# Implicit Learning of Arithmetic Principles

Richard W. Prather & Martha W. Alibali

*Abstract*—**Past research has investigated children's knowledge of arithmetic principles over development. However, little is known about the mechanisms involved in acquiring principle knowledge. We hypothesize that experience with equations that violate a to-be-learned principle will lead to changes in equation encoding, which in turn will promote acquisition of principle knowledge. Adults' knowledge of an arithmetic principle was evaluated before and after a training session in which some participants were exposed to equations that violated the principle. Participants who were exposed to temporally proximal principle violations increased their knowledge more than participants who were exposed to widely spaced violations. Learners with low principle knowledge post-training were also poor at encoding key features of the equations. Thus, variations in experience lead to variations in principle learning, and encoding is an important component of principle knowledge.**

*Index Terms*—**Cognitive Development, Implicit learning**

## I. Introduction

Principles can be defined as regularities or general rules within a domain. For example, one principle that applies to the domain of mixture problems is that the concentration of the final solution must be in between the concentrations of the two initial solutions. Learners have been shown to use principles in a variety of domains, including counting [1], proportional reasoning [2], artificial grammar learning [3, 4], and language acquisition [5-7]. Studies from the domain of language acquisition [6, 7] provide a representative example of research on principle learning. In these studies, infant participants were exposed to linguistic input (sometimes for as little as 2 minutes) that contained regularities. Participants were then tested with stimuli that either corresponded to or violated those regularities. Looking-time evidence indicates that infants are capable of learning linguistic regularities in this manner (however, for alternative views see [8] and [9]). These linguistic regularities can be thought of as principles of language.

The focus of the present study is arithmetic principles, specifically the *Relationship to Operands* principle. Generally stated, this principle describes the relationships between the operands and the result in an arithmetic equation. The exact relationship varies depending on the operation. In simple addition equations (A + B = C), the sum (C) must be greater than the two addends (A and B). In simple subtraction equations (A - B = C), the difference (C) must be less than the subtrahend (A), however it may have any relationship with the minuend (B). In simple multiplication equations (A x B = C), the product (C) must be greater than both operands (A and B). In simple division equations (A ÷ B = C), the quotient (C) must be smaller than the dividend (A), however it may have any relationship with the divisor (B). These are all considered examples of the *Relationship to Operands* principle, though the details for each operation differ slightly.

A large amount of literature is devoted to learners' knowledge of arithmetic principles. Most of this literature focuses on characterizing learners' knowledge at particular points of development. This includes infants [10], preschool age children [11], older children [12], and adults [13]. This sort of inventory of knowledge and its development is certainly necessary and useful. However it leaves unanswered the question of how principle knowledge is acquired. This is not to say that no study to date has in part addressed the mechanisms of principle acquisition; however this issue has not been the focus of the vast majority of the literature.

### A. *Principle learning in artificial grammars*

A large amount of research has investigated the learning of regularities in artificial grammars. This research often investigates the implicit learning of these regularities. In a typical artificial grammar learning study, participants view a large amount of input that corresponds to a particular artificial grammar [4], [14]-[17]. These grammars involve strings of letters of varying length, such as ABFE or ABBQW, that correspond to some predetermined finite state grammar. Participants may be instructed to memorize the examples or otherwise pay attention to them, and they are not told that there are any regularities in the stimuli to learn. After the initial exposure, participants are told that a regularity was present, and their knowledge of the regularity is assessed via their evaluation of novel examples that either violate or are consistent with the regularity. The general consensus is that participants are often able to learn what seem like complex regularities. Most relevant to the current research is the idea that participants are able to learn regularities in a domain through exposure to examples that are structured in specific ways. In artificial grammar studies, principle learning is usually accomplished via experience with principle-consistent examples as opposed to examples that violate the principle.

R. W. Prather is at the University of Wisconsin, Madison, WI, 53706 (phone: 608-263-0757; e-mail: rwprather@wisc.edu).

M. W. Alibali, is at the University of Wisconsin, Madison, WI, 53706 (e-mail: mwalibali@wisc.edu).

The first author is funded by an Institute of Education Sciences Interdisciplinary Training Program predoctoral fellowship.

### B. *Arithmetic principle learning*

Though research in other domains is certainly suggestive regarding mechanisms of acquisition of arithmetic principles, research specifically on arithmetic would be more informative. Few studies specifically address changes in arithmetic principle knowledge.

