PRACTICING VERSUS INVENTING WITH CONTRASTING CASES: THE EFFECTS OF TELLING FIRST ON LEARNING AND TRANSFER

<Journal of Educational Psychology, in press>

Daniel L. Schwartz[1], Catherine C. Chase[1], Marily A. Oppezzo[1], & Doris B. Chin[2,3]

[1]School of Education, Stanford University

[2]Human-Sciences and Technology Advanced Research (H-STAR) Institute, Stanford University

[3]*Corresponding author:*
Doris B. Chin
H-STAR Institute
450 Serra Mall, Building 160
Stanford, CA 94305-2055
Email: dbchin@stanford.edu
Phone: 650-736-1514

Abstract

Being told procedures and concepts before problem solving can inadvertently undermine the learning of deep structures in physics. If students do not learn the underlying structure of physical phenomena, they will exhibit poor transfer. Two studies on teaching physics to adolescents compared the effects of "telling" students before and after problem solving. In Experiment 1 (N=128), students in a tell-and-practice (T&P) condition were told the relevant concepts and formulas (e.g., density), before practicing on a set of contrasting cases for each lesson. Students in an invent-with-contrasting-cases (ICC) condition had to invent formulas using the same cases and were only told afterwards. Both groups exhibited equal proficiency at using the formulas on word problems. However, ICC students better learned the ratio structure of the physical phenomena and transferred more frequently to semantically unrelated topics that also have a ratio structure (e.g., spring constant). Experiment 2 (N=120) clarified the sources of the effects, while showing that ICC benefitted both low- and high-achieving students.

Keywords: transfer, science education, physics education, contrasting cases, inventing, proportional reasoning, tell and practice

In a review of the Third International Mathematics and Science Study (TIMSS), Hiebert and Stigler (2004) noted that instruction in the United States largely takes a form that we will label “Tell-and-Practice” (T&P). Teachers or texts first explain concepts and their formulaic expression, and then students practice on a set of well-designed problems. It is a convenient and efficient way to deliver accumulated knowledge. Nevertheless, very many scholars are working on instructional alternatives. Catrambone (1998) summarizes a prevailing concern with T&P, “Students tend to memorize the details of how the equations are filled out rather than learning the deeper, conceptual knowledge” (p. 356).

Many T&P alternatives use some form of guided discovery, including problem-based learning (Barrows & Tamblyn, 1976), project- and design-based activities (Barron et al, 1998), inquiry (Edelson, Gordin, & Pea, 1999), and modeling (Lesh & Doeer, 2003). The mechanics of these alternatives withhold didactic teaching at first, lest it undermine the processes of discovery. However, there are unresolved debates on the timing of explicit instruction (see Kirschner, Sweller, & Clark, 2006; Tobias & Duffy, 2009). Sweller (1988), for instance, proposes that withholding explicit instruction incurs a needless cognitive load that detracts from learning. In contrast, Bransford, Franks, Sherwood and Vye (1989) propose that students need to first experience the problems that render told knowledge useful. The persistence of these debates has been attributed to the lack of experimental evidence (Mayer, 2009). One reason for the lack of evidence is that T&P and guided discovery can differ on so many dimensions that it is difficult to maintain fidelity to their respective instructional models while isolating the causal differences between them.

For example, Schwartz and Bransford (1998) asked students to invent graphs to characterize simplified data sets from classic psychology experiments. The students learned and transferred more from a subsequent lecture than other students who first summarized an explicit chapter on the experiments and then heard the same lecture. Schwartz and Bransford (1998) argued that there is a "time for telling." For novices, explicit instruction is effective if they have been prepared with appropriate experiences. But if they have not been prepared, then explicit instruction provides much less value. Supporting this claim, Schwartz and Martin (2004) found that students who invented a variance formula for a set of specially designed cases exhibited superior transfer from a subsequent worked example, as compared to students who instead completed T&P instruction before the worked example. In response to these findings, Sweller (2009) correctly questioned "the appropriateness of the control groups used in these studies. If multiple factors are varied simultaneously, as they were in these experiments, does this procedure not break the 'vary one thing at a time' rule essential to all randomized, controlled experiments, resulting in it being impossible to determine exactly what caused the effects?" (p. 53). In other words, it is impossible to know if the results were due to the specially designed cases, the timing of explicit instruction, or both.

