

Aiding Transfer in Statistics: Examining the Use of Conceptually Oriented Equations and Elaborations During Subgoal Learning

Robert K. Atkinson
Arizona State University

Richard Catrambone
Georgia Institute of Technology

Mary Margaret Merrill
Mississippi State University

Consistent with a subgoal-learning approach, the authors hypothesized that learners who studied statistics examples with conceptually oriented equations would transfer more successfully to novel problems compared with learners who studied examples using computationally oriented equations. The conceptually oriented equations were designed to capture the relationship between the concept and its computation, whereas the computationally oriented equations were designed to simplify calculations. This hypothesis was supported across 2 experiments. The authors also examined the implications of providing learners with elaborations of the procedures illustrated in the examples either before or after they studied them. The location of the elaborations had no apparent effect. Overall, these results demonstrate that solution procedures organized around appropriate conceptually oriented subgoals are easier to adapt for novel problems than procedures built around computationally friendly, but conceptually opaque, steps.

A considerable amount of research has examined the transfer success learners have after studying training materials such as those containing step-by-step instructions (Kieras & Bovair, 1984; Smith & Goodman, 1984), examples (e.g., Reed, Dempster, & Ettinger, 1985; Ross, 1987, 1989; Zhu & Simon, 1987), or both (Fong, Krantz, & Nisbett, 1986). Although there are a few exceptions (e.g., Fong et al., 1986; Zhu & Simon, 1987), the usual finding from such research is that learners can carry out new procedures or solve new problems that are quite similar to those on which they were trained; however, learners have difficulty when the novel cases involve more than minor changes from what they encountered in the training materials (e.g., Bassok, Wu, & Olseth, 1995; Catrambone, 1995, 1996, 1998; Novick & Holyoak, 1991).

This difficulty with transfer seems to stem from a tendency by many learners to memorize the steps of how equations are filled

out rather than learning the critical conceptual knowledge that is implicit in the details. That is, learners often represent the problem-solving procedures of training problems or worked-out examples as a set of linear steps rather than forming a nonlinear (e.g., hierarchical) representation—a more flexible representation, which could enable them to successfully solve novel problems (Dufresne, Gerace, Hardiman, & Mestre, 1992; Singley & Anderson, 1989). However, if learners attempt to apply a linear representation to solve a new problem, then they will frequently not know how to modify the problem-solving procedures in order to solve the novel problem or will carry out an inappropriate procedure.

Subgoal Learning

A more fruitful approach to problem solving would be to organize one's problem-solving knowledge in some way that generalizes across problems in a domain. One type of knowledge structure that appears to aid procedural generalization is one organized by subgoals, which are meaningful conceptual pieces of an overall solution procedure (Catrambone, 1998). Because problems within a domain often share a common set of subgoals, albeit the steps for achieving the subgoals vary from problem to problem within the domain, learning a set of subgoals is useful because subgoals can guide learners when they attempt to solve novel problems. Subgoals provide this assistance by helping learners to identify which parts of a previously learned solution procedure need to be modified in order to solve a novel problem (Catrambone, 1996, 1998).

Because novices prefer to learn from examples (e.g., LeFevre & Dixon, 1986; Pirolli & Anderson, 1985), it seems important theoretically and practically to develop an approach to example construction that helps learners go beyond memorizing sets of

Robert K. Atkinson, Division of Psychology in Education, Arizona State University; Richard Catrambone, School of Psychology, Georgia Institute of Technology; Mary Margaret Merrill, Department of Counselor Education and Educational Psychology, Mississippi State University.

Mary Margaret Merrill is now at the Department of Psychology, Louisiana State University at Shreveport.

A subset of the results from Experiment 1 was reported at the 23rd Annual Meeting of the Cognitive Science Society, Edinburgh, Scotland, August 2001. This research was supported by the Office of Naval Research, Cognitive Sciences Division, Grant N000140110459 awarded to Robert K. Atkinson and by Air Force Office of Scientific Research Grant F49620-98-1-0362 awarded to Richard Catrambone.

Correspondence concerning this article should be addressed to Robert K. Atkinson, Division of Psychology in Education, Arizona State University, P.O. Box 870611, Tempe, Arizona 85287-0611. E-mail: robert.atkinson@asu.edu

steps and instead acquire subgoal knowledge that can be flexibly applied to a wide variety of problems in a domain. According to the subgoal-learning model (Catrambone, 1995, 1996, 1998), one approach to constructing an example in order to make its goal structure explicit involves features such as the use of labels for groups of steps or visually isolating groups of steps in example solutions. Example solutions that are segregated or labeled encourage learners to self-explain how the steps go together. One result of this self-explanation process is the formation of subgoals (Catrambone, 1998). The subgoal-learning model can be summarized as follows: (a) By using a cue, such as segregating a set of grouped steps or labeling them, learners are encouraged to treat a collection of solution steps as a group. (b) After grouping the steps, learners typically attempt to explain to themselves why those steps go together in order to determine its purpose. (c) Through the process of self-explanation, learners form a goal that captures the purpose of that set of solution steps.

Factors That May Influence Subgoal Formation

As previously mentioned, according to the subgoal-learning model, several instructional manipulations—such as the use of labels or visual isolation—have been found to make the goal structure of a problem's solution explicit. However, this does not exclude the possibility that other factors may influence subgoal formation. Two potential factors include the type of equations used in examples and the use of elaborations. The goal of our work was to examine the effects of these factors on initial learning and subsequent transfer in order to help inform the subgoal-learning model.

Conceptually Oriented Versus Computationally Oriented Equations

In the domain of mathematical problem solving, learners are often provided both conceptually oriented and computationally oriented solution approaches. For example, 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 these two noticeably distinct types of formulas (i.e., conceptual and computational). In fact, many statistics texts written for the behavioral sciences present both types of formulas during the treatment of these two statistical analyses (e.g., Gravetter & Wallanu, 2000; Kiess, 2002; Shavelson, 1996; Sprinthal, 2000). 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). The conceptual formula for SS in a *t* test,

$$\sum (X - \bar{X})^2,$$

translates directly into the sum of (Σ) squared deviations $(X - \bar{X})^2$. This clearly captures how the variance term measures the variability of all the scores around 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 that can lead to more efficient calculations. The computational formula is intended to simplify calculations by working directly with raw scores instead of relying on a measure of central tendency (i.e., mean) in its computation (Gravetter & Wallanu, 2000, p. 121). However, this convenience comes at a cost: The terms of the computational formula do not translate directly into the SS, thereby clouding the true meaning behind this very fundamental component of variability.