In one of the few studies to address principle acquisition [18], Dixon and Bangert had participants engage in a task for which knowledge of an arithmetic principle was required for success. They found that increased principle knowledge was predicted by *successive* correct trials on the task, and not by total number of correct trials. That is, learners who were successful in a temporally proximal way were more likely to learn the principle than learners whose success was spread out. It should be noted that Dixon and Bangert used a "natural learning" paradigm in which there were no manipulations intended to influence learning. Their results suggest that it would be valuable to investigate the acquisition of arithmetic principles under more controlled conditions.

Dixon and Bangert theorized that the mechanism by which learners gain arithmetic principle knowledge involves changes in learners' internal representations [18]. Internal representational change as a mechanism for learning is consistent with some theories of cognitive development such as Representational Redescription [19]. Prior research in a different domain (gear problems) also suggests that changes in learners' internal representations lead to changes in behavior [20]. Dixon and Bangert argue that in order to change representation, the learner needs to have the relevant regularity highlighted repeatedly within a short time span. Once a particular regularity is highlighted enough, the learner will integrate that regularity into his or her representation of the domain. Thus, repeated temporally proximal experiences that highlight a specific arithmetic principle should lead to the acquisition of that principle by the learner. Dixon and Bangert interpreted this to mean that arithmetic principle acquisition would be predicted by temporally proximal principle-consistent experiences, and their findings supported this hypothesis. This view has also been supported to some extent by prior work on temporal proximity and arithmetic learning [21].

The current study expands on Dixon and Bangert's initial investigation of arithmetic principle acquisition, drawing on what is known about learning of regularities in other domains. The general goals of the current research are twofold: first, to articulate a theory of arithmetic principle acquisition, and second, based on that theory, to investigate which aspects of input facilitate learning of arithmetic principles.

### C. *Encoding as a learning mechanism*

Previous research has attributed changes in learners' knowledge to changes in their *encoding* or *representation* of the problem domain [19], [22]. What exactly is meant by *encoding* or *representation*? For a given stimulus, whether it be a simple arithmetic equation, a chess board scene, or a physics problem, there are many features that can be encoded. Consider the example of a simple equation, $8 + 3 = 11$. There are many features of the equation that could be noted: the color of the numerals, the order of the numbers, the type of operation, the value of the operands, and the relative magnitude of the operands and the sum. An experienced arithmetic learner should be able to encode these characteristics from the external stimulus. The learner uses these encoded characteristics to form an internal representation of the equation. A learner's representation of a particular problem may draw, not only on information encoded from the given problem, but also on other sources of knowledge from long-term memory, such as prior knowledge of common problem schemas.

Problem encoding and representation vary as a function of domain knowledge or expertise. Studies comparing novices and experts in several domains (e.g., chess, physics) have shown that experts more accurately encode relevant displays. For example, Chase and Simon found that expert chess players reconstructed complex chess positions better than novices [23]. Experts did not generally have better memory, but they encoded and represented the positions differently from the novices.

In the domain of arithmetic, research has linked deficits in encoding with difficulties in solving certain types of problems. For example, many children have difficulty solving equations in the form $A + B + C = A + \_\_$. McNeil and Alibali hypothesized that learners have trouble accurately encoding equations in this unfamiliar format [22]. Indeed, they found that learners who had difficulty solving these types of equations also showed poor encoding of the equations.

The current research addresses encoding and representation as they relate to arithmetic principle knowledge. More specifically, the present study investigates the influence of several factors in changing learners' encoding of arithmetic equations. Problem encoding is somewhat fluid and may change over the course of development [19], in the short term given the right experience [22], or with the development of expertise in a domain [23]. We hypothesize that these changes in encoding are the mechanism that leads to increased arithmetic principle knowledge.

Consider learners' encoding of simple arithmetic equations. The learner's encoding of the equations involves noting and prioritizing a set of features. The learner may note features such as the value of the operands, the operation, the order of the operands, specific characteristics of the operands such as evenness or oddness, and so forth. Some characteristics may be deemed very important while others are virtually ignored. In the case of the *Relationship to Operands* principle, the characteristic that is crucial to encode is the relative magnitude of the operands and the result. Learners whose encoding prioritizes the relative magnitude of the operands and the result should demonstrate knowledge of the *Relationship to Operands* principle. Learners whose encoding does not prioritize the relative magnitudes of the operands and the result should not show knowledge of *Relationship to Operands*.

In this research, we consider two characteristics of the learners' input that may affect encoding and principle learning: the number of principle violations and the temporal proximity of those violations.