The following studies highlight a guided discovery pedagogy called Inventing with Contrasting Cases (ICC). The contrasting cases are designed to help students find "deep structure" (we say more below). The contrasting cases can also serve as practice problems for T&P. Therefore, they provide an opportunity to maintain both instructional fidelity and experimental control. In the following two studies, all conditions used the same sets of contrasting cases. The principal difference between conditions was that T&P students were told the concepts and solution methods beforehand, and then practiced with the materials. Students in the ICC condition invented their own solutions for the same materials and were not told about the conventional solutions and concepts until the end. Given the close comparison, the current experiments begin to identify the specific effects of telling first on subsequent learning and transfer compared to one form of guided discovery.

Inventing with Contrasting Cases

ICC is similar to other instructional approaches that ask students to generate symbolic representations or models of data (e.g., Bowers, Cobb, & McClain, 1999; Lehrer & Schauble, 2004; Lesh & Doerr 2003). The unique aspects of ICC are that the inventing tasks take a few minutes rather than days or weeks, and they use pre-defined contrasting cases. We start by explaining the contrasting cases. Afterwards, we describe the inventing activity.

Contrasting Cases

Contrasting cases, which originally came from work in perceptual learning (Gibson & Gibson, 1955), are instructional materials designed to help students notice information they might otherwise overlook. Like tasting wines side-by-side, contrasts can improve discernment (e.g., Biederman & Shiffrar 1987; Bransford, Franks, Vye, & Sherwood, 1989; Eisner, 1972). Rittle-Johnson and Star (2007, 2009), for example, found that students developed more procedural flexibility in algebraic manipulations if they compared alternative solutions side-by-side than if they saw them sequentially. By contrasting different solutions, students noticed what made each procedure unique and useful.

[Figure 1 here – Clown Worksheet]

In addition to supporting procedural flexibility with symbolic representations, contrasting cases can be configured to highlight empirical regularities. Figure 1 shows a set of contrasting cases designed to help students learn about density. The narrative of the worksheet is that each row represents a company that ships clowns to events. A given company always packs its clowns into busses by the "same amount," but each company packs their clowns to different degrees.

The systematic variability across the cases is intended to help students notice the ratio structure of density. Gick and Paterson (1992), for example, demonstrated that the inclusion of "near misses" – contrasts that differ on a single dimension – improve schema induction and transfer. In Figure 1, the busses in the upper- and lower-right corners have the same number of clowns, but vary in size (number of bus compartments). This should help students notice that bus size is important, and not just the total number of clowns.

The cases in Figure 1 contain at least three levels of structure. The first level is the surface features, or irrelevant structure. Two examples of surface features are the type of clown and the lines delineating the bus exterior. These incidental details are irrelevant to the concept of density. If this were the only information students gained from the cases, they would not have learned much about density.

A second level of structure is the density used *within* a given company. For example, the company in the first row uses two busses with different numbers of clowns and bus sizes, but both busses use a density of one clown per compartment. This level of information better instantiates the idea of deep structure, because the density is a fixed relation between mass and volume, even though the amounts of mass and volume vary.

The third level is the structure of ratio as it occurs *across* the companies. While the specific densities differ for each company, all three use proportionate ratios. This last level of structure is termed the "invariant under transformation" (Gibson, 1969). The invariant of a ratio structure persists in all the cases, despite the varying ratios for each company.

An invariant under transformation is a specific class of deep structure that differs from the qualitative relations (Forbus, 1984) often investigated in research on analogical transfer (e.g., an atom is like the solar system, Gentner & Markman; 1997). For this class of deep structure, there is a lawful parametric relation (ratio) that must be preserved across changes to the quantities, and a given ratio is defined by its position along a continuum of possible ratios. The contrasting cases in Figure 1 are designed on the assumption that picking up the invariant of a ratio structure across situations is the key to effective transfer, and that without this recognition, it seems unlikely students would transfer to other ratio phenomena such as speed or springiness.

Inventing

Students need a productive task orientation to benefit from the contrasting cases. The literature provides several good examples. Williams and Lombrozo (2010) propose orienting students with an "explanation" directive. In their research, participants received artificial stimuli that comprised contrasting cases of category instances. Participants who were asked to explain category membership exhibited better learning and transfer compared to thinking aloud or simply describing the instances. Similarly, Schmidt, De Volder, De Grave, Moust, and Patel (1989) asked students to work in small groups to explain the contrasting case of a blood cell that expands in water but shrinks in salt water. These explanation activities prepared students to learn from a relevant text compared to students who had not done the group explanation activities. Taking a different approach, Gentner, Lowenstein and Thompson (2003) found that students who received directives to draw analogies across negotiation strategies transferred more than students who did not.

