Title

Subgoal Learning and the Effect of Conceptual vs. Computational Equations on Transfer

Permalink

https://escholarship.org/uc/item/4401k2vn

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 22(22)

Authors

Atkinson, Robert K.
Catrambone, Richard

Publication Date

2000

Peer reviewed

Subgoal Learning and the Effect of Conceptual vs. Computational Equations on Transfer

Robert K. Atkinson (atkinson@ra.msstate.edu)
Department Counselor Education and Educational Psychology; Box 9727;
Mississippi State, MS 39762, USA

Richard Catrambone (rc7@prism.gatech.edu)
School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332-0170, USA

Abstract

Subgoal learning is examined through the use of equations that are designed to encourage a conceptual rather than computational approach to solving problems (conducting statistical tests). Learners who studied conceptually-oriented examples transferred more successfully to novel problems compared to learners who studied computationally-oriented examples. These results extend prior work on subgoal learning by demonstrating another technique for aiding subgoal learning.

Introduction

Research suggests that learners typically struggle when they are obligated to solve problems that have different procedural requirements than those demonstrated by training problems or worked-out examples, even if those differences are relatively slight (e.g., Catrambone, 1995, 1996, 1998; Novick & Holyoak, 1991; Reed, Dempster, & Ettinger, 1985). This difficulty may stem in part from the fact that learners often represent the problem solving procedures of training problems or worked-out examples as a set of linear steps rather than forming a hierarchical representation that could permit them to successfully solve novel problems (Dufresne, Gerace, Hardiman, & Mestre, 1992; Singley & Anderson, 1989)

Educators and researchers alike are concerned with this problem. In fact, the Committee on Developments in the Science of Learning (1999) recently suggested that "a major goal of schooling is to prepare students for flexible adaptation to new problems and settings [and that] students' abilities to transfer what they have learned to new situations provides an important index of adaptive, flexible, learning" (pp. 223). Research indicates, however, that this goal is rarely achieved (Chi, Feltovich, & Glaser, 1981; Larkin, McDermott, Simon, & Simon, 1980).

Presumably, emphasizing the structure of an example through instruction will increase flexible transfer by helping the learner look beyond the surface features of the example and test problem to find the goal-related features that can be used to solve the problem. Thus, instead of committing to memory the details of equations as the basis for one's problem solving knowledge, a more productive approach would be to organize this knowledge in such a way that it could support generalizations across problems in a domain. One type of knowledge structure that appears to offer the promise of enhancing this type of procedural generalization is one organized around subgoals.

Subgoal-Oriented Instruction

As used in the present paper, a subgoal denotes a meaningful conceptual piece of an overall solution procedure. Subgoals are particularly useful to learners because they can assist them in solving novel problems since problems within a domain often share a common set of subgoals, albeit the steps for achieving the subgoals vary from problem to problem within a domain. Once learners become familiar with the typical subgoals in a domain, this knowledge can assist them in identifying which part of a previously-learned solution procedure needs to be modified in order to solve a novel problem (Catrambone, 1996, 1998).

Recently, a line of research has emerged examining the efficacy of subgoal-oriented instruction (Catrambone, 1995, 1996, 1998). In particular, this line of research has explored several techniques for designing examples that help learners to form subgoals to represent the purpose of steps in an example's solution. Across a series of studies, Catrambone investigated the impact of making the goal structure of an example's solution explicit by using manipulations such as the use of solution step labels or visually isolating parts of example solutions. These studies indicated that if examples are designed in such a way as to encourage subgoal learning, then learners are more likely to correctly solve new problems that involve the same subgoals but require new steps for achieving them.

These studies also suggest that example solutions that are segregated or labeled encourage learners to self-explain how the steps go together. One result of his self-explanation process is the formation of subgoals (Catrambone, 1998). This work parallels research in the text-comprehension literature on the effects of signals

or cues on text-processing strategies (e.g., Lorch & Lorch, 1995; Meyer & Rice, 1989). Just as organizational signals in text induce learners to change their text-processing strategy by cueing the important text content and its organizational structure, worked-example labels are intended to increase the likelihood that learners will discern the hierarchical conceptual structure of the problem contained in the example.

Factors that May Influence Subgoal Formation

As previously mentioned, several structural manipulations have been found to successfully make the goal structure of a problem's solution explicit, such as by the use of labels or visual isolation. But, there might be other factors that influence subgoal information. For instance, two potential factors are the nature of equations used in examples and the presence of conceptual elaborations.