## II. Research Questions

### A. Does the presence of principle violations affect learning?

Learning of regularities in most domains involves learning by positive examples only [3, 24]. That is, participants who learn regularities in these studies see only principle-consistent examples. Thus, to the learner, there is no direct information regarding what is inconsistent with the regularity being learned. Given that there is relatively little research on the process of arithmetic principle acquisition, there is very little work on how principle violations and principle-consistent experiences may contribute to arithmetic principle learning.

The current study investigates learning from a mixture of violations and principle-consistent evidence. The idea is that a mix of positive and negative evidence will highlight the relevant regularities for the learner. Contrasting stimuli that are consistent with a particular principle with stimuli that are inconsistent with the principle may help the learner encode the relevant regularity.

### B. Does the temporal proximity of the relevant input affect learning?

Previous investigations [18] have suggested that temporal proximity of relevant input facilitates learning of arithmetic principles. Examples that highlight a particular regularity may be more effective when presented to the learner in a blocked sequence. For example, a learner who receives three examples that highlight *Relationship to Operands* in a row may be better off than a learner who sees those same three examples interleaved between less informative input.

Two previous studies [21, 25] have investigated knowledge change though "blocked" vs. "mixed" practice with the arithmetic principles. In these studies, participants solved sets of three-term arithmetic problems in which the *Inversion* principle was either relevant (a + b – b) or irrelevant (a + b – c). Participants in the blocked learning group solved only relevant equations, while participants in the mixed group solved relevant and irrelevant equations.

Both studies concluded that participants who received blocked practice learned the principle better than participants who receive mixed practice. Although this may seem to support proximity of relevant input as important to learning, it is unclear whether the total amount of relevant input may also have been a factor. In both cases, blocked practice included twice the amount of relevant input as mixed practice.

One goal of the current study is to examine whether temporal proximity affects principle learning. Principle learning may be facilitated by input that highlights the regularity repeatedly within a small amount of time. Thus the temporal proximity of the relevant input may facilitate learning.

### C. Does equation encoding relate to principle knowledge?

We hypothesize that learners' encoding of arithmetic equations will relate to their knowledge of arithmetic principles. Specifically, in the case of the *Relationship to Operands* principle, encoding of the relative magnitudes of the operands and the result will relate to knowledge of the principle. The idea is that noting relative magnitudes as an important feature is required for understanding the *Relationship to Operands* principle. We hypothesize that learners who encode this feature of equations are more likely to have knowledge of the principle than learners who do not.

## III. Method

### A. Participants

Adult participants (*N* = 119) were recruited through a university course extra credit pool.

### B. Procedure

The focus of the study was on participants' acquisition of the *Relationship to Operands* principle for division. Participants took part individually in one experimental session that included six tasks, described below.

#### 1) Evaluation task – pre-training

The evaluation task was based on that used by Dixon et al. (2001) to assess knowledge of principles. Participants viewed sets of solved division equations, presented one at a time on sheets of paper. Each set consisted of nine equations presented in a 3 × 3 matrix. Participants were told that each set had been solved by "a hypothetical student who is learning arithmetic". Participants were also told that all of the equations were incorrect, but that they might believe that some students made better attempts at arithmetic than others. Participants were asked to rate each attempt on a scale from 1 to 7, with 1 indicating *very bad* and 7 indicating *pretty good*. This task was not timed.

#### 2) Training task

The training task involved serial presentation of stimuli on a computer screen similar to that used in artificial grammar learning studies [3]. Participants were instructed that they were to view a display of equations solved by two hypothetical students and to decide which student understood arithmetic better.

Stimuli included a mix of correct, incorrect non-violation and incorrect violation equations (all division). Equations were marked by color as correct (green) or incorrect (red). The proportion and placement of the violation equations depended on the condition to which the participant was assigned. There were two factors manipulated in a 2 x 2 factorial design: number of violation equations (high, low) and temporal proximity of violations (blocked, interleaved). Thus the four conditions were as follows: high-blocked, high-interleaved, low-blocked, and low-interleaved.

In addition to these conditions, there was also group of participants who viewed no violations at all. In this condition, all of the equations in the training stimuli were consistent with the arithmetic principle.

*3) Evaluation task – post-training*

Following the training, the evaluation task was presented again using a different set of stimuli.

