace, J. G. The role of quantification operators in the development of conservation of quantity. *Cognitive Psychology,* 1973, **4**, 301–327.

Lawson, J.; Baron, J.; & Siegel, L. The role of number and length cues in children's quantitative judgments. *Child Development,* 1974, **45**, 731–736.

Piaget, J. *The child's conception of number.* New York: Norton, 1952.

Pufall, P. B.; Shaw, R. E.; & Syrdal-Lasky, A. Development of number conservation: an examination of some predictions from Piaget's stage analysis and equilibration model. *Child Development,* 1973, **44**, 21–27.

Schaeffer, G.; Eggleston, V. H.; & Scott, J. L. Number development in young children. *Cognitive Psychology,* 1974, **6**, 357–379.

Smither, S. J.; Smiley, S. S.; & Rees, R. The use of perceptual cues for number judgment by young children. *Child Development,* 1974, **45**, 693–699.

Von Gast, H. Der Umgang mit Zahlen und Zahlgebilden in der frühen Kindheit. *Zeitschrift für Psychologie,* 1957, **161**, 1–90.

Vygotsky, L. S. *Thought and language.* Cambridge, Mass.: M.I.T. Press, 1962.