Does Practice and Knowledge Equal Knowledge and Practice?

Miriam Rosenberg-Lee (mrosenbe@andrew.cmu.edu)
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213 USA

Marsha C. Lovett (lovett@cmu.edu)
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213 USA

Abstract

Debate surrounds the value of procedural practice in learning conceptual material in mathematical domains (Schoenfeld, 2004). We investigated whether purely procedural practice could lead to conceptual gains and explored cognitive load theory as a mechanism for those gains (Sweller, 1988). In a laboratory experiment, 93 undergraduates practiced a procedure by solving 30 problems of an algebra analog and were given conceptual tests before, during and after practice. The conceptual tests tapped students' understanding of the underlying structure of the domain. Participants' conceptual knowledge increased with procedural practice, particularly between no practice and some practice. Consistent with cognitive load theory performance on the conceptual test after practice was significantly related to procedural performance at the end of practice. However, this relationship between procedural learning and conceptual learning only held if participants had been alerted earlier in practice to the conceptual nature of the task. These results are consistent with the proposal by Rittle-Johnson, Siegler, & Alibali, (2001) that there is an iterative relationship between conceptual understanding and procedural skill.

Introduction

The role of practice in mathematics is a matter of intense controversy in the current curriculum reform debates. The National Council of Teachers of Mathematics' 1989 Curriculum and Evaluation Standards for School Mathematics launched the present debate by advocating for, among other reforms, the de-emphasis of rote practice and rote memorization of rules and algorithms (Schoenfeld, 2004). Consistent with this approach were the opinions of education researchers such as Robert Davis who wrote in 1986:

> If "mathematics" is seen as conformity to memorized rituals, if it is taught without meaning … if meaninglessness compels a slow pace and a vast investment in repetition, and if routine calculation is the main goal, very little mathematics will be included in the curriculum (pp. 272-273).

However, this shift of emphasis did not seem warranted to all parents and an anti-reform movement developed, advocating for back to basics. The 'traditionalists' actively sought mathematicians to support their side, having them send letters to state decision makers (Schoenfeld, 2004). David Ross, a mathematician at Kodak Research Labs states: "The best way to advance students' conceptual thinking about mathematics is to have them master the traditional algorithms" (2001, ¶13). The traditionalists contend that practice is a necessary prerequisite without which deep understanding of mathematical concepts is impossible. Conversely, reformers fear that emphasizing practice can only lead to superficial understanding of mathematical concepts.

Clearly with such opposing perspectives, consensus may be difficult to reach. Even before the debate reached its current level, Schoenfeld (1994) suggested that these issues must be settled through empirical research. He argued that a particular question worth investigation is: "how much mastery of some basics is required for competent, flexible performance on more demanding tasks." (¶ 19) That phrasing of the question highlights the underlying goal of both sides of this debate, namely, competent flexible performance on 'demanding' tasks.

Competent and flexible performance requires two different kinds of proficiency: competent performance implies that calculations are fast and effortless; whereas, flexible performance suggests not just solving familiar problems but also being able to tackle novel problems built on the same principles. These two proficiencies can be likened to the distinction between procedural and conceptual knowledge. According to Rittle-Johnson, Siegler, & Alibali, (2001), procedural knowledge is "the ability to execute action sequences to solve problems" (p. 346) and conceptual knowledge is "an implicit or explicit understanding of the principles that govern a domain and of the interrelations between units of knowledge in a domain" (p. 346-7). For example, in learning multiplication, a student could understand the concept that multiplication is equivalent to multiple additions, and this knowledge could be independent of knowing the procedure "add number A to itself B times." Thus, the debate can be reformulated as: does procedural practice result only in procedural improvements, or can it also lead to conceptual learning?

If procedural practice can lead to conceptual learning, what mechanism(s) would be at work? Evidence from the worked example literature suggests that actually working through sample solutions and solving problems, as opposed to reading declarative instructions, is instrumental in 'understanding' a domain (Chi, & Bassok, 1989, Schworm, & Renkl, 2002). In this view the benefit from practice comes from being able to *map* the declarative instructions to the various steps of the problems. This perspective predicts that some practice is better than none, but doesn't address the effects of increasing practice, nor the quality of that practice on conceptual learning.