*4) Verification task*

Participants viewed solved arithmetic equations one at a time on a computer monitor, and were asked to judge whether each equation was correct (e.g., "54 ÷ 6 = 9") or incorrect (e.g., "15 ÷ 5 = 2"). Presentation time was brief (1300 milliseconds) to make the task more challenging. Participants were asked to respond as quickly and accurately as they could, and both speed and accuracy were recorded. Incorrect trials included both principle violations and non-violations, because we intended to use the data to generate an additional measure of participants' knowledge of principles. However, there were no systematic differences in speed or accuracy on violation and non-violation trials and no systematic relations with performance on the other tasks. Therefore, data from this task will not be considered further.

*5) Encoding task*

Participants viewed equations presented briefly (650 ms) on a computer screen. After each equation the participant saw a series of four letters. The participant was then asked to answer a question about the letters (e.g., Was the first letter a *z*?) and then to answer a question about the equation. Questions about the equation were of four types: (a) identity (e.g., Was the first number 27?), (b) relationship (e.g., Was the first number bigger than the third?), (c) operation (e.g., Was the operation division?) or (d) parity (e.g., Was the third number odd?). The question types were mixed throughout the trials, thus, for any given trial the participant did not know what type of question they would be asked.

*6) Word problem task*

Participants were presented with ten multiple choice word problems, including 5 problems that involved division and 5 that involved multiplication. Each problem conveyed a story scenario such as "Mike bakes 56 cookies on 7 trays of the same size. How many cookies were on each tray?" For each problem, participants selected from a set of five equations the one that could be used to solve the story problem.

## IV. RESULTS

### *A. Do participants show knowledge of the principle before training?*

Participants' knowledge of the principle was inferred based on their performance on the evaluation task; they were inferred to have knowledge if they judged students who violated the principle more harshly than students who did not violate the principle.

To examine whether participants showed knowledge of the principle before training, we compared their average ratings of violation sets and non-violation sets on the pre-training evaluation task. Overall, participants rated violation sets lower than non-violation sets (1.99 vs. 2.32), $t(119) = 6.50$, $p < .01$. Thus, participants had some knowledge of the principle before training; however, the mean difference was quite small, and participants therefore had room to improve.

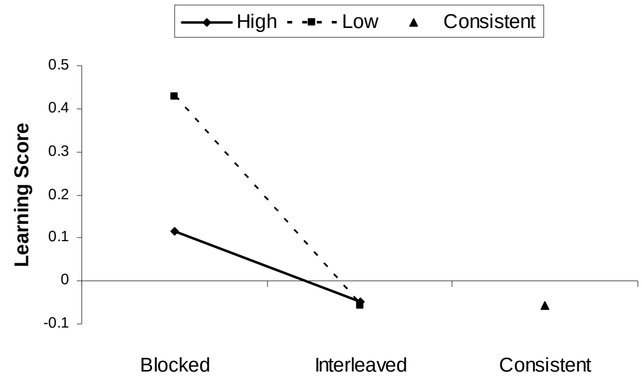


Figure 1. Learning scores for participants in each of the training conditions.

### *B. Does participants' principle knowledge change with training? If so, what factors is learning dependent on?*

For each participant a "learning" score was calculated. This was done by comparing the difference in ratings of non-violation and violation sets in the evaluation task pre and post training. For example, if a participant rated non-violations higher than violations by 0.05 on the pre-training evaluation task and by 0.70 on the post-training evaluation task, that participant's learning score would be 0.65. The data analysis compares learning scores across conditions.

Recall that participants viewed input that varied along two factors, number of violations (high vs. low), and proximity (blocked, interleaved). We conducted a 2 (number of violations: high or low) x 2 (proximity: blocked or interleaved) ANOVA with learning scores as the dependent measure. The data are presented in Figure 1. A significant effect of proximity was found, $F(1, 83) = 4.21$, $p = .04$. Participants in the blocked conditions had higher learning scores ($M = .269$) than participants in the interleaved conditions ($M = -0.052$). Number of violation equations did not significantly affect improvement scores, $F(1, 83) = .919$, $p = .34$. Participants in the high-amount conditions did not have significantly different learning scores from participants in the low-amount conditions (*M*s = 0.032 and 0.186, respectively). The interaction of the two factors was not significant, $F(1, 83) = 1.04$, p = .31. In sum, improvement on the evaluation task was dependent on proximity of the violations, not the amount of violations.

To examine whether principle violations facilitated learning more than principle-consistent examples, we performed a planned comparison between the blocked group and the principle consistent group. This comparison was marginally significant, $t(73) = 1.89$, p = .061.