Conceptual vs. Computational Equations. The process of calculating sum of squared deviation scores or sums of squares (SS) for the variance terms in t-tests and analyses of variance (ANOVAs) can involve two noticeably distinct types of formulas: conceptual and computational. According to Gravetter and Wallanu (2000), the conceptual formula is useful "because the terms in the formula literally define the process of adding up the squared deviations" (p. 121). For instance, the conceptual formula for SS in a t-test, $\sum(X - \overline{X})^2$, translates directly into the sum of ($\sum$) squared deviations $(X - \overline{X})^2$. This clearly captures how the variance term measures the amount of spread about the mean.

In contrast, the computational formula for SS, $\sum X^2 - \frac{(\sum X)^2}{N}$, permits the learner to calculate SS directly from raw scores which can lead to more efficient calculations. However, there is a notable drawback to this convenience: the computational formula conceals the true meaning behind SS. Unlike the conceptual formula, a learner cannot directly translate the terms in the computational formula into a sum of squared deviation scores. As a result, the leaner may not grasp that this formula is designed to measure the amount of spread about the mean.

On the one hand, the computational approach might aid performance on problems that are just like the examples that illustrated the approach, but might make far transfer difficult. That is, in the computational approach, the equation is streamlined for doing the calculations, but "hides" what is really going on. On the other hand, the conceptual approach, although typically more cumbersome computationally, clearly shows how the variance is related to the difference of each mean from the grand mean. Therefore the conceptual approach might aid far transfer by making it easier for the learner to determine how to adapt relevant parts of the procedure.

Thus, we hypothesize that conceptually-oriented equations will be more effective than computationally-oriented equations at helping learners acquire knowledge structured around the goal-related features of the problems they study and this translates to superior far transfer performance.

Conceptual Elaborations. Another factor that appears to have the potential to influence subgoal formation is the use of elaborations in example-based instruction and, in particular, conceptual elaborations. The literature contains examples of several types of elaborations that vary in the degree to which they elaborate the problem at hand. They range from elaborations involving problem solutions (Lovett, 1992) to those that focus on rules and procedures (Catrambone, 1996; Reed & Bolstad, 1991; Reed et al., 1985).

To date, the success of these various elaborations has been mixed. Although Lovett (1992) found that far transfer was facilitated by elaborated solutions, Reed and his colleagues (Reed & Bolstad, 1991; Reed et al., 1985) have found virtually no evidence to suggest that rule-based instructional elaborations—those that elaborate on the purpose and appropriateness of applying a rule or procedure in a given problem-solving context—are beneficial to learners.

In one study, Catrambone (1996) examined the relative benefits of rule-based instructional elaborations versus subgoal labels. In this study, Catrambone manipulated two factors: subgoal labels (present or absent) and rule-based elaborations (present or absent), where the elaborations consisted of supplemental material describing an alternate representation or equation that could be used to solve the problems the participants were studying. He found that the labeling manipulation enhanced transfer while the rule-based elaboration manipulation did not.

The rule-based elaborations used in the Catrambone (1996) study, however, offered "what to do" knowledge not "what it means" knowledge. This distinction is important in light of research suggesting that rules conveying "what to do" knowledge might provide little help to learners for developing a deep understanding of the rule-based system they are studying whereas knowledge about "what it means" may facilitate this depth of understanding (Riesbeck & Schank, 1989). For instance, an elaboration that describes what is meant by the term "variance" (see Appendix for an example) might be more effective than one dedicated to elaborating the procedural aspect of the variance formula.

In sum, the impact of conceptual elaborations containing "what it means" knowledge in the context of subgoal-oriented instruction remains an open question.

Overview of Study

The aim of the study was to compare the effectiveness of conceptual and computational equations, and the use of elaboration, on performance. Performance was assessed in two ways: the time spent studying the training examples and correctness of solutions on near and far transfer problems.

Experiment

Method

Participants and Design. Participants were 215 students drawn from several educational psychology courses at a small, northeastern college who participated in the experiment for course credit. The participants were randomly assigned to one cell of a 2 x 2 x 2 factorial design. The first factor was the characteristics of the variance formulas (conceptual or computational) in the t-test example, the second was the characteristics of the variance formulas (conceptual or computational) in the ANOVA example, and the third was conceptual elaboration (elaboration or no elaboration) in the examples, described below.

Training Phase. Participants received an instructional booklet containing a general overview of statistical hypothesis tests and two training examples, one representing a t-test and another representing the use of an ANOVA for the same 2-group comparison. The introduction to statistical hypothesis tests described the utility of these procedures and provided an overview of the four-step hypothesis testing process common to both tests. Each training example was preceded by an overview of the test that it exemplified. This explanation described the purpose of the test without going into detail regarding how to perform the test's calculations.