Cognitive load theory provides a mechanism that connects amount and quality of procedural practice to conceptual knowledge acquisition. Cognitive load theory, developed by Sweller, (1988) proposes that at the beginning of learning at task, participants will have little cognitive capacity for learning the conceptual aspects (schema, in his words) of a domain, because that capacity will be engaged in executing the procedure necessary to solve the problem. However, procedural practice results in the automatization of the problem-solving procedure, thus freeing up cognitive resources for conceptual learning. If this *mechanism* were at work, we would expect to see a connection between the development of automaticity and conceptual learning.

While we are primarily concerned with the effects of procedural practice on conceptual learning, much research has examined the opposite direction, that is, conceptual learning on procedural performance (cf. Anderson, 1993). Similarly, there has been considerable debate about which type of knowledge is first to emerge during development. However, Rittle- Johnson et al., point out that the pathway need not be unidirectional. Instead they propose that learning results from the iterative effects of conceptual on procedural and procedural on conceptual. In this research, we hope to tease apart the effects of procedural practice on conceptual learning, but also to look at the combination of conceptual knowledge and procedural practice on later conceptual learning

The task used to investigate these questions is a variant of the Blessing and Anderson (1996) task known as Symbol Fun. Symbol Fun is an analog of algebra where the operators and operands are replaced by symbols (i.e. ® for + and # for *). Participants are presented with a series of rules that can be applied to given strings, which, when applied correctly will result in isolating 'x' on the left-hand side, or in algebraic terms, solving for 'x'. The procedural aspects of this task consist of correctly applying the rules to isolate 'x' on the left-hand side. Importantly, the rules are presented so that the task appears to be pure symbol manipulation. However, given that the rules conform to those of algebra, there is considerable conceptual material here: the mapping of the symbols to their algebraic counterparts, the goal structure of solving for 'x', the inverse relationship between pairs of operators, i.e., addition and subtraction and multiplication and division. Thus, our task has clearly defined conceptual content, but can be performed without reference to that material. This organization allows us to examine what, if any, of the conceptual material students will learn given that they are only required to produce the correct set of symbol-manipulation steps. That is, we can ask: can purely procedural practice lead to the kind of conceptual understanding educators seek? Moreover, by examining the relationships between the procedural and conceptual measures we can investigate possible mechanisms for one leading the other.

Methods

Participants

Ninety-three Carnegie Mellon University undergraduates (mean age = 19.4, 53 female) were given course credit towards a research requirement for participating in this experiment.

Procedure

Participants in this task received an introduction to the rules of Symbol Fun by computer. Then they completed 30 trials of computerized practice on Symbol Fun, followed by a conceptual test. Participants also completed a demographic information form which included math SAT score and number of math courses taken at the high school and college level. We also collected responses to a Need For Cognition (Cacioppo, Petty, Feinstein, & Jarvis, 1996) questionnaire. We were primarily interested in the effects of different amounts of procedural practice on conceptual learning, so we divided the participants into four groups and varied the amount of practice they received prior to taking the conceptual test for the first time. Thus, Group 1 was tested before and after all training, whereas Group 2 was tested after 10 trials and again after the training was complete. Group 3 was tested after 20 trials of practice and again at the end of training. Finally, Group 4 was test after 30 trials of training; however, this was the end of practice, so this group only completed the test once. The complete experimental design can be found in Table 1.

Materials

Problem-Solving Training The algebra analog, Symbol Fun, used in this experiment was only slightly modified from the task described in Blessing & Anderson (1996). In Symbol Fun the operators and operands of algebra are replaced by symbols (see Table 2). For example, the equation:

$$- x + B = D$$

becomes:

£ ρ ® ♪ ↔ ¥

The goal of the task is to isolate the ρ on the left-hand side, (i.e. solve for 'x'). There are three major operations in Symbol Fun which could be used to accomplish the goal: (a) adding an operand-operator pair to both sides, (b) canceling two operator-operand pairs when the operand was the same and the operators were inverses and (c) eliminating an operator from in front of the variable.