In ICC, students are asked to invent a quantitative index. For the cases in Figure 1, students invent an index for density, disguised as a "crowdedness index" for each company.

They have to come up with a single, common way to compute a consumer index that works for each company, so that consumers can compare the crowdedness used by each company (because crowded clowns are grumpy clowns). The directive to make an indexing scheme differs from asking students to explain, draw an analogy, or simply compare and contrast, because it specifically pushes students towards a singular, quantitative explanation of the deep structure. Ideally, the inventing task also recruits a confluence of productive psychological processes including explanation, analogy, and compare and contrast. One question addressed by the current research is whether inventing is a useful task orientation, or whether it would be simpler and more direct to just tell students about density and have them practice with the cases.

Learning Predictions

In debates over the timing of explicit instruction, studies that have found advantages for "telling first" have focused on the learning of multi-step procedures. For example, Klahr and Nigam (2004) found an advantage for explicit instruction over unsupported discovery for the control-of-variable procedure. Tuovinen and Sweller (1999) found that worked examples improved database programming compared to free exploration, especially for novices. While unsupported exploration has low fidelity to most models of guided discovery, these results indicate that explicit instruction can improve procedural knowledge. However, success at procedural learning does not address a major challenge for many topics of instruction. In science, students also need to learn the structure of scientific phenomena.

Our hypothesis is that explicit instruction beforehand may not lead students to learn the deep structure very effectively. When learning science, students often focus on the surface features of single instances rather than the deep structures that define a phenomenon across instances (e.g., Goldstone & Son, 2005). For example, students may hear about density and learn its formula, and then facilely map the variables of the formula to discrete features of the subsequent problems. In short, they may simply divide the relevant values to compute density. Though successful, they may not understand the significance of ratio in the structure of density. Thus, we are making the possibly unintuitive prediction that telling students the structure of a phenomenon, in this case with a verbal explanation and an encapsulating mathematical formula, may limit their learning of the structure from subsequent practice problems.

Initial support for this hypothesis comes from the previously mentioned "time for telling" studies (Schwartz & Bransford, 1998; Schwartz & Martin, 2004). Kapur (2008, 2010) also compared "delay of structure" treatments with T&P instruction in a series of studies on Newtonian physics and rate-time-distance problems. In the delay of structure treatments, students did not initially receive explicit instruction that ensured they could solve the problems. The students often failed, but it was a "productive failure," because the students exhibited better overall learning. One interpretation is that the T&P students, who were told how to solve the problems beforehand, had no need to search for the deep structure. In contrast, the delay of structure condition may have led students to search for the deep structure so they could solve the problems.

To gather more direct evidence on the effects of T&P for the learning of structure, the following studies took a simple measurement approach. Students practiced or invented with a set of contrasting cases, and a day later, they redrew the cases. Our first prediction was that T&P students would *not* reproduce the ratio structure of the problems as well as ICC students who received the same problems without being told first.

A second prediction was that T&P students would transfer poorly. Gick and Holyoak (1983) demonstrated that if people do not learn the deep structure, they rarely exhibit spontaneous transfer to problem isomorphs with different surface features. Therefore, we predicted that at posttest, T&P students would transfer relatively poorly to scientific phenomena that involve ratio but not density (e.g., spring constant), despite being able to solve procedural problems on density. People can have knowledge of a mathematical formula, or any formal theory, without recognizing the structure it describes. For example, Michael, Klee, Bransford, and Warren (1993) showed that clinical psychology students could recite the relevant theories but could not apply those theories to patients. They had not learned to recognize the structure of symptoms in patients.

Experiment 1

This study examines 8th-grade students' learning of structure and their spontaneous transfer from lessons on density and speed to problems on surface pressure and the spring constant. In terms of physical phenomena, these four are quite different. Nevertheless, they still share the deep structure of a ratio between distinct physical properties. Density is described as mass over volume; speed is distance over time; surface pressure is force over area; and the spring constant is force over displacement.

Ratio is a critical deep structure for many physical phenomena. By adolescence, most children have the cognitive wherewithal to reason about ratio (Siegler, 1981), but this does not mean they spontaneously recognize or learn the ratio structure in novel situations. If students do not pick up the ratio structure during instruction, they should do poorly on transfer problems that have a ratio structure. In contrast, helping students understand the prevalence of the ratio structure should open up a deeper understanding of many physics concepts.

Methods

Participants.