Unlike the conceptual formula, the computational formula does not allow a learner to directly translate the terms into SS scores. As a result, the learner might not grasp that this formula is designed to measure the amount of spread about the mean and thus might have difficulty adapting this formula for a novel case. This approach might cause difficulties for novel problems—problems for which the formula cannot be applied without modification, as the underlying solution rationale is not obvious. Thus, this approach may be restricted to solving a narrow range of problems that fall into predefined problem categories that correspond to solution formulas.

Whereas a computationally oriented solution approach somewhat clouds the concept it is designed to calculate, the relationship between the concept and its computation is kept clearly evident in a conceptually oriented solution approach. This approach is considered better for understanding the concept and is often used as a teaching method because it keeps the real meaning of the term in clear focus (Sprinthal, 2000). This approach focuses on the main concepts represented in a solution procedure that might be more effective even if the solutions involve multiple or more cumbersome calculations. Thus, there might be a trade-off between the cognitive effort necessary to acquire a solution approach (and perhaps the efficiency of that approach) and the transfer distance that is covered by that approach.

In sum, there exists a possible trade-off between learning efficiency and transfer success across the two types of equations. For instance, the computational approach—in which the equations are more “compact”—might aid initial procedure acquisition and performance on problems that are just like the examples that illustrated the approach (i.e., near-transfer problems), but might make far transfer difficult. Conversely, the conceptual approach, although typically more cumbersome computationally, will aid far transfer in statistical learning by making it easier for the learner to determine how to adapt relevant parts of the procedure for novel problems.

Subgoal-Enhancing Elaborations

Another factor that has the potential to influence subgoal formation during example-based instruction is the use of elaborations. The literature contains instances of several types of elaborations that vary in the degree to which they elaborate the problem or subgoals at hand. The success of these various elaborations has been mixed, however. Although Lovett (1992) found that far transfer was facilitated by elaborated solutions, Reed and his colleagues (Reed & Bolstad, 1991; Reed et al., 1985) found little 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—aid learners.

Catrambone (1996) examined the relative benefits of rule-based instructional elaborations and subgoal labels by manipulating two factors: subgoal labels (present or absent) and rule-based elaborations (present or absent). In this case, the elaborations consisted of supplemental material describing an alternate representation or equation that could be used to solve the problems the participants were studying. Although the subgoal factor was significant, the elaboration factor was not significant. Consequently, Catrambone (1996) concluded that the presence of subgoal labeling enhanced transfer, whereas the presence of rule-based elaborations did not.

The rule-based elaborations used in the Catrambone (1996) study offered learners “what to do” knowledge and not “what it means” knowledge. This distinction is important in light of research suggesting that rules conveying what-to-do knowledge provide relatively little help to learners attempting to develop a deep conceptual understanding of a rule-based system, 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* might be more effective for gaining understanding than one dedicated to elaborating the procedural aspect of the variance formula.

If one were inclined to examine the impact of using elaborations during subgoal-oriented instruction, the question would arise about where to locate them relative to the subgoals in order to optimize their effectiveness. Unfortunately, there is no clear answer to this question. On the one hand, Catrambone (1996) examined the efficacy of presenting learners with elaborations prior to exposure to subgoal-oriented examples. Consequently, the elaborations served as a preface or an overview of the representations that the participants would subsequently see in the examples. On the other hand, although not explicitly stated, one can infer that Reed and Bolstad (1991) provided their participants with elaborations after exposure to examples. These elaborations served to encapsulate how the participants could modify the examples in an effort to enhance transfer.

It is conceivable that the placement of elaborations directly impacts their effectiveness. Thus, the mixed success encountered in the prior research may be at least partly a function of where the elaborations were located relative to the examples. Lovett’s (1992) success in fostering far transfer through elaborations provides some justification for integrating elaborations into examples despite the previously described research that failed to establish empirical support for elaborations before or after examples (e.g., Catrambone, 1996; Reed & Bolstad, 1991; Reed et al., 1985). There are, however, other potential sources of support for presenting elaborations before or after examples. Because this work generally parallels research involving the effects of advance organizers and postsummaries on learning (e.g., Ausubel, 1960; Hartley & Trueman, 1982; Lorch, Lorch, & Inman, 1993; Mayer, 1979, 1983; West & Fensham, 1976), the expository instruction literature may provide additional insight on this issue. For instance, on the one hand, elaborations containing what-it-means knowledge could—by being presented prior to subgoal-oriented worked examples—function in a similar capacity as an advanced organizer, which according to Ausubel (1968), serves to “provide ideational scaffolding for the stable incorporation and retention of more detailed and differentiated material that follows” (p. 148). On the other hand, according to Ormrod (1999), postlesson summaries serve multiple functions, including helping learners to (a) review

material, (b) determine which of the many ideas they have studied are important, and (c) pull key ideas into a more cohesive organizational structure.

In sum, the relative effectiveness of placing elaborations before or after a learner studies a set of instructional material remains unsettled. On the one hand, elaborations offered as a preinstructional aid might, like an advanced organizer, provide a useful cognitive structure for receiving new material. However, this approach is not without its drawbacks. For instance, preexample elaborations run the risk of appearing too abstract and noncontextualized. On the other hand, a postexample elaboration might—similar to postlesson summaries—supplement or reinforce what was illustrated in the examples, thereby facilitating the integration and application of recently acquired procedural knowledge. Thus, learners might also demonstrate learning gains from elaborations presented after subgoal-oriented examples.

Overview of Experiments

The purpose of the present research was to examine, in the domain of statistics (specifically, *t* tests and ANOVAs), the effects on initial learning and subsequent transfer due to examples demonstrating conceptually oriented subgoals versus computationally oriented subgoals and the use and position of elaborations. Performance was assessed in two ways: (a) time spent studying the instructional material that included worked-out examples and (b) correctness of solutions on near- and far-transfer problems. Specifically, the aim of Experiment 1 was to compare the relative impact on problem-solving performance of subgoals structured around either conceptual or computational equations as well as the relative efficacy of elaboration overviews presented to learners either before or after they studied an example. Although we hypothesized that subgoals oriented around conceptual equations would aid far transfer, we predicted that near-transfer performance for the computational group would be superior because the equations in that condition are more compact. Similarly, we predicted faster learning times in the computation condition, again on the notion that the equations are more compact. Essentially, we assumed a plausible trade-off between learning efficiency and near-transfer success, with far-transfer success across the subgoals structured to accommodate two different types of equations. The purpose of Experiment 2 was to replicate and extend the results of Experiment 1 through the complementary use of concurrent verbal protocols (CVP) and retrospective debriefings (RD), referred to as the *CVP-RD methodology* (Taylor & Dionne, 2000).