Table 1: Experimental Design

	Test 1	Test 2	Demographics/NFC
Group1	0 Trials	30 Trials	After Test 2
Group2	10 Trials	30 Trials	After Test 2
Group3	20 Trials	30 Trials	After Test 2
Group4	30 Trials	-	After Test 1

These three operations correspond to the nine rules of Symbol Fun. Rule 1 specifies that any operator-operand pair can be added to both sides of the ↔ symbol. Rules 2 - 5 govern the cancellation of operator-operand-operator-operand sequences, two for each of the operator inverse pairs (addition-subtraction and division-multiplication). Finally rules 6 – 9 describe the elimination of each of the operators in front of the ρ symbol. The participants are introduced to the nine rules from a computer interface with one rule per screen. The screens are similar to one another, with a schematic of how the rule applies, a short test description of the rule, and an example of an application of that rule.

Symbol Fun is not a perfect match to algebra; for example, the division-multiplication operator pair acts more like the addition-subtraction operator pair than in standard algebra, thus you could have an equation like: /x = *A + B. And there is an order of operations so that one can't remove a symbol before 'x' until all the constants have been removed from the left-hand side. However, unlike the Blessing and Anderson (1996) version, more than one rule can apply at a time.

In the computer display the current problem appeared in the upper left-hand corner. In the lower left-hand side each of the nine rules were displayed and the corresponding "Go" and "See Full Rule" buttons were on the lower right-hand side of the screen. Pull down menus within the rules allowed participants to insert symbols into their chosen rule or to specify operations to the right-hand side of the symbol strings. If participants selected an applicable rule, using the correct pull down(s), the computer provided a green "Good" for feedback and output the result of their rule. If they selected either the wrong rule or the right rule but the wrong pull down(s) they received a red "Try again". If they got the step wrong a second time the computer instructed them on an applicable rule they could have used and provided the result of that rule. A correct selection for the last step was signaled by a green "Excellent".

Procedural practice consisted of 30 trials of Symbol Fun grouped in three blocks of ten trials. Within each block, there were two 1-step problems and 4 each of 2- and 3-step problems. The problems were randomized within each block, and the constants used in each problem were randomly generated.

Conceptual Testing A major difference from the Blessing and Anderson (1996) procedure and the one used here is the addition of an assessment of learners' conceptual knowledge about the domain. The extent of this knowledge was operationalized by a four item test. The *categorization* question asked participants to group the symbols and label the groups. This question measured their understanding of the functional differences between operands and operators. The *valid expressions* question asked whether a novel string of symbols was admissible in the domain: two were admissible, but were ordered in an unusual way, the other two directly violated the role of operators or operands. The *order of operations* question asked participants to identify legal starting moves; one move was not legal because it could only be completed after the others had been completed. The *inverse operators* question asked participants to generalize the function of novel operators. The question demonstrate how the novel operators could be cancelled in one order; to solve the problem the operators had to be added and cancelled in the opposite order.

Two isomorphic versions of the test were used, and the order in which participants received them was counterbalanced. Whenever participants saw the test for the second time, the experimenter pointed out the similarity and assured them that same or different responses were acceptable. Participants had ten minutes to work on the conceptual test and were warned when they had a minute left.

Table 2: Mappings of Algebra to Symbol Fun

Algebraic Symbol	**Task Symbol**
+	®
-	£
*	#
/	©
=	↔
Constant	@
Constant	♪
Constant	♥
Constant	¥
Variable	ρ

Results and Discussion

Procedural Performance

Not surprisingly, with practice, participants get faster and more accurate at Symbol Fun. Figure 1 displays reaction time data for the first step of each trial. First steps were used to ensure a homogenous sample, as trials differed in the number of steps, and participants received them in a random order. Decreases in reaction times are consistent with the Power Law of Practice (Newell & Rosenbloom, 1981) in that the data is better fit by a power function (R^2 = .905) than a linear function (R^2 = .360) or an exponential function (R^2 = 0.697). Recall that the experiment is organized into blocks of 10 trials with a subset of the participants stopping between blocks to complete the conceptual test. Participants tended to be much slower on their first trial back from the tests, so those trials were removed from the sample (see Figure 1, Trials 11 and 21). To further ensure a homogeneous sample, only correct trials were used for the remaining analyses. A repeated measures general linear model analysis revealed that using this sample there was a significant speed up in performance ($F(2, 83) = 36.694$, $p < .001$), but there were no differences between the groups ($F(3, 83) = .974$, $p = .409$) nor an interaction between group and trial in practice ($F(6, 83) = .310$, $p = .931$).