Participants (*N*=128) were eighth-graders from four classes at a diverse middle school (35% Asian, 25% Latino, 22% Filipino, 11% White, 4% African-American and 3% other; 37% qualify for free lunch programs). Within each class, students were assigned to the T&P or ICC conditions through stratified random assignment based on their cumulative class scores in science. Treatments occurred in separate rooms with rotating instructors to ensure there were no teacher effects (each instructor taught both T&P and ICC for each lesson). In both conditions, students worked in self-elected and sometimes-changing pairs, consistent with regular class practice. All tests were taken individually.

Design and Instruction.

This section describes the complete treatments for both conditions. The experiment had two phases that served to test the two predictions. The first phase tested the prediction that T&P students would not learn the structure of density as well as ICC students. The second phase of the study tested the prediction that T&P does not support transfer as well as ICC instruction. Figure 2 shows the full timeline for instruction throughout the study. Measures of student learning were interleaved at four points (shown centered and bolded).

[Figure 2 here – Study timeline]

Phase 1: On Day 1, T&P students received an instructional page titled "Finding Density" (see Figure 3a). It contained some everyday examples, the formula, and a worked example that was different from the contrasting cases they would receive. Students took turns reading sections of the page aloud, while the rest of the class followed along. Student questions were answered by referring them to the instruction page. Afterwards, students were told, "*On the next page, compute the density for each company that busses clowns.*" Students then practiced on the contrasting cases worksheet in Figure 1. The instructional page was kept by students for reference.

[Figure 3 here – Instructional pages]

In the ICC condition, students did not receive any density instruction beforehand, but instead, they received an instruction page titled "Inventing an Index" (Figure 3b). It described an index in everyday language and gave several examples, such as class grade and batting average. It developed the narrative of companies shipping clowns and the need for a crowdedness index. The instructional page did not provide any examples for how to compute an index, but it provided a set of constraints at the end of the page (i.e., a company only gets one index value, the same procedure should be used to find the index for each company, and a larger index means more crowded). The instructional page was read aloud and questions were answered in the same fashion as the T&P condition., Then, students were told to "*invent a procedure for computing a crowded clown index for each company.*" They used the same contrasting cases worksheet the T&P group received. The instructional page was kept by students for reference.

The critical measurement for phase 1 came on Day 2. Twenty-four hours after completing the clown worksheet and prior to any further potentially contaminating instruction or experiences, students were asked to reconstruct the worksheet. This "memory test" provided a measure of the structure they had learned from the previous day's lesson (Chase & Simon, 1973; Jee, Gentner, Forbus, Sageman, & Uttal, 2009). Students received a sheet that stated: "*Yesterday, you received a sheet that had an activity with clowns and busses. Use the space to redraw how the sheet looked the best you can.*" Students were given 10 minutes to complete their drawings. We coded two aspects of their recreations. The first was whether students reconstructed the ratio structures. We call this "deep structure" recall, and students received 1 deep structure point for each pair of busses that shared one ratio between them, unique from the other pairs of busses (3 pts = 100%). The ratios did not have to be the same as the original worksheet, but they had to be proportional across pairs of busses. The second coding captured students' memory for surface features. Students received 1 point for each of the following surface features that we had *a priori* seeded into the cases (6 pts = 100%): elaborated clown features; clowns positioned on lines between bus compartments; different line styles for the busses; wheels that did not correspond to the compartments; a company name; and, incidental text features such as "Name __________." Our prediction was that surface feature recall would be the same in both conditions, but that structural recall would be worse in the T&P condition. This would indicate that the effect of T&P is specific to deep structure. A primary coder evaluated this and all other measures, and a secondary coder evaluated 25% of the data for each measure. Measures were blind-coded, and agreement was greater than 98% for each measure for both studies.

Phase 2: Students completed three more lessons using T&P or ICC on Days 2 and 5. The lessons were designed to cover discrete and continuous versions of both density and speed. (The full set of materials is available at http://aaalab.stanford.edu/transfer.html.) Similar to Day 1, the T&P condition always received and read aloud an instructional page that included everyday examples of the target concept, an introduction to the relevant formula, and a worked example that used a different cover story from the contrasting cases. The ICC instructions also kept to their format from Day 1. Students were asked to invent an index for the contrasting cases. However, it was no longer necessary to explain the concept of "inventing an index" or introduce the constraints on the index (e.g., each company only gets a single index score). After their respective instructional pages, students received the worksheets, which had the same data and contrasting cases regardless of condition. The insets of Figure 2 show one case from each of these worksheets. The popcorn task had students find the popping rate of different machines (discrete speed). The gold task had students find the purity of gold used by different companies (continuous density). The racecar task had students find the speeds of different cars and further determine which cars have the same speed (continuous speed) and therefore come from the same company. (Note that this additional instruction for the cars task requests the T&P and ICC students to explicitly compare across cases for the first time.)