Experiment 1

This experiment was designed to address two primary research questions: (a) Will a learner be more capable of adapting a subgoal-oriented solution organized around conceptual equations compared with a solution procedure that emphasizes computational efficiency? (b) Do the learning effects of subgoal elaborations differ according to whether they are placed before or after the relevant example?

Method

Participants. The participants were 112 lower division undergraduates (46 men and 66 women) drawn from several psychology and educational

psychology courses at Mississippi State University who participated in the experiment for course credit. Participants had not taken any previous statistics courses.

Design. Participants were randomly assigned in equal proportions to the cells of a 2×2 between-subjects factorial design ($n = 28$ per condition). The first factor was the characteristics of the variance formulas (conceptual or computational) in the t test and ANOVA examples. The second was the position of the elaborative material (preceding the examples or following the examples).

Training phase. During the training phase, the participants received a demographic questionnaire and a 10-page instructional booklet. The demographic questionnaire asked each learner to provide information (e.g., GPA, ACT scores, number of college mathematics courses in progress or completed) that could be used to gauge the learner's level of prior knowledge. Participants also received an instructional booklet containing a general introduction to statistical hypothesis tests and two training examples, one representing a t test (see Figure 1) and another representing the use of an ANOVA for the same two-group comparison. The introduction to statistical hypothesis tests described the utility of these procedures and provided a summary 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 overview described the purpose of the test without going into detail regarding how to perform the test's calculations. These introduction and overview materials were the same for all participants.

With respect to the t test example, the variance formulas were either conceptual or computational in nature (see Figure 2). That is, the formulas presented for the variance either captured how the variance term measures the amount of spread about the mean (i.e., conceptual) or they permitted the learner to calculate SS directly from raw scores while concealing the true meaning behind SS (i.e., computational). Similarly, with regard to the ANOVA example, the variance formulas (for SS between [SSB] and SS within [SSW]) were either conceptual or computational in character.

Both examples were supplemented with elaborations that provided what-it-means knowledge. That is, they were designed to describe the meaning behind the various formulas used in the two hypothesis tests. For the t test, elaborations were provided about the variance formula, the pooled-variance formula, and the t statistic. For example, the following was provided as the elaboration for the variance formula in the conceptual condition:

A variance is a measure of how much the scores that make up a group deviate from the mean of a group. . . [P]art 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.

For the ANOVA, elaborations were provided about the SSW, the SSB, and the F statistic. Depending on the condition, these elaborations were presented either before each example (t test elaboration before the t test example and ANOVA elaboration before the ANOVA example) or after each example in the context of the overall hypothesis testing procedure. 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 (see Figure 2). Second, each of the six calculational elements in the two examples was either labeled or visually isolated. The t test elements 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 elements were to find or do preliminary calculations, SSB, SSW, mean squares between, mean squares within, and F statistic.

Test phase. The six-page test booklet contained three test problems for the participants to solve. The first test problem required participants to conduct a t test; this served as a near-transfer assessment (see Figure 3). The second problem required them to conduct a two-group ANOVA; this also served as a near-transfer assessment (see Figure 3). The third problem required a three-group ANOVA; this served as far-transfer assessment because it required the learner to adapt the equations for variance (see Figure 4). The extension is a relatively straightforward, modular extension of the conceptual equations. However, the extension is less straightforward in the computational equations because it involves changes to the "interior" of the equations that do not map as neatly to the additional group. 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.

Procedure. Participants were asked to fill out the demographic questionnaire and then study carefully the instructional booklet containing the

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

Figure 1. Problem statement for the t test example.

SAMPLE OF COMPUTATIONAL *T* TEST VARIANCE CALCULATION

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

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

SAMPLE OF CONCEPTUAL *T* TEST VARIANCE CALCULATION

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

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

Figure 2. Sample of computational and conceptual *t* test variance calculations.

training examples because after studying it they would be asked to solve three 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 rather than memorizing formulas.

Materials were administered to participants in groups ranging in size from 5 to 30 participants. The participants worked for approximately 75 min and were asked to show all of their work.

Scoring. 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 solution element. The three test problems each contained six calculational elements. The

PROBLEM STATEMENT FOR THE *T*-TEST PROBLEM AND 2-GROUP ANOVA PROBLEM

The superintendent of a rural school district wants to determine if a new method of teaching, called Direct Instruction, will produce higher achievement scores for students enrolled in her district's elementary classrooms when compared to the teaching method traditionally used by district teachers. An experiment was conducted to compare the two methods of instruction using students from one of the small, rural elementary schools in the district. Each student was randomly assigned to one of two instructional methods for the fall term. The criterion for measuring achievement was a student's score on a standardized test. The results are shown below.

Traditional		Direct Instruction	
Student	Score	Student	Score
1	78	1	90
2	69	2	86
3	83	3	91
4	75	4	82
5	81		

INSTRUCTIONS FOR *T*-TEST PROBLEM (NEAR TRANSFER)
Determine if there is any difference between the standardized test scores of the students exposed to the traditional and Direct Instruction teaching methods using a *t*-test.

INSTRUCTIONS FOR 2-GROUP ANOVA PROBLEM (NEAR TRANSFER)
Determine if there is any difference between the standardized test scores of the students exposed to the traditional and Direct Instruction teaching methods using an analysis of variance (ANOVA).

Figure 3. Problem statement for the *t* test problem and two-group ANOVA problem.

PROBLEM STATEMENT FOR 3-GROUP ANOVA PROBLEM

Suppose the same superintendent discovers another new method of teaching, called Core Knowledge, and wants to compare this one with the other novel method of teaching, Direct Instruction, as well as the traditional method. Another experiment was conducted to compare the three methods of instruction using students from one of the small, rural elementary schools in the district. For this experiment, each student was randomly assigned to one of *three* instructional methods for the fall term. Again, the criterion for measuring achievement was a student's score on a standardized test. The results are shown below.

Traditional		Direct Instruction		Core Knowledge	
<u>Student</u>	<u>Score</u>	<u>Student</u>	<u>Score</u>	<u>Student</u>	<u>Score</u>
1	68	1	89	1	95
2	72	2	87	2	88
3	81	3	92	3	81
4	79	4	83	4	85
5	84			5	92

INSTRUCTIONS FOR 3-GROUP ANOVA PROBLEM (FAR TRANSFER)

Determine if there is any difference between the standardized test scores of the students exposed to the traditional, Direct Instruction, and Core Knowledge teaching methods using an analysis of variance (ANOVA).

Figure 4. Problem statement for the three-group ANOVA problem.

correct numerical answer to the element was awarded 1 point. For example, the correct answer to the second element in the *t* test problem, Group 1 variance, was 30.2. If a participant's problem-solving protocol contained this answer, she was given a point. The maximum score for each of the three test problems was 6.