We also examined whether number of violations affected learning within the blocked conditions. The difference in learning scores between the high-blocked and low-blocked conditions was not significant, $t(41) = 1.26$, $p = 0.21$.

### C. Is there a relationship between principle knowledge and equation encoding?

We hypothesized that there would be a relationship between learners' principle knowledge and their encoding of the relative magnitudes of the numbers in the equations. The most relevant comparison is between learners' post-training principle knowledge and their encoding scores. We calculated encoding scores for each of the four types of questions in the encoding task: *relationship*, *identity*, *parity* and *operation*. Based on participants' post-training principle knowledge scores, participants were divided into three equal size groups, with high, medium and low principle knowledge. For these groups, scores on the post-training evaluation task were 1.14, 0.32, and -0.29, respectively. For each of the four question types, we ran a one-way ANOVA and conducted post hoc comparisons.

For both *relationship* and *identity* encoding scores, the ANOVA was significant, $F(1, 116) = 3.66$, $p = .03$ and $F(1, 116) = 5.94$, $p = .003$, respectively. For both *relationship* and *identity* encoding scores, participants in the low principle knowledge group performed significantly more poorly than participants in the medium or high knowledge groups (see Figure 2). For *relationship* encoding scores the low knowledge group ($M = 0.71$) scored significantly lower than the medium group ($M = 0.81$, $p = .023$), though not significantly lower than the high group ($M = 0.78$, $p = .21$). For *identity* encoding scores, the low knowledge group ($M = 0.63$) scored significantly lower than the medium group ($M = 0.73$ $p = .05$) and the high group ($M = 0.78$, $p = .003$). It is worth noting that if the identity of the numbers is accurately encoded, the relative magnitude of the operands and the result can be inferred.

Neither *parity* nor *operation* encoding scores yielded significant results, $F(1, 116) = 0.384$, $p = .68$ and $F(1, 116) = 0.087$, $p = .91$. Thus, there were no significant differences between the principle knowledge groups in encoding the parity of the numbers or the operations used in the equations.

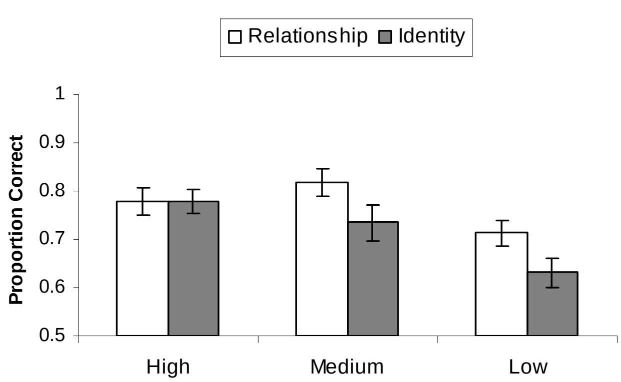


Figure 2. Relationship and Identity encoding scores for participants with high, medium and low principle knowledge post-training.

Note that the parity of the numbers is not relevant to the *Relationship to Operands* principle. Further, regardless of which operation is involved, the relative magnitudes of the operands and result are crucial to encode; thus, variations in encoding operation as a function of principle knowledge were also not expected.

These results suggest that learners with greater arithmetic principle knowledge encode specifically the relative magnitudes and the identities of the numbers in arithmetic equations better than learners with less principle knowledge.

### D. Are there consequences to having principle knowledge?

Because the training was in division only, we looked at participants' performance on the division items on the word problem task. We conducted a 2 (number of violations) x 2 (proximity) ANOVA with division sub-scores as the dependent measure. Neither factor was significant, proximity $F(1, 83) = 2.39$, $p = .12$, number of violations $F(1,83) = .37$, p = .54. Participants in the blocked conditions were correct on 78% of the division items while participants in the interleaved condition were correct on 71% of the items.

## V. Conclusion

We have shown that learners' arithmetic principle knowledge is affected by their experience. Specifically, increases in arithmetic principle knowledge are related to temporally proximal principle violations. Simple exposure to such violations led to gains in knowledge. This result is not what would be expected based on much of the implicit learning literature, which generally focuses on learning via principle-consistent examples only.

The current experiment extends prior research [18] by utilizing an experimental manipulation of exposure to equations, as opposed to a correlational approach. In addition, this work examines the relation between principle knowledge and equation encoding.