In addition to the instructional pages, a second treatment difference involved the timing of an explicit lecture on ratio, shown in bold italics in Figure 2. T&P students received a lecture on the importance of ratio in physics at the beginning of phase 2, on Day 2. This lecture was separate from the instructional pages the T&P students received before each of the remaining three lessons. The lecture directly explained the prevalence of ratio in physics including density, speed, force, and more, and it explicitly indicated the analogous ratio structures in the clown worksheet, an approach demonstrated to support transfer (Brown & Kane, 1988; Gick & Holyoak, 1983; Gentner, Loewenstein, & Thompson, 2003). The lecture helped to evaluate the common intuition that if one just told the students that ratio was important and explicitly pointed out, they would begin to learn the structural information from the subsequent contrasting cases and use it later at transfer.[1] ICC students received this same lecture, but on Day 8, after they had completed all the inventing tasks. The lecture also served as the ICC students' explicit instruction on the concepts of density and speed.

The last piece of instruction for both groups was a worksheet of word problems, given on Day 8. This gave students practice applying the density and speed formulas. Across the four days of instruction, the ICC instruction required approximately 10 more minutes of instructional time. The T&P students used this time completing extra word problems on this last day.

To test the effects of T&P and ICC on transfer, three learning measures were taken during phase 2. The first, an immediate transfer task, was given at the beginning of Day 8, after ICC students had completed the four ICC cycles but had not yet received the ratio lecture. This permitted us to evaluate transfer for the ICC students *before* they had received any direct instruction, whereas the T&P students had completed all their instruction (except practice on word problems). Students received a sheet that showed diagrams of four aerosol cans with internal springs pressing on plates. Students were asked to “describe” the plate pressure used in the aerosol cans. The word “describe” had not been used previously in either treatment. Students received 1 pt for each instance that was described with a ratio (e.g., springs over plate area; 4 pts = 100%).

The second measure was a word problem test given on Day 11. The test comprised 3 problems each for density and speed. An example problem is: *Brenda packs 120 marshmallows into 4 soda cans. Sandra packs 300 marshmallows into 11 soda cans. Whose soda cans are more densely packed?* This test served two purposes. First, as is characteristic in transfer research, it is important to ensure that a lack of transfer is not due to a simple failure to learn the concept or procedure. Therefore, we wanted to be sure the T&P students had learned how to compute ratio quantities and could do so in the context of solving standard word problems. The second purpose was to find out whether the ICC condition put the students at risk for not learning how to apply the formulas, given that they did not “practice” on the contrasting cases and received much less overall direct instruction on the density and speed concepts and their formulas.

[Figure 4 here]

The third phase-2 measure was a delayed transfer task that occurred 21 days after all students had completed their instructional treatments. Ideally, this delay would determine whether the effects of the treatments were somewhat lasting, and it would determine if the final lecture helped or hurt the ICC students. All students attempted the transfer problem embedded in a biology test. Figure 4 shows that the transfer problem was on the spring constant. Students were asked to "*determine the stiffness of the mat fabric for each trampoline*." Students received 1 pt for each trampoline described by ratio (e.g., number of people over displacement; 4 pts = 100%).

<u>Results</u>

[Figure 5 here – samples of clown recall]

<u>Redrawing Test: Encoding of Surface Features and Deep Structure.</u>

Figure 5 shows excerpts of student recreations of the clown worksheet. The examples show what surface and deep feature recall looks like for this task. The left columns of Table 1 show the average percentages of recall for deep structures and surface features in Experiment 1. The T&P students did not recreate the deep structure as well as the ICC students. There was a strong interaction between treatment and the type of feature recalled; $F(1,120) = 11.9$, $p < .001$. There was a treatment effect for deep structure; $F(1,120) = 11.0$, $p = .001$, $d = .30$, but not for surface features; $F(1,120) = 1.5$, $p = .223$, $d = .11$. There were no significant correlations between surface and deep recall; $r_{ICC} = -.21$, $p = .10$, $r_{T\&P} = .09$, $p = .52$. Interestingly, 19% of the T&P students wrote down the formula, which was not present on the original worksheet, compared to only 3% of ICC students. Recall of the ratio structure did not differ appreciably for those T&P students who wrote down the formula (M = 57%, SD = 33.6%) versus those who did not (M = 51%, SD = 43.9%); $F(1,56) = 0.2$, $p = .65$. This result highlights that remembering the structure of the formula does not entail knowing the structure of the phenomena to which the formula refers.