Because most elements contained subcomponents, the binary system allowed us to award partial credit for each element. This permitted us to capture the proportion of the element's solution—for those participants who did not have the correct numerical answer for the element—that was correct. For instance, the equation associated with the second element (i.e., Group 1 variance) in the conceptual condition was coded for the presence or absence of seven subcomponents, 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 subcomponents, he or she was awarded a 0.86 for the element. If the element was correct except for a trivial math error, the participant received full credit (1 point) for that particular element.

Results

To validate the scoring system, 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 2×2 analysis of covariance (ANCOVA) was conducted on the study times of the two examples (including elaborative material) as well as on the correctness measures for the three test problems. According to the procedure described by Hays (1994), each measure was tested for homogeneity of regression, and the results were found to be nonsignificant—all F s < 2.2 . For statistical control purposes, the data were adjusted for one covariate, specifically the composite score on the ACT. This was selected as a covariate over the other demographic variables we collected (e.g., GPA, number of college-level math courses completed or in

progress) for several reasons, including (a) the ACT is designed to measure general educational development for college-age learners in several skill areas including English, mathematics, reading, and science reasoning; (b) participation in this experiment required many of these skills; (c) the participants were lower division undergraduates who presumably had taken the ACT recently; (d) the participants did not consistently report GPA; and (e) although number of college mathematics courses was consistently reported, the number of college mathematics courses did not consistently correlate with any of the dependent measures used in this experiment. Table 1 presents the unadjusted means and standard deviations derived from the experiment.

Study times for t test example plus elaboration. There was no effect on the combined reading time for the *t* test example and elaboration as a function of equation type or position of elaboration.

Study times for ANOVA example plus elaboration. There was a significant effect of position of elaboration on the time to read the ANOVA example and elaboration, $F(1, 107) = 7.51$, $MSE = 8.49$, $p = .01$. Cohen's *f* statistic for these data yields an effect size estimate of .27, which corresponds to a medium effect. Participants reading the elaborations prior to the examples ($M = 7.57$ min) were faster than their peers who read the elaborations after exposure to the examples ($M = 9.01$ min).

Performance on t test problem (near transfer). There was no effect of equation type, $F(1, 107) = 0.77$, $MSE = 2.35$, $p = .38$, on problem-solving success, which suggests that the participants exposed to the conceptual *t* test example performed similarly to those who studied the computational version. However, the two-way interaction between type of equations and position of elaboration was significant, $F(1, 107) = 5.63$, $p = .019$. Cohen's *f* statistic for these data yields an effect size estimate of .23, which corresponds to a medium effect.

Table 1
Study Times of Examples and Scores on Problems as a Function of Example Type and Position of Overview (Experiment 1)

Variable	Conceptual solution				Computational solution			
	Preexample elaboration		Postexample elaboration		Preexample elaboration		Postexample elaboration	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Times (minutes)								
<i>t</i> test example + elaborative material	10.39	3.41	12.07	4.06	11.27	3.42	11.48	2.85
ANOVA example + elaborative material	6.98	1.69	8.73	2.95	8.25	3.61	9.34	3.36
Scores (max. = 6)								
<i>t</i> test problem	4.64	1.71	5.14	1.47	4.75	1.75	4.53	1.76
Two-group ANOVA problem	3.46	1.89	3.79	1.44	2.53	1.62	2.33	1.57
Three-group ANOVA problem	3.32	1.80	3.26	2.00	2.23	1.83	2.35	1.68

Note. ANOVA = analysis of variance; max. = maximum.

Subsequent analysis demonstrated that there was a simple main effect for type of equation at the postexample level of the location of elaboration factor, $F(1, 107) = 5.36, p < .025$. Cohen's f statistic for these data yields an effect size estimate of .31, which corresponds to a medium effect. Specifically, the postexample elaboration participants who received the conceptual equations displayed higher scores than their counterparts who were exposed to the computational equations. The remaining simple main effects were not significant.

Performance on two-group ANOVA problem (near transfer). Participants in the conceptual condition significantly outperformed those in the computational condition, $F(1, 107) = 20.93, MSE = 1.91, p < .001$. Cohen's f statistic for these data yields an effect size estimate of .44, which corresponds to a large effect. There was, however, no main effect of position of elaboration, $F(1, 107) = 0.03, MSE = 2.35, p = .86$. The interaction between type of equations and position of elaboration was significant, $F(1, 107) = 7.10, p < .02$. Cohen's f statistic for these data yields an effect size estimate of .25, which corresponds to a medium effect.

Subsequent analysis demonstrated that—comparable with the outcome for the *t* test problem—there was a simple main effect for type of equation at the postexample level of the location of elaboration factor, $F(1, 107) = 25.96, p < .001$. Cohen's f statistic for these data yields an effect size estimate of .69, which corresponds to a large effect. Again, participants in the conceptual condition displayed higher scores than their counterparts in the computational condition. The remaining simple main effects proved to be nonsignificant.

Performance on three-group ANOVA problem (far transfer). Participants in the conceptual condition significantly outperformed those in the computational condition, $F(1, 107) = 11.80, MSE = 2.40, p < .001$. Cohen's f statistic for these data yields an effect size estimate of .33, which corresponds to a medium effect. There was no effect of position of elaboration or an interaction between elaboration position and type of formulas used in the examples.

Supplemental analysis. Because time studying the instructional material may account for the differences cited above, we conducted several 2×2 ANCOVAs on the correctness measures for the *t* test problem, two-group ANOVA problem, and three-group ANOVA problem using total study time (study time for *t* test example plus elaboration plus study time for ANOVA exam-

ple plus elaboration) as a second covariate (ACT scores being the first covariate). The results of the ANCOVAs corroborated the aforementioned statistical conclusions, which implies that study time was not a confounding variable in this instance.

Because there are documented gender differences in mathematical thinking (e.g., Fennema, Carpenter, Jacobs, Franke, & Levi, 1998), we elected to examine for the possibility of an interaction between gender and the orientation—conceptual versus computational—of the formulas in the examples. To accomplish this, we conducted several posteriori 2×2 ANCOVAs, with type of equations and gender as independent variables along with the composite ACT score as a covariate. For each of the three correctness measures (*t* test problem, two-group ANOVA problem, and three-group ANOVA problem), this interaction was nonsignificant—all F s < 1 . Also, the main effect of gender was nonsignificant across the three correctness measures.