Half of the participants were exposed to examples that contained conceptual elaborations designed to provide "what it means" knowledge. That is, they were designed to describe the conceptual meaning behind the various formulas used in the two hypothesis tests. The other half of the participants studied examples in which the elaborations were not present.

With respect to the t-test example, the variance formulas were either conceptual or computational in nature. Similarly, with regard to the ANOVA example, the variance formulas were either conceptual or computational in character.

Regardless of the instructional manipulations, the examples contained a number of invariant structural features. First, all of the equations used across both tests were converted to their verbal equivalents so that they were devoid of any statistical notations. Second, each of the six calculational subgoals in the two examples was either labeled or visually isolated.

The t-test subgoals were to find: sample mean for group 1, variance for group 1, sample mean for group 2, variance for group 2, pooled variance, and t-statistic. The ANOVA subgoals were to find/do: preliminary calculations, sum of squares between, sum of square within, mean squares between, mean squares within, and f-value.

The Appendix shows samples of the materials from the examples.

Test Phase. The test booklet contained three test problems for the participants to solve. The first test problem required the participant to apply a t-test. The second problem required them to apply an ANOVA to a 2-group situation and the third problem asked them to apply an ANOVA to novel situation involving three groups. Thus, the first two problems were near transfer while the third problem involved more far transfer in that it required the learner to adapt the equations for variance. The extension is a more straightforward, modular extension of the conceptual equations. However, the extension is less straightforward in the computational equations since it involves changes to the "interior" of the equations.

The test booklet also included two sheets that participants could refer to, one containing the condition-specific formulas for the t-test (conceptual or computational) and the other containing the condition-specific formulas for the ANOVA. Although these sheets represented the formulas in the sequence in which they were applied in the training examples they did not contain any of the values from those examples.

A binary scoring system was developed to score the problem-solving protocols. This system was designed to award participants with points for the accuracy with which they achieved each subgoal. The three test problems each contained six calculational subgoals. The correct numerical answer to the subgoal was awarded one point. For example, the correct answer to the second subgoal in the t-test problem, correct group 1 variance, was 30.2. If a participants' problem-solving protocol contained this answer, he/she was given a point.

Since most subgoals contained subcomponents, the binary system allowed us to award partial credit. This permitted us to capture the proportion of the subgoal's solution—for those participants who did not have the correct numerical answer for the subgoal—that was correct. For instance, the equation associated with the second subgoal (i.e., correct group 1 variance) in the conceptual condition was coded for the presence or absence of seven components, ranging from whether each value was present in the formula to whether the equation had the correct denominator. In this example, if a participant's problem-solving protocol had six of the seven components, he/she was awarded a .86 for the

subgoal. If the subgoal was correct except for a trivial math error, the participant received full credit (one point) for that particular subgoal.

Procedure. Participants were asked to study carefully the instructional booklet containing the training examples since after studying it they would be asked to solve several problems. They recorded the amount of time they spent studying each example. The participants were informed that they would not be able to refer to any of the examples while solving the problems but that they would have a copy of the formulas. This constraint was designed to increase the likelihood that participants would focus their attention on studying the examples and how they were solved.

Participants were run in groups ranging in size from 5 to 30 participants. Participants worked for approximately 75 minutes and were asked to show all their work.

Results

To validate the scoring system that was developed, two raters independently scored a random sample of 10% of the problem-solving protocols and agreed on scoring 98% of the time. Disagreements were resolved by discussion. One experimenter independently scored the remaining problem-solving protocols.

A 2 x 2 x 2 analysis of variance was initially conducted on the study times for the two examples (i.e., t-test and 2-group ANOVA) and the correctness measures for the three test problems, using elaboration, type of t-test formulas, and type of ANOVA formulas as grouping factors. There was no systematic effect of elaboration on correctness and so, in the interest of clarity and brevity, this factor will not be discussed below in the context of correctness. Table 1 presents the mean scores for each condition on the correctness measures for the three test problems.

Training Times for T-Test Example: There was a significant main effect of elaboration, $F (1, 207) = 10.11$, $MSE = 10.8$, p < .01, which indicated that the participants presented with the elaborated material ($M = 8.11$ min.) spent more time studying the examples compared to participants who studied unelaborated materials ($M = 6.73$ min.). There were no other significant main effects or interactions.

Training Times for ANOVA Example: There were no significant main effects or interactions for training times on the ANOVA example.

Performance on T-Test Problem (Near Transfer): There was a significant main effect of t-test formula, $F (1, 211) = 9.18$, $MSE = 1.32$, p = .009, which indicated that the participants exposed to the conceptual t-test example outperformed those who studied the computational version. There was no effect on performance as a function of the version of the ANOVA example studied and there was no interaction between the factors.