[Table 1 here – means from Exp. 1 & Exp. 2]

<u>Transfer Effects.</u>

By the first transfer task on surface pressure, T&P students had received lessons on discrete and continuous versions of density and speed, and they had received a separate lecture on the importance of ratio in physics that explicitly mapped the analogy across density, speed, and other physics concepts. For each lesson, they had received a unique worked example and then practiced on a new set of cases. The means in Table 1 indicate that the T&P students used ratio to describe surface pressure significantly less often than the ICC students, who had not yet been told about the formulas or the importance of ratio; $F(1,94) = 6.8$, $p = .011$, $d = .27$. (Due to an implementation error, data from both of the conditions in one class could not be included in the preceding analysis of immediate transfer. The delayed transfer analysis, which is next, included all the students.)

[Figure 6 about here – examples of transfer performances.]

Figure 6 provides prototypical examples of how students answered the delayed transfer problem on the spring constant (trampoline). (The immediate transfer problem on surface pressure had the same classes of response.) To receive credit for transfer, students had to produce answers similar to Figure 6d, which indicates a ratio between people and rungs. The delayed transfer problem occurred 3 weeks after both groups had completed the final word problems and the ICC students had received their lecture on ratio. Table 1 indicates that the ICC students performed significantly better than the T&P students on the delayed transfer test; $F(1,124) = 13.6$; $p < .001$, $d = .33$.

Further analysis indicated that learning the ratio structure during the clown task was important for transfer, regardless of condition. (See Klahr & Nigam (2004) on common causes of transfer despite instructional differences.) A three step regression analysis used performance on the delayed transfer problem (trampoline) as the dependent measure. In the first step, structural recall of the clown task was regressed against transfer performance; $F(1,119) = 5.79$, $p = .018$, $r = .22$. Students who recalled the deep structure of the clowns were more likely to transfer. In the second step of the analysis, experimental condition was added to the regression equation. It improved the prediction of the model, $F_{change}(1, 118) = 8.84$, $p = .004$, $r_{model} = .34$. Thus, students in the ICC condition were more likely to transfer, even when taking into account their structural recall of the clown worksheet. One possible explanation is that ICC students who did not induce the ratio structure during the clown activity learned it over the three subsequent inventing activities. Finally, in the third step, the interaction term of condition by clown recall was added to the equation. It did not improve the fit; $F_{change}(1, 117) = 0.91$, $p = .342$, $r_{model} = .35$. The lack of an interaction indicates that finding the clown structure improved transfer for both conditions. The primary difference was that students in the ICC condition found the structure more often. Overall, these results should not be interpreted as demonstrating that finding the structure of the clown worksheet was sufficient for transfer. A more cautious interpretation is that those who found the structure on the first lesson were more likely to find structure on the second, third, and fourth lessons, which would have a cumulative effect for improving transfer.

One potential concern with the focus on ratio is that students may have been learning math, but they may not have been learning physics. A *post-hoc* coding of the transfer problems sheds light on this concern. We coded whether students ranked the trampoline cases correctly in terms of their physical property of stiffness. The instructions did not have an explicit directive to rank the cases, but the students implicitly did so, for example by writing "stiffest," or a specific ratio, next to a trampoline. Students who used ratios for over half of the cases correctly ranked the relative stiffness of the fabrics 96% of the time. Students who did not use ratios succeeded 1% of the time, χ^2 (1, N=126) = 113.7, $p < .001$. (The results are similar for the surface pressure problem; χ^2 (1, N=96) = 66.1, $p < .001$.) Students who did not use ratios tended to rank the stiffness of the fabric based on the value of a single feature such as the number of people on the trampoline (Fig. 5a) or the number of rungs the fabric stretched (Fig. 5b). Thus, without ratios, students had very little chance of forming a qualitative understanding of the physics, because they focused on only one feature in the physical situation. As an analogy, students will often conflate density and mass, because they do not appreciate that density is a structural relation between two physical properties.

<u>Word Problem Test: Computation.</u>

The transfer differences were not due to differences in students' knowledge of the formulas. Table 1 indicates that the T&P and ICC conditions exhibited similar percentages correct on the word problem test; $F(1, 124) = 1.7$, $p = .20$.