Discussion

Will a learner be more capable of adapting a subgoal-oriented solution organized around conceptual equations compared with a solution procedure that emphasizes computational efficiency? The present experiment provides support for the hypothesized advantage of conceptually oriented equations as aids to adapting a solution procedure to accommodate a novel problem. Specifically, the results suggest that subgoals organized around conceptual equations aid far-transfer performance relative to subgoals organized around computational equations. In addition to the statistically significant difference, the medium effect size associated with this result indicated that the conceptual equations also created a meaningful difference between the two conditions. The fact that the conceptual group was able to transfer to the three-group ANOVA problem (far transfer) more successfully is consistent with the hypothesis that these participants were better able to modify the equation to calculate variances for three groups compared with the computational participants. As predicted, it is easier to adapt the procedure for finding individual deviations from a group mean (the conceptual approach) than to modify the internal structure of the computational equation. For computational learners, the internal components are presumably more opaque and thus

they are less likely to be able to determine how to adapt them successfully for novel problems.

Do subgoal elaborations placed prior to examples lead to more dramatic learning effects relative to elaborations placed subsequent to examples? The location of the elaborations does not seem to consistently impact performance. The absence of a consistent effect on problem-solving performance due to the pre- and postmanipulation involving the elaborative material suggests several possibilities. One is that the material simply was not very effective and therefore its positioning did not matter. Such a claim could be tested by a replication in which there are conditions that do not receive the elaborative material. A second explanation is that positioning of such material does not matter for learning in this domain, at least when learners study the types of examples we provided. A third possibility is that because the elaborative material essentially recapitulates the information provided in the examples (with a bit more depth) but without numbers, learners may have been less inclined to pay attention to them. This might hold true even for learners in the precondition because they knew they would be studying concrete examples at some point; prior work demonstrated learners' preference for studying examples over "abstract" information (see LeFevre & Dixon, 1986). The fact that participants in the precondition worked through the ANOVA training materials faster than those in the postcondition might be due largely to the former group finding the elaborations difficult to comprehend and therefore reading them quickly in anticipation of reaching concrete examples. Unfortunately, we did not collect reading times on the elaborations separately from times on the examples.

A caveat remains, however. As mentioned earlier, there was an unanticipated interaction between conceptual and computational equations and the use of pre- and postexample elaborations on both near-transfer problems. Therefore, the main effect for equation type—that is, that subgoals organized around conceptual equations aid far-transfer performance relative to subgoals organized around computational equations—needs qualification. Specifically, the results of this experiment suggest that, for near-transfer problems, participants provided with conceptually oriented equations outperformed their peers presented with computationally oriented equations, albeit this advantage was observed only when elaborations were presented to learners after—as opposed to before—they studied an example.

One possible explanation for this result is that the elaborations provided to the learners in the conceptual condition served to—much like postlesson summaries—clarify the structure of the examples and provided learners with an additional conceptual framework for identifying the example's essential information. Because the relationship between the concept and its computation was kept clearly evident in the conceptually oriented solution approach, the better learning gains associated with postexample elaborations suggest that those elaborations helped these learners by (a) assisting them in determining which of the many ideas they have studied were important, (b) supplementing and reinforcing what was illustrated in the examples using conceptually oriented subgoals, and (c) assisting them in pulling key ideas into a more cohesive organizational structure. On the other hand, the preexample elaborations may have appeared too abstract and noncontextualized for the learners to understand without having first studied a conceptually oriented example.

Experiment 2

The results of Experiment 1 clearly establish that subgoals organized around conceptual equations aid far-transfer performance relative to subgoals organized around computational equations. There was also an unexpected suggestion in the results that a conceptual subgoal organization aided near transfer. To determine the robustness of the conceptual effect, we conducted a second experiment again comparing conceptual and computational equations. In this follow-up experiment, however, we did not include elaborations as an experimental manipulation because we wanted to focus our efforts on more closely examining the conceptual effect. Against this background, we elected to use a slightly different experimental format, one designed to provide a more complete account of the phenomena, namely, the dual CVP-RD methodology (Taylor & Dionne, 2000). With these two modifications in place, we once again examined a question raised in the previous experiment, namely, will a learner be more capable of adapting a subgoal-oriented solution organized around conceptual equations compared with a solution procedure that emphasizes computational efficiency?

Method

Participants. The participants were 30 undergraduates (3 men and 27 women) drawn from several psychology and educational psychology courses at Mississippi State University who participated in the experiment for course credit. Participants had not taken any previous statistics courses.

Design. Participants were randomly assigned in equal proportions to either the conceptual-formula condition or the computational-formula condition.

Training phase. Similar to Experiment 1, participants in this experiment received an instructional booklet containing a brief description of statistical hypothesis tests, two training examples, one representing a *t* test and another representing the use of an ANOVA for the same two-group comparison. With respect to the two training examples, the variance formulas were either conceptual or computational in nature.

Test phase. The test problems were identical to Experiment 1.

Procedure. The procedure was identical to Experiment 1, with one notable exception. Instead of administering the material to participants in groups, the students in the present experiment were run individually, and each session was videotaped. Specifically, this experiment was modeled after the dual CVP-RD methodology of Taylor and Dionne (2000), whereby the participant is asked to talk out loud while studying the instructional material and solving subsequent problems, followed by a short question-and-answer session in which the participant is interviewed regarding his or her problem-solving performance.

For the verbal-protocol portion, participants were asked to verbalize their thoughts concurrently while studying the instructional material and solving the three problems on the posttest. The participants were asked to talk aloud and verbalize anything that came to their mind. They were not instructed to provide unusual information. Instead, they were asked to provide a report of what they were thinking more or less while they were thinking it. To assist the participants with the thinking (talking)-aloud procedure, they were trained using a warm-up problem (playing the experimenter in a game of tic-tac-toe). During the session, if the participants lapsed into silence for more than 15 s, the experimenter instructed them to "Please keep talking." After the participants finished the posttest, the experimenter asked several questions, including asking them to describe, in their own words, the meaning of four terms they encountered, including variance, pooled variance, SSB, and SSW.

Scoring. Scoring of the problem-solving protocols was identical to Experiment 1. Scoring of the concurrent verbal protocol collected during

the posttest and retrospective interview was based on several considerations. First, we approached the scoring of the *t* test problem (i.e., Problem 1 of the posttest) by focusing on whether the participants explicitly mentioned that they were attempting to find the variance (or sample variance) for Group 1 and for Group 2—labels provided to both conceptual and computational conditions. We hypothesized that if participants were explicitly noting the subgoals they were trying to achieve while solving the *t* test problem, then they might also be more likely to explicitly note the relevant subgoals (SSB and SSW) for the ANOVA problems.