Conceptual Performance

Practice also leads to improvements in conceptual knowledge; however, the pattern of results is somewhat more complicated than for the procedural data. Figure 2 represents participants' performance on the first and second conceptual test, divided up by group. Recall that the groups differed in when they received the conceptual test for the first time (see Table 1). The Test 1 line (solid) represents the score on the conceptual test with increasing practice. Qualitatively, we see that conceptual performance increases with practice, but with diminishing returns. An ANOVA confirms this observation, in that the groups are significantly different ($F(89,2) = 17.002$, $p < 0.001$); however, this difference is driven by Group 1 in that the remaining groups are not significantly different from each other ($F(2,69) = 2.212$, $p = .117$).

The dashed line in Figure 2 represents performance by groups 1, 2 & 3 on Test 2. These groups of participants all had thirty trials of practice before they saw this second test, but they differ in when they saw the conceptual test for the first time. Seeing Test 1 at different points doesn't appear to affect performance on Test 2 ($F(68,2) = 1.25$, $p =.303$).

Participants clearly improve from Test 1 to Test 2 (paired-$t(68) = -5.300$, $p<0.001$). However, again this effect is driven by performance of Group 1 (paired-$t(22) = -7.443$ $p<0.001$) and to a lesser extent Group2 (paired-$t(22) = -2.117$, $p = 0.046$), but not Group 3 (paired-$t(22) = -.703$, $p = .490$). This may seem somewhat contradictory to the result mentioned above, that performance on Test 2 does not depend on group. However, we are looking at the change in performance between tests. Since the groups are starting at different levels of test performance, but ending at the same level we would expect some difference in their improvement. Indeed, combining all these variables in a repeated measures general linear model, we see that there is a significant effect of Test 1 vs Test 2 ($F(68, 1) = 42.565$, $p<0.001$), and of group (i.e., how much practice they had when they took their first test) ($F(68, 2) = 4.062$, $p=0.022$). Most importantly, the interaction between group and test number is also significant, ($F(68,2) = 18.527$, $p < .001$) confirming that practice before Test 1 matters for Test 1, but not for Test 2.

Interestingly, performance on Test 1 and Test 2 are not correlated ($r(69) = .168$, $p = .167$) suggesting that these two scores represent different measures of conceptual performance. However, there does seem to be an increasing relationship between the tests as they are taken closer together. Thus for Group 1, when the tests are taken 30 trials apart, their correlation is not significant($r(23) = .08$, $p = .716$). For Group 2.where the tests are 20 trials apart, the correlation is higher but still not significant $r(23) = 0.343$ $p = .109$). Finally for group 3, where the tests are only 10 trials apart, the two tests are significantly correlated ($r(23) = .522$, $p = .006$).