Discussion

The T&P and ICC students had learned to apply formulas to word problems equally well. Nevertheless, T&P instruction was sub-optimal for helping students learn the structures that are important for transfer. T&P students exhibited a relatively low rate of recreating the deep structure of the contrasting cases at recall, and they also transferred less often than the ICC students, both before and after the ICC students had received a lecture on ratio in physics. Our favored explanation is that many T&P students focused on applying what they had been told, and missed the deep structure in the problem situations.

The T&P students' poor recall of the ratio structure of the crowded clowns cannot easily be attributed to a lack of interest or general memory effects. If the T&P results were due to general effects, then the T&P students' memory for both surface features and deep structures should have been lower than ICC, but the T&P and ICC students demonstrated equal surface feature recall.

Regardless of condition, students who found and recalled the ratio structure of the crowded clown worksheet transferred better. However, more of the ICC students found the ratio structure, and therefore, more ICC students transferred. Under this explanation, the ICC students' deeper understanding of ratio led to the transfer. An alternative explanation is that the ICC students had "learned to learn" (e.g., Brown & Kane, 1988). They developed the strategy of looking across multiple problems to find a deep structure. By this alternative, the ICC students did not transfer the ability to recognize the ratio structure, but instead, they transferred the idea of looking for patterns given multiple cases. The transfer problems included multiple cases which may have cued the same strategy they had used during instruction. If true, it means that ICC students should not transfer the use of ratios to situations that only involve a single instance. Experiment 2 tests this prediction.

Experiment 2

What led to superior levels of transfer for the ICC students? Experiment 2 was designed to help answer this question by examining a) which behaviors helped students learn the structural information, and b) what learning led to the transfer. We assume the contrasting cases themselves were an important element that aided students in learning the deep structure. But they could not have been sufficient because the T&P students also used them. Our hypothesis is that the inventing directive orients students to search the cases for a common deep structure on which to base their index. In contrast, T&P instruction focuses students on applying solutions, one problem at a time. This reduces their chances of finding the deep structure across the cases. To investigate further, we videotaped a subset of students. We coded the videotapes specifically for how many times the students transitioned among the cases. The prediction was that students in the ICC condition would shuttle among the cases as they searched for the structure on which to base the index. In contrast, the T&P students would be more inclined to complete each case separately and sequentially, making fewer transitions between cases. As Gentner, Lowenstein, and Thompson (2003) concluded from their studies on analogical encoding, “learners cannot be counted on to spontaneously draw appropriate comparisons, even when the two cases are presented in close juxtaposition” (p. 403).

The second issue is the question of what transferred. One hypothesis is that students transferred the ability to recognize ratios. An alternative hypothesis is that students transferred the strategy of looking across multiple cases. To test these alternatives at transfer, half of the students from each condition received the original four-case trampoline problem, while the other half received the same scenario but with only one trampoline. If the ICC students transfer on the basis of having a problem with multiple cases, like those at instruction, then the ICC advantage over T&P should disappear for the one-case transfer problem. If the ICC students transfer on the basis of ratio, then they should exhibit greater transfer than T&P students for both the one- and four-case transfer problems.

A final practical question addressed by Experiment 2 is whether ICC and T&P instruction interacts with students' prior achievement. There are concerns that lower-achieving students may be better served by the greater guidance of T&P instruction. For example, Hiebert and Wearne (1996) found that early elementary students with low initial understanding did not fare very well with invention tasks when learning place value. Similarly, Tuovinen & Sweller (1999) compared adult learning through free exploration versus learning from worked examples. Free exploration was only effective for adults with high prior knowledge. Therefore, we collected permission to analyze students' cumulative class grades in science to see if it would indicate that the T&P instruction was relatively more effective for the low-achieving students.

Methods

Experiment 2 followed a nearly identical protocol with 120 eighth-grade students at the same school. There were four primary changes. (1) The experiment used instructors blind to the hypotheses. (2) Twenty-four pairs of students were selected at random to be videotaped in class while they worked on the clown worksheet (3 pairs x 4 classes x 2 treatments). (3) The timeline was condensed by removing the first transfer task on surface pressure and combining the word problem and delayed transfer test into a single posttest, given one week after instruction. (4)

Students in each treatment received either the 1- or 4-case trampoline transfer problem (spring constant) at random.