We approached the coding of the two ANOVA problems in a similar fashion. For the two-group ANOVA problem (i.e., Problem 2), we coded for whether the participants explicitly mentioned that they were trying to find SSB and SSW. We also scored for whether the participants made a spontaneous comment about how SSW was like the sample variances from the *t* test problem. Such a comment would indicate a deeper understanding of variance, and this might be related to being in the conceptual condition. For the three-group ANOVA problem (i.e., Problem 3), we scored it just like the two-group problem (except for the comment about noting that SSW is like variance from the *t* test problem). We also noted whether participants made any comments about encountering difficulties adapting their solutions to accommodate a third group.

For the comments generated during the retrospective interview, we scored whether participants indicated that sample variance measured how “different” (or other similar words) each score was from the mean. Also, we scored whether participants indicated that SSB measured how different each group mean was from the grand mean. Finally, we scored whether participants indicated that SSW measures how different each score was from the group mean. Again, such comments would indicate a deeper understanding of variance. In sum, a binary scoring system was developed to score videos generated as a result of the CVP-RD methodology. Specifically, the concurrent verbal protocols were coded for whether the participant (a) used a label like *variance for Group 1* while solving the *t* test problem; (b) used a label like *variance for Group 2* while solving the *t* test problem; (c) used a label like *SSB* while solving the two-group ANOVA problem; (d) used a label like *SSW* while solving the two-group ANOVA problem; (e) made a comment relating SSW, while solving the two-group ANOVA problem, to the group variances from the *t* test problem; (f) used a label like *SSB* while solving the three-group ANOVA problem; (g) used a label like *SSW* while solving the three-group ANOVA problem; and (h) expressed confusion over how to adapt the formulas to accommodate the novel three-group situation. Additionally, the verbal protocols generated during the retrospective interview were coded for whether the participant noted that (i) variance assesses how different each score is from the group mean, (j) SSB assesses how different each group mean is from the grand mean, and (k) SSW assesses how different each score is from the group mean. Participants were awarded 1 point for the presence of each of these features. Thus, the maximum combined score (concurrent protocol and retrospective interview) for each individual was

10 (Item h was not included in the total). This score provides a converging measure on how well participants understood variance.

Results

To validate the scoring system associated with the CVP-RD methodology, two raters independently scored a random sample of 20% of the verbal protocols and agreed on scoring 100% of the time. One experimenter independently scored the remaining verbal protocols. As with the previous experiment, an ANCOVA was conducted on the study times of the two examples as well as on the correctness measures for the three test problems. Similar to Experiment 1, for statistical control purposes, the data were adjusted for one covariate, the composite score on the ACT. Each measure was tested for homogeneity of regression, and the results were found to be nonsignificant—all $F_s < 1.2$. Table 2 presents the primary results from the experiment.

Study times for t test example. There was no effect on the reading time for the *t* test example as a function of equation type, $F(1, 27) = 0.19$, $MSE = 10.87$, $p = .66$, which suggests that the participants assigned to the conceptual condition spent approximately the same amount of time studying the *t* test example as their peers in the computational condition.

Study times for ANOVA example. There was no effect on the reading time for the ANOVA example as a function of equation type, $F(1, 27) = 0.01$, $MSE = 11.53$, $p = .94$, which suggests that the participants assigned to the conceptual condition spent approximately the same amount of time studying the ANOVA example as their peers in the computational condition.

Performance on t test problem (near transfer). There was not a reliable effect of equation type, $F(1, 27) = 3.64$, $MSE = 1.28$, $p = .067$, which suggests that the participants exposed to the conceptual examples performed similarly to those who studied the computational versions.

Performance on two-group ANOVA problem (near transfer). As in Experiment 1, participants in the conceptual condition significantly outperformed those in the computational condition, $F(1, 27) = 5.76$, $MSE = 1.55$, $p < .05$. Cohen’s f statistic for these data yields an effect size estimate of .45, which corresponds to a large effect.

Performance on three-group ANOVA problem (far transfer). As in Experiment 1, participants in the conceptual condition significantly outperformed those in the computational condition, $F(1, 27) = 4.69$, $MSE = 2.33$, $p < .05$. Cohen’s f statistic for these data

Table 2
Study Times of Examples and Scores on Problems as a Function of Example Type (Experiment 2)

Variable	Conceptual solution		Computational solution	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Times (minutes)				
<i>t</i> test example	9.47	4.07	9.33	2.29
ANOVA example	7.80	3.73	7.73	2.89
Scores (max. = 6)				
<i>t</i> test problem	5.34	1.05	4.83	1.35
Two-group ANOVA problem	4.57	0.70	3.73	1.70
Three-group ANOVA problem	4.24	1.15	3.28	1.91

Note. ANOVA = analysis of variance; max. = maximum.

yields an effect size estimate of .42, which corresponds to a large effect.

Verbal-protocol analysis. Overall, participants in the conceptual condition ($M = 5.67$, $SD = 1.80$) produced a significantly greater number of coded self-explanations than those in the computational condition ($M = 3.67$, $SD = 1.80$), $F(1, 27) = 9.25$, $MSE = 3.31$, $p = .005$. Cohen's f statistic for these data yields an effect size estimate of .59, which corresponds to a large effect. On closer examination, this difference in favor of the conceptual condition could be attributed to 3 of the 10 categories included in the coding scheme. First, the participants assigned to the conceptual condition used the *SSB* label in the two-group ANOVA problem more frequently than their counterparts in the computational condition, $\chi^2(1, N = 30) = 4.62$, $p = .032$. In fact, in the conceptual condition, 100% (15/15) of the participants used label *SSB* in the two-group ANOVA problem compared with 27% (4/15) of their counterparts in the computational condition. Second, participants assigned to the conceptual condition used the label *SSW* in the two-group ANOVA problem more frequently than their counterparts in the computational condition, $\chi^2(1, N = 30) = 9.13$, $p = .003$. As with the previous category, in the conceptual condition, 100% (15/15) of the participants used the label *SSW* in the two-group ANOVA problem compared with 47% (7/15) of their peers in the computational condition. Finally, the participants assigned to the conceptual condition produced a comment relating *SSW* to group variance of a t test more frequently than the participants assigned to the computational condition, $\chi^2(1, N = 30) = 4.62$, $p = .032$. That is, in the conceptual condition, 27% (4/15) of the participants used comments relating *SSW* to group variance of t test compared with 0% (0/15) of their counterparts in the computational condition.

Supplemental analysis. Similar to the previous experiment, we conducted several ANCOVAs on the correctness measures for the t test problem, two-group ANOVA problem, and three-group ANOVA problem using total study time (study time for t test example plus study time for ANOVA example) as an additional covariate. The results of the ANCOVAs corroborated the aforementioned statistical conclusions.