Performance on 2-Group ANOVA Problem (Near Transfer): There were no significant main effects for this dependent measure; t-test: $F (1, 211) = 1.06$, $MSE = 2.09$, $p = .31$; ANOVA: $F (1, 211) = 0.21$, $p = .65$. However, the two-way interaction between t-test equations and ANOVA equations was significant, $F (1, 211) = 5.52$, $p < .02$. Examination of the mean scores suggest a disordinal interaction, that is, the effects of the t-test factor reverse themselves as the levels of the ANOVA factor change. Specifically, for the participants provided with conceptual t-test formulas, the conceptual ANOVA group obtained a higher score than the computational group. For participants provided with the computational t-test formulas, the computational ANOVA group obtained a higher score than the conceptual group.

Performance on 3-Group ANOVA Problem (Far Transfer): There were no significant main effects for this dependent measure; t-test: $F (1, 211) = 2.55$, $MSE = 2.65$, $p = .11$; ANOVA: $F (1, 211) = 0.10$, $p = .75$. The interaction was significant, $F (1, 211) = 6.01$, $p < .02$. Examination of the mean scores revealed the same disordinal interaction found in the 2-group problem. That is, for the participants provided with conceptual t-test formulas, the conceptual ANOVA group obtained a higher score than the computational t-test formulas. For participants provided with the computational t-test formulas, the difference was reversed.

Discussion

The overall performance differences among the groups can be summarized as follows: the combined t-test

Table 1: Scores on Test Problems as a Function of T-Test and ANOVA Examples

	T-Test Conceptual		T-Test Computational	
	ANOVA Conceptual	ANOVA Computational	ANOVA Conceptual	ANOVA Computational
T-Test Problem (max = 6)	5.43	5.46	4.99	5.07
2-Group ANOVA Problem (max = 6)	4.16	3.61	3.50	3.87
3-Group ANOVA Problem (max = 6)	4.02	3.55	3.12	3.74

conceptual and ANOVA conceptual condition tended to outperform the other conditions consisting of the other possible combinations of t-test formulas and ANOVA formulas on near and far transfer problems. There was little evidence of improved generalization by any group as a function of having been provided with elaborations. The results suggest that the first example sets the tone for the interpretation of the second example and performance on the far transfer problem. If the first t-test example used conceptual equations, then performance on the far transfer (3-group) ANOVA problem was particularly aided if the ANOVA example was also conceptual. If the first t-test example was computationally-oriented, the performance on the far transfer ANOVA problem was better if the ANOVA example was also computationally-oriented. Thus, it appears that in order for a learner to acquire a more subgoal-oriented approach to these problems, the best pedagogical approach would be to make both examples use conceptual equations. Even if the ANOVA example was conceptually-oriented, its benefits on the far transfer ANOVA problem were reduced if the initial t-test example was not also conceptual. Consistency in the examples appears to be important for subgoal learning.

The results advance prior work on subgoal learning by demonstrating that generalization can be enhanced through the nature of the equations used in examples. Thoughtfully-designed examples that include conceptually-oriented equations seem to an effective way to help learners solve novel problems.

Two caveats remain, however. First, under certain circumstances, the conceptual formula represents the most direct way of calculating sum of squares. In particular, when a data set consists of a small number of whole numbers and its mean is a whole number (which characterizes the data used in the present study), the resulting deviation score will be a whole number, which allows the learner to avoid the computational burden of decimals or fractions. Thus, one could argue that the advantage of the conceptual group in the present study therefore appears to be computational, rather than in increasing understanding. This suggests that a follow up study should explore the impact of presenting computational and conceptual equations to learners in situations in which the latter is clearly more cumbersome computationally (e.g., resulting means are not whole numbers and/or data set contains decimals).

Second, while the present results are consistent with the claims about benefits to transfer for learners who acquire useful subgoals (e.g., Catrambone, 1996, 1998), subgoal-learning was demonstrated only indirectly here. Thus, an important extension of the present work is to add converging measures, such as talk-aloud protocols, to determine if the transfer advantage can be clearly tied back to subgoal learning.

References

Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87*(1), 5-17.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(4), 1020-1031.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 12* (4), 355-376.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121-152.

Dufresne, R.J., Gerace, W.J., Hardiman, P.T., & Mestre, J.P. (1992). Constraining novices to perform expertlike problem analyses: Effects on schema acquisition. *The Journal of the Learning Sciences, 2,* 307-331.

Gravetter, F.J., & Wallnau, L.B. (2000). *Statistics for the behavioral sciences*. Belmont, CA: Wadsworth.

Larkin, J., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science, 208*, 1335-1342.