Results

The right side of Table 1 shows the results for Experiment 2. The data replicated Experiment 1 closely. There were no differences between treatments for surface feature recall and word problem performance, but the ICC condition showed a substantial advantage for deep structure recall and transfer performance. While science achievement had an overall effect on performance, it did not interact with the instructional conditions. The right-most columns show the Experiment 2 data partitioned by a median split on the students' cumulative grades in their science class. The treatment effect was sufficiently strong that lower-achieving ICC students descriptively outperformed higher-achieving T&P students at transfer, 40.9% vs. 32.9%.

To statistically dissect the possible interactions of treatment by achievement, separate ANCOVAs were performed on the four measures in Table 1. The analyses used the achievement covariate, the treatment factor, and the treatment by covariate interaction. Two students in the T&P condition did not have cumulative grades and were omitted from these analyses.

For deep structure recall, ICC significantly outperformed T&P; $F(1, 111) = 8.7$, $MSE = 0.12$; $p = .004$, $d = .28$; achievement was a significant predictor of deep structure recall; $F(1, 111) = 23.7$, $p < .001$; but, there was no achievement by treatment interaction; $F(1, 111) = 0.15$, $p = .70$. For surface feature memory, there was no difference by treatment; $F(1, 111) = 2.8$, $MSE = 0.02$, $p = .10$, $d = .15$; there was no effect of achievement; $F(1, 111) = 1.6$, $p = .22$; and, there was no interaction; $F(1, 111) = 0.09$, $p = .77$. As in Experiment 1, there were no significant correlations between surface feature and deep structure recall; $r_{ICC} = .09$, $p = .49$; $r_{T\&P} = -.18$, $p = .17$.

With respect to posttest performance, the ICC students exhibited greater transfer; $F(1, 111) = 10.6$, $MSE = 0.20$, $p = .002$, $d = .31$; there was an effect of achievement on transfer; $F(1, 111) = 9.2$, $p = .003$; but there was no treatment by achievement interaction; $F(1, 111) = 0.18$, $p = .67$. For the word problems, there were no treatment differences, $F(1, 111) = 0.21$, $MSE = 0.03$, $p = .65$, $d = .04$; there was a very strong effect of achievement; $F(1, 111) = 34.0$, $p < .001$; but again, there was no interaction of treatment by achievement; $F(1, 111) = 0.09$, $p = .76$.

In this study, we also coded the percentage of correct answers students produced for each of the four contrasting cases worksheets they received during instruction (these data were not available for Experiment 1). The T&P condition did modestly better. An ANOVA compared the average percentage correct per worksheet and demonstrated a significant effect of condition; $M_{ICC} = 84.7\%$ (SD = 22.7%), $M_{T\&P} = 91.8\%$ (SD = 12.2%), $F(1, 113) = 4.5$, $p = .037$. The following disaggregates worksheet performance across the four lessons: Clowns (discrete-density): $M_{ICC} = 85.1\%$ (SD = 31.7%), $M_{T\&P} = 91.5\%$ (SD = 23.6%), $F(1, 113) = 1.52$, $p = .22$. Popcorn (discrete-speed): $M_{ICC} = 82.7\%$ (SD = 31.8%), $M_{T\&P} = 90.4\%$ (SD = 23.2%), $F(1, 113) = 2.19$, $p = .14$. Gold (continuous-density): $M_{ICC} = 85.7\%$ (SD = 33.5%), $M_{T\&P} = 96.6\%$ (SD = 13.4%), $F(1, 113) = 5.33$, $p = .023$. Racecar (continuous-speed): $M_{ICC} = 85.1\%$ (SD = 29.1%), $M_{T\&P} = 88.7\%$ (SD = 28.8%), $F(1, 113) = 0.44$, $p = .51$. In isolation, only the Gold worksheet exhibited a significant treatment difference. This is driven by an uptick in T&P performance, which one can speculate is the result of the worksheet's overt numerical presentation of weight and volume, which in turn, made it easier for the T&P students to map in the density formula. In sum, despite doing somewhat worse on the basic classroom assignments, the ICC students did better at transfer and equally well on the word problem test. In absolute terms, the ICC students exhibited a high degree of success at the inventing task.

[Figure 7 here – one and four trampoline problems]

What learning transferred? Figure 7 indicates a consistent advantage for the ICC condition for both 1- and 4-case problems; $F(1,112) = 10.8$, $p = .001$, $d = .31$. Although the 4-case problem was solved more frequently than the 1-case problem, the difference was not significant; $F(1, 112)=2.7$, $p = .105$. Importantly, there was no treatment by number-of-cases interaction; $F(1,112) = 0.5$, $p = .481$, which indicates that the ICC students transferred their understanding of ratio. If ICC students had only transferred a strategy of inventing across multiple cases