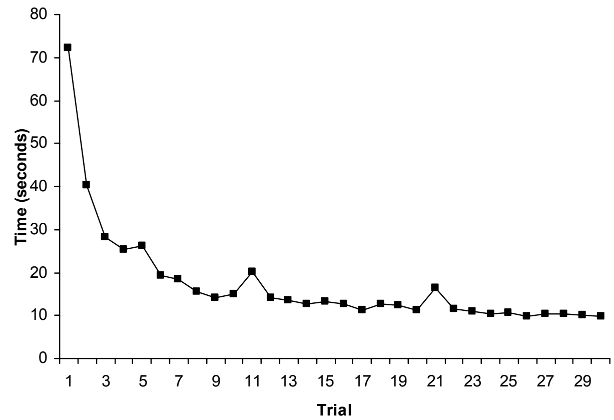

Figure 1: First step Reaction times for all participants

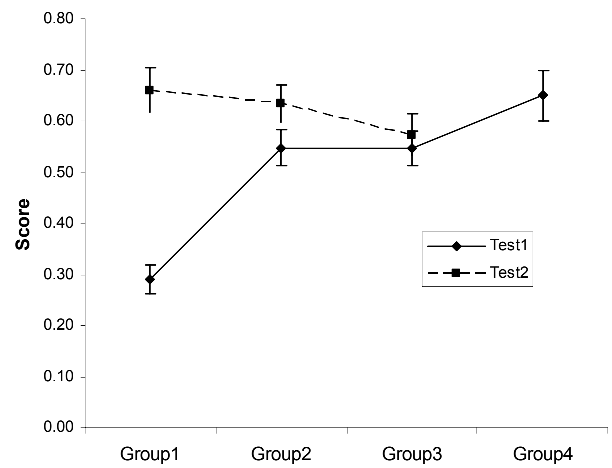

Figure 2: Conceptual performance by group

Connecting Procedural and Conceptual

We have seen that practice leads to improvements in both conceptual and procedural performance, but what is the relationship between these two forms of learning? In order to assess this question we performed regression analyses on the two dependent/outcome measures of conceptual learning (Tests 1 and 2) to determine which measures of procedural performance best account for these outcome measures. The six procedural measures used as independent/predictor variables included accuracy and reaction times for the three blocks. Accuracy and reaction times do not correlate within blocks, suggesting they measure different aspects of procedural proficiency (Block 1, $r(87) = -.180$, $p =.091$, Block 2, $r(89) = -.109$, $p = .304$m, Block 3, $r(90) = -.093$, p

= .376). We also included two between-subject factors that represent the effects of practice. The first factor coded practice *incrementally*, namely, the time participants saw Test 1 (i.e. 0, 10, 20, 30 trials of practice). The second factor coded practice *dichotomously*, that is, whether the participant had had no or some procedural practice before taking the first conceptual test (i.e. Group 1 vs. the remaining groups). Finally we had three demographic variables: math SAT score, Need For Cognition (NFC) score, and number of math courses at the high school and university level. Need For Cognition is a measure of an individual's desire to makes sense of situations. Unfortunately we did not have SAT information for seven participants, either because they hadn't taken the test, or couldn't remember their score. These participants were excluded from these analyses. Nevertheless, results from analyses of all participants' data excluding the math SAT variable are consistent with the results presented here.

Table 3 lists the model summaries of these analyses. Consistent with the results above, there was a strong effect for having had any practice before taking Test 1; however, in addition to that, both Block 1 accuracy and NFC cognition were key predictors of Test 1 performance. In contrast, the main predictor of Test 2 performance was reaction time during Block 3, followed by Block 3 accuracy, and finally, having had practice before seeing Test 1. That is, performance on the first test, when the task had seemed purely procedural, depended primarily on having had some exposure to the task, then on how well (i.e. accurately) participants performed the task, and finally on an individual difference variable, namely, one's desire to make sense of the task. Yet for Test 2 performance, after participants had been oriented by Test 1 to the fact that the task has a conceptual structure, the most important factor in their performance was reaction times at the end of practice, just before they took Test 2. That is, participants' later procedural fluency (not just their task accuracy) is a predictor of their later conceptual performance. We can relate these results to cognitive load theory which suggests that practice leads to automaticity and the freeing of cognitive resources which can then be used for conceptual learning. Although we do not have a direct measure of cognitive load, we can regard Block 3 reaction times as a proxy for automaticity. Thus, consistent with cognitive load theory we find that participants who are more automatic procedurally do better conceptually at Test 2.