It also appeared worthwhile to explore the relationship between the participants' verbal-protocol performance and their problem-solving performance. This was fueled in part by the unexpected finding that, despite outperforming the participants assigned to the computational condition on both the two-group (near-transfer) and three-group (far-transfer) ANOVA problems, the conceptual condition participants produced significantly more self-explanations related to *SSB* and *SSW* while solving only the two-group problem. As a result, we wanted to know the nature and strength of the relationship between the participants' frequency of explicitly noting the subgoals they were trying to achieve at any point while solving the two ANOVA problems and their far-transfer performance.

To address this issue, we created two new variables: one that represented the participants' use of the *SSB* label across both the two-group and the three-group ANOVA problems and one that captured the participants' use of the *SSW* label across the same two problems. We then examined the relationship between these two new categories and far-transfer performance for each condition using several Pearson product-moment correlations. There was no statistically significant correlation between far-transfer perfor-

mance and the frequency of explicitly noting the subgoals participants were trying to achieve in the computational condition. In contrast, there was a statistically significant relationship between far-transfer performance and use of the *SSB* label across the two ANOVA problems, $r(28) = .56$, $p = .001$, as well as the use of the *SSW* label, $r(28) = .56$, $p = .001$, for the conceptual condition participants. On the basis of Cohen's (1992) effect size guidelines for Pearson product-moment correlation coefficients, both of these correlations represent large effects.

Discussion

This experiment replicates and extends the advantage of conceptually oriented equations found in the previous experiment. Specifically, the results from the problem-solving protocols replicate the finding from Experiment 1 that subgoals organized around conceptual equations can aid both near- (two-group ANOVA) and far-transfer (three-group ANOVA) performance relative to subgoals organized around computational equations. Moreover, the large effect sizes associated with the results of the near-transfer problem, far-transfer problem, and verbal-protocol analysis clearly underscore the meaningful difference in problem-solving performance created between the two conditions by the subgoals oriented around conceptual equations.

In addition, the results of the verbal-protocol analysis are consistent with the results from the problem-solving protocols. Overall, participants presented with conceptual equations generated more self-explanations—at least the transfer-relevant ones we coded for—than their peers in the computational condition. The conceptual-condition participants learned more appropriate subgoals than the computational-condition participants. For instance, all of the participants in the conceptual condition used both the *SSW* label and the *SSB* label while solving the two-group ANOVA problem (i.e., Problem 2). Moreover, 4 of the 15 conceptual participants—in contrast to 0 computational participants—made a spontaneous comment about how *SSW* was like the sample variances from the t test problem. This finding is particularly significant given that the participants were not cued to look for similarities across the t test and ANOVA problems.

The supplemental analysis also revealed a significant relationship between participants' verbal-protocol performance and their problem-solving performance for the participants assigned to the conceptual condition. In particular, for these participants, we found a strong, direct relationship between the frequency with which they explicitly noted the subgoals they were pursuing while solving the two ANOVA problems and their problem-solving performance on the far-transfer item. This provides additional evidence supporting the advantage of subgoals structured around conceptually oriented equations.

General Discussion

The results of this study advance prior work involving the subgoal-learning model by demonstrating that generalization can be enhanced through the nature of the equations used in examples. Specifically, thoughtfully designed examples that include subgoals organized around conceptually oriented equations seem to be an effective way to help learners solve novel problems. Additionally, it appears that subgoals structured around conceptually oriented

equations play a much more significant role in transfer compared with the effects of elaborations designed to enhance the subgoals.

There are several caveats, however. The study produced several unexpected findings, including the lack of any hypothesized differences in reading (learning) times between the two equation groups and the lack of a hypothesized advantage of subgoals structured around computationally oriented equations on the near-transfer t test problem across both experiments. Although we correctly hypothesized that subgoals oriented around conceptual equations would aid far transfer, we incorrectly predicted that near-transfer performance for participants in the computational group would be superior because the equations in that condition are more compact—instead, the near-transfer t test performance associated with the two types of equations was statistically equivalent to each other, whereas it was the conceptual-condition participants who were superior on the two-group ANOVA problem.

We offer an initial conjecture for why the two conditions performed similarly on the first near-transfer problem (t test) but differently on the second (two-group ANOVA). It is possible that learners are likely to notice similar subgoals across problem types. The similarity of the variance subgoals in the t test and ANOVA problems, for learners in the conceptual condition, is fairly high and therefore learners in the conceptual condition would be relatively likely to notice the similarity, at least compared with those in the computational condition. The observation of such a similarity might help problem-solving performance. This observation—and the associated benefit in problem solving—would not occur, though, until the second problem is reached, that is, the two-group ANOVA problem.

If the conjecture above is correct, this would suggest that the subgoal-learning model should be modified to suggest that not only do learners tend to self-explain the purpose of a set of solution steps (i.e., identify the subgoal they achieve), perhaps they also tend to self-explain any conceptual similarities they see across subgoals from discrete examples. The analysis of the verbal protocols collected during problem solving provides preliminary support for this account. That is, approximately one third of the learners in the conceptual condition made a spontaneous comment about how the SSW in the two-group ANOVA problem was similar to the sample variances from the t test problem. It would appear that subgoals containing conceptually oriented equations can foster interexample comparisons and that the divergence in problem-solving performance between the two types of equations can only occur after such a comparison can be made. This suggests that if the order of the t test and two-group ANOVA problems were reversed in the test phase, then the two groups would perform similarly on the two-group ANOVA problem, whereas the conceptual group would outperform the computational group on the t test problem. We also predicted faster learning times in the computation condition, again based on the notion that the equations associated with that group are more compact, which did not materialize. Why the computational equations do not lend themselves to increased learning efficiency relative to the conceptual equations remains a question for further exploration.

Although there are several issues we plan to explore, there are two issues we plan to address initially. First, under certain circumstances, the conceptual formula represents the most direct way of calculating SS. Although it does not characterize the data used in the present study, when a data set consists of a small number of

whole numbers and its mean is a whole number, 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 appears to be partly due to the nature of the numbers in the test problems rather than increased 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., data set contains decimals). If we were to find that more “messy” training examples reduce the performance difference between the two groups, then it would make sense to explore whether “clean” conceptual examples can consistently produce superior far transfer over both clean and messy computational examples.