We found that training that included violations led to superior learning; however, it is unclear whether violations in and of themselves are beneficial. The training stimuli included a mix of violations, incorrect non-violations, and correct equations. Thus, learners always viewed violation equations in the context of principle-consistent equations. This allows for comparison between violations and non-violations, and it may be this comparison, rather than the presence of violations per se, that facilitates learning of the principle. Further, this comparison may be easier for learners when violations are presented in blocks. Learners must generalize to some degree across specific examples to note the regularities present, and it may be easier to generalize when the relevant examples are closely spaced. It should be noted that the equations in the training stimuli were marked as correct and incorrect, not as violations and non-violations. Thus, from the present data, it is not possible to determine whether comparing violations to incorrect non-violations or to correct equations is more effective.

We initially hypothesized that accurate equation encoding of specific problem features underlies arithmetic principle knowledge. The results suggest that learners with relatively low principle knowledge were also poor at encoding the relative magnitudes of the numbers in the equations and the identity of the numbers in the equations. However, the range of encoding scores was not great, making group differences difficult to discern. In future studies, it would be desirable to use a more challenging encoding task.

Although accurate equation encoding may be required for knowledge of the principle, it does not seem to be the case that accurate encoding automatically leads to principle knowledge. This is suggested by the fact that high principle knowledge participants did not encode the problems significantly more accurately than the medium principle knowledge group. It may be that accurate encoding is necessary, but not sufficient, for principle knowledge.

The results of this experiment suggest that accurate encoding of relevant problem features is a key aspect of learning the *Relationship to Operands* principle. However, future studies will be needed to test this claim directly. The general idea that changes in encoding lead to changes in knowledge may generalize to other arithmetic principles and to principles in other domains. However, the specifics of what information needs to be encoded and of how to facilitate changes in encoding will certainly vary.

Further research is progressing along several lines. We are pursuing additional behavioral and computational modeling work that aims to more fully characterize the relationship between equation encoding and principle knowledge. Finally, we are also conducting studies using a similar paradigm with children in classroom settings, to explore whether educational interventions that involve implicit learning might facilitate principle learning in classroom settings.

## References


[1] Gelman, R., & Gallistel, C. (1978). *The Child's Understanding of Number.* Cambridge, MA: Harvard University Press.

[2] Dixon, J. A., & Moore, C. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology, 32,* 241-253.

[3] Altmann, G., Dienes, Z., & Goode, A. (1995). Modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(4), 899-912.

[4] Knowlton, B. J., & Squire, L. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*(1), 169-181.

[5] Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321-324.

[6] Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old-infants. *Science,* **274**, 1926-1928.

[7] Gomez , R. & Gerken, L. (1999) Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition,* 70(2), 109-135.

[8] Marcus, G. (1999). Language acquisition in the absence of explicit negative evidence. *Cognition, 73,* 293-296.

[9] Altmann, G. (2002). Statistical learning in infants. *Proceedings of the National Academy of Sciences, 99*(24), 15250-15251.

[10] Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749-750.

[11] Starkey, P. (1992). The early development of numerical reasoning. *Cognition, 43,* 93-126.

[12] Dixon, J. A., Deets, J., & Bangert, A. (2001). The representations of the arithmetic operations include functional relationships. *Memory and Cognition, 29*(3), 462-477.

[13] Prather, R. W., & Alibali, M. (2008). Understanding and using principles of arithmetic: Operations involving negative numbers. *Cognitive Science, 32*(2), 445-457.

[14] Servan-Schreiber, E., & Anderson, J. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*(4), 592-608.

[15] Whittlesea, B., & Dorken, M. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General, 122*(2), 227-248.

[16] Gómez, R. L. & Schvaneveldt, R. W. (1994). What is learned from artificial grammars? Transfer tests of simple associative knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(2), 396-410.

[17] Tunney, R., & Altman, G. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*(3), 614-639.

[18] Dixon, J. A., & Bangert, A. (2005). From regularities to concepts: The development of children's understanding of a mathematical relation. *Cognitive Development, 20,* 65-86.

[19] Karmiloff-Smith, A. (1992) *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, Mass.: MIT Press/Bradford Books.

[20] Dixon, J. A. & Bangert, A. (2002) The prehistory of discovery: Precursors of representational change in solving gear system problems. *Developmental Psychology, 38,* 918-33.

[21] Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General, 127,* 377-397.

[22] McNeil, N. M. & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science, 28,* 451-466.

[23] Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.

[24] Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47,* 172-196.

[25] Stern, E. (1992). Spontaneous use of conceptual mathematical knowledge in elementary school children. *Contemporary Educational Psychology, 17,* 266-277.