Lorch, R., & Lorch, E.P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology, 87*, 537-544.

Lovett, M.S. (1992). Learning by problem solving versus by examples: The benefits of generating and receiving information. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 956-961), Hillsdale, NJ: Erlbaum

Meyer, B.J.F., & Rice, E. (1989). Prose processing in adulthood: The text, the reader and the task. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adult and later life* (pp. 157-194). New York: Cambridge University Press.

Novick, L.R., & Holyoak, K.J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(3), 398-415.

Reed, S.K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 106-125.

Reed, S.K., & Bolstad, C.A (1991). Use of examples and procedures in problem solving. *Journal of Educational Psychology, 17*(4), 753-766.

Riesbeck C.K. & Schank, R.C. (1989). *Inside case-based reasoning*. Hillsdale, NJ: Erlbaum.

Singley, M.K., & Anderson, J.R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard Univ. Press.

Appendix

Sample Materials from T-Test Example and ANOVA Example

SAMPLE OF PROBLEM STATEMENT:
A car manufacturer that makes a car called the Jupiter just came out with a new model, the Jupiter XL. Some of the modifications made to the car are expected to improve the mpg (miles per gallon) rating of the car while other modifications are not. The manufacturer has hired your firm, an independent consumer research firm, to test the new model. To determine if there is any difference between the mpg rating of the old and new models, you collect a random sample of 5 cars of the old model and 6 cars of the new model. You drive the cars along the same city route and record the average mpg rating of each car. Here are the data:

	Old Model		New Model
Car	**MPG**	**Car**	**MPG**
1	30	1	37
2	34	2	36
3	34	3	40
4	29	4	36
5	33	5	34
		6	33

SAMPLE OF CONCEPTUAL ELABORATION FOR THE COMPUTATIONAL T-TEST VARIANCE CALCULATION
A variance is a measure of how much the scores that make up a group deviate from the mean of a group. Even though it is not obvious in the calculation below, part of the calculation of the variance involves computing the difference between each score in a group and the mean for the group. Thus, a variance measures the variability of scores around a mean.

SAMPLE OF COMPUTATIONAL T-TEST VARIANCE CALCULATION

$$s_1^2 = \text{sample variance for group 1} = \frac{\text{sum of squared scores in group 1} - \dfrac{(\text{sum of the scores in group 1})^2}{\text{number of scores in group 1}}}{(\text{number of scores in group 1}) - 1}$$

$$= \frac{(30^2 + 34^2 + 34^2 + 29^2 + 33^2) - \dfrac{(30 + 34 + 34 + 29 + 33)^2}{5}}{5 - 1} = \frac{5{,}142 - \dfrac{(160)^2}{5}}{4} = \frac{22}{4} = 5.5$$

SAMPLE OF CONCEPTUAL T-TEST VARIANCE CALCULATION

$$s_1^2 = \text{sample variance for group 1} = \frac{(\text{1st score - mean})^2 + (\text{2nd score - mean})^2 + \cdots + (\text{last score - mean})^2}{(\text{number of scores in group 1}) - 1}$$

$$= \frac{(30 - 32)^2 + (34 - 32)^2 + (34 - 32)^2 + (29 - 32)^2 + (33 - 32)^2}{5 - 1} = \frac{22}{4} = 5.5$$

SAMPLE OF COMPUTATIONAL ANOVA SUM OF SQUARES (BETWEEN) CALCULATION
$SSB = \text{sum of squares between groups}$

$$= \left[\frac{(\text{sum of scores in group 1})^2}{\text{number of scores in group 1}} + \frac{(\text{sum of scores in group 2})^2}{\text{number of scores in group 2}} + \cdots + \frac{(\text{sum of scores in last group})^2}{\text{number of scores in last group}} \right] - \frac{(\text{sum of scores in all groups})^2}{\text{number of scores in all groups}}$$

$$= \left[\frac{(160)^2}{5} + \frac{(216)^2}{6} \right] - \frac{(376)^2}{11} = 12{,}896 - 12{,}852.36 = 43.64$$

SAMPLE OF CONCEPTUAL ANOVA SUM OF SQUARES (BETWEEN) CALCULATION
$SSB = \text{sum of squares between groups}$

$= \text{number of scores in group 1 } (\text{the mean for group 1} - \text{grand mean})^2 +$

$\quad \text{number of scores in group 2 } (\text{the mean for group 2} - \text{grand mean})^2 + \cdots +$

$\quad \text{number of scores in last group } (\text{the mean for last group} - \text{grand mean})^2 = 5(32 - 34.18)^2 + 6(36 - 34.18)^2 = 43.64$