Second, we plan to pursue an additional issue raised by our contention that the conceptual approach seems to lead to more dramatic learning effects than the computational approach. Traditionally, a conceptual approach is presented—with varying degrees of clarity—in textbooks along with a computational approach. One might wonder why students struggle with transfer in statistics learning if the conceptual approach is presented—assuming it is presented clearly—in textbooks. One possibility is that when students are shown both approaches, they tend to ignore or gloss over the conceptual one in order to get to the “real” equations, which are the ones they will most likely use. Thus, an interesting experiment might be to compare a conceptual-plus-computational-example condition with a pure-conceptual condition to a pure-computational condition. The conceptual-plus-computational condition could be further subdivided into conceptual-followed-by-computational solution and computational-followed-by-conceptual solution in order to check whether the first set of equations dominates over the second set or whether one dominates the other regardless of order. If one were to find that the conceptual-plus-computational conditions perform like the pure-computational condition, this would be consistent with the claim that students more or less ignore the conceptual version in favor of the computational one, thereby implying that textbooks designers are making a strategic mistake in their treatment of the subject matter.

References

- Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal materials. *Journal of Educational Psychology, 51*, 267–272.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart & Winston.
- Bassok, M., Wu, L. L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretive effects of content on problem-solving transfer. *Memory and Cognition, 23*, 354–367.
- Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87*, 5–17.
- Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1020–1031.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127*, 355–376.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.
- 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.

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher*, 27, 6–11.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18, 253–292.

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

Hartley, J., & Trueman, M. (1982). The effects of summaries on the recall of information from prose: Five experimental studies. *Human Learning*, 1, 63–82.

Hays, W. L. (1994). *Statistics* (5th ed.). New York: Harcourt.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, 8, 255–273.

Kiess, H. O. (2002). *Statistical concepts for the behavioral concepts* (3rd ed.). Boston: Allyn & Bacon.

LeFevre, J., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction*, 3, 1–30.

Lorch, R. F., Lorch, E. P., & Inman, W. E. (1993). Effects of signaling topic structure on text recall. *Journal of Educational Psychology*, 85, 281–290.

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.

Mayer, R. E. (1979). Twenty years of research in advance organizers: Assimilation theory is still the best predictor of results. *Instructional Science*, 8, 133–167.

Mayer, R. E. (1983). Can you repeat that? Qualitative and quantitative effects of repetition and advance organizers on learning from science prose. *Journal of Educational Psychology*, 75, 40–49.

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

Ormrod, J. E. (1999). *Human learning* (3rd ed.). Columbus, OH: Merrill.

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples

in the acquisition of recursive programming skill. *Canadian Journal of Psychology*, 39, 240–272.

Reed, S. K., & Bolstad, C. A. (1991). Use of examples and procedures in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 753–766.

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, 106–125.

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

Ross, B. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629–639.

Ross, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456–468.

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences*. Boston: Allyn & Bacon.

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

Smith, E. E., & Goodman, L. (1984). Understanding written instructions: The role of an explanatory schema. *Cognition and Instruction*, 1, 359–396.

Sprinthall, R. C. (2000). *Basic statistical analysis* (6th ed.). Boston: Allyn & Bacon.

Taylor, K. L., & Dionne, J. (2000). Accessing problem-solving strategy knowledge: The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educational Psychology*, 92, 413–425.

West, L. H., & Fensham, P. J. (1976). Prior knowledge or advance organizers affective variables in chemical learning. *Journal of Research in Science Teaching*, 13, 297–306.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, 4, 137–166.

Received May 9, 2002

Revision received April 14, 2003

Accepted April 24, 2003 ■

United States Postal Service
Statement of Ownership, Management, and Circulation

1. Publication Title: Journal of Educational Psychology

2. Issue Frequency: Quarterly

3. Issue Date: October 2003

4. Issue Number: 4

5. Number of Issues Published Annually: 4

6. Annual Subscription Price: Print \$65, Online \$127, Print & Online \$127

7. Complete Mailing Address of Known Office of Publication (Street, city, county, state, and ZIP+4):
750 First Street, N.E., Washington D.C. 20002-4242

8. Complete Mailing Address of Headquarters or General Business Office of Publisher (Not printer):
750 First Street, N.E., Washington, D.C. 20002-4242

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor (Do not leave blank):
 Publisher: American Psychological Association, 750 First Street, N.E., Washington, D.C. 20002-4242
 Editor: Benjamin D. Zumbo, Department of Special Education, Benjamin Building, University of Maryland, College Park, MD 20742
 Managing Editor: Susan Knapp, American Psychological Association, 750 First Street, N.E., Washington, D.C. 20002-4242

10. Owner (Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the name and address of the individual owner. If the publication is published by a partnership or other unincorporated firm, give the name and address of each individual owner. If the publication is published by a proprietor, give the name and address.)
 Full Name: American Psychological Association, Complete Mailing Address: 750 First Street, N.E., Washington, D.C. 20002-4242

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box.
 Full Name: NONE, Complete Mailing Address: _____

12. Tax Status (For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one):
 Has Not Changed During Preceding 12 Months
 Has Changed During Preceding 12 Months (Publisher must submit explanation of change with this statement)

PS Form 3526, October 1999 (See Instructions on Reverse)

13. Publication Title: Journal of Educational Psychology

14. Issue Date for Circulation Data Below: September 2003

15. Extent and Nature of Circulation	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies (Net press run)	5460	5250
(1) Paid and/or Requested Circulation		
(a) Paid in Advance Subscriptions (Based on Form 3541) (Include advertiser's proof and exchange copies)	755	705
(b) Paid in Advance Subscriptions (Based on Form 3541) (Include advertiser's proof and exchange copies)	3567	3532
(c) Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution		
(d) Other Classes Mailed Through the USPS		
b. Free Distribution Outside the Mail (Carriers or other means)		
(1) Outside-County as Stated on Form 3541	4322	4237
(2) In-County as Stated on Form 3541	174	196
(3) Other Classes Mailed Through the USPS		
c. Free Distribution Outside the Mail (Carriers or other means)		
(1) Outside-County as Stated on Form 3541	174	196
(2) In-County as Stated on Form 3541	4496	4433
(3) Other Classes Mailed Through the USPS		
d. Total Free Distribution (Sum of 15b and 15c)	4666	4633
e. Total Distribution (Sum of 15a and 15d)	5460	5250
f. Copies not Distributed	96.1	95.6

16. Publication Statement of Ownership: Publication required. Will be printed in the December 2003 issue of this publication. Publication not required.

17. Department or Division, Publisher, Business Manager, or Owner: APPA, Washington, DC, American Psychological Association Date: 1/11/03

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including fines and imprisonment).

Instructions to Publishers

- Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
- In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use these spaces if more space is required.
- Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
- Item 15h, Copies not Distributed, must include (1) nonreturnable copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3) copies for office use, leftovers, spoiled, and all other copies not distributed.
- If the publication has had Federal authorization as a general or requester publication, the Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
- In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
- Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Postoffice authorization.

PS Form 3526, October 1999 (Reverse)