Table 3: Regression Summaries

Variable	St. Beta	t	Sig
Test 1			
Practice (dichotomous)	.627	7.760	<.001
Block 1 Accuracy	.311	3.668	<.001
NFC	-.196	-2.319	.023
Test 2			
Block 3 Reaction Times	-.429	-3.602	.001
Block 3 Accuracy	.295	2.576	.013
Practice (dichotomous)	-.305	-2.563	.013

We have seen that Test 1 and Test 2 performance seem to be related to different factors. However Test 1 and Test 2 differ in two ways: the presence of the earlier test (which could orient participants to the conceptual structure of the task) and, on average, the amount of procedural practice at test time. To address which of these is more responsible for the Test 1 vs. Test 2 differences in predictors of conceptual learning, we compare Group 1 at Test 2 and Group 4 at Test 1. These are the two groups that performed all the practice trials without interruption. They differ only in that Group 1 has already seen the conceptual test and Group 4 has not. Moreover, there are no differences between these two groups' scores on their conceptual test after practice ($t(45) = .155$, $p = .878$). Procedurally, there are no differences between these two groups on their accuracy throughout the task. Generally, Group 4 tends to be faster than Group 1 although the difference is only significant on the last block ($t(45) = 2.564$, $p = .014$).

We performed regression analyses on these two conceptual test scores with our six procedural measures as independent variables. For Group 1 (Test 2), the only factor selected was Block 3 reaction times, ($R^2 = .451$), whereas for Group 4 the only factor selected was Block 1 accuracy ($R^2 = .661$). Thus, for these two groups, the key difference of having seen a prior conceptual test (or not) is enough to produce different predictors of the conceptual test scores at the end of training. That is, for Group 1 participants who have seen the test before, performance is related to procedural proficiency at the end of practice, whereas for Group 4 participants, taking the test for the first time, their conceptual score are only related to initial accuracy. This suggests that the first conceptual test may be a conceptual intervention of sorts and that it provides conceptual knowledge that changes the relationship between procedural skill and conceptual learning. This is akin to Rittle-Johnson et al.'s (2001) view that there is an iterative relationship between procedural skill and conceptual learning. Here, we have evidence of some (very minor) conceptual training (i.e., simply taking the first test) changing the conceptual impact of additional procedural practice.

These results suggest that cognitive load can explain gains in conceptual learning from procedural performance, but that this relationship requires some awareness of the conceptual nature of the task. However, we have not manipulated this relationship directly. Particularly, we have assumed that faster reaction times imply reduced cognitive load. A dual task paradigm where improvements in Symbol Fun were offset by increased cognitive load from the other task would demonstrate whether reducing cognitive load is integral to this type of learning. Interestingly, while we have seen that there is a differential relationship between procedural practice and conceptual learning with a small conceptual intervention, there was no difference in conceptual performance. An outstanding question is whether this shift in role for procedural performance has any educational benefits.

Acknowledgments

This research was supported by Natural Sciences and Engineering Research Council of Canada PGS D3.

References

Anderson, J. R. (1993). *Rules of the Mind.* Hillsdale, NJ: Erlbaum.

Blessing, S. & Anderson, J. R. (1996). How people learn to skip steps. Journal of Experimental Psychology: *Learning, Memory and Cognition,* 22, 576-598.

Davis, Robert B. (1986). Conceptual and procedural knowledge in mathematics: A summary analysis. In: Hiebert, James (Ed.), *Conceptual and procedural knowledge: The case of mathematics.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cacioppo, J., Petty, R., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bullletin*, 119, 197-253.

Chi, M. T. H., & Bassok, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.

National Council of Teachers of Mathematics: Commission on Standards for School Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: The Council**.**

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition*. Hillsdale, NJ: Erlbaum.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, 93, 346-362.

Ross (1994). *Math Wars*. Retrieved July 2nd 2004 from http://www.ios.org/articles/dross_math-wars.asp

Schoenfeld (1994). *What Do We Know about Mathematics Curricula?* Retrieved July 2nd 2004 from http://www-gse.berkeley.edu/faculty/aschoenfeld/WhatDoWeKnow/What_do_we_Know.html

Schoenfeld (2004). *Math Wars*. Retrieved July 2nd 2004 from http://www-gse.berkeley.edu/faculty/aschoenfeld/Math_Wars.pdf

Schworm, S. & Renkl, A. (2002). Learning by solved example problems: Instructional explanations reduce self-explanation activity. In W. D. Gray & C. D. Schunn (Eds.), *Proceeding of the 24th Annual Conference of the Cognitive Science Society* Mahwah, NJ: Erlbaum.

Sweller, J., Cognitive load during problem solving: Effects on learning, *Cognitive Science,* 12, 257-285 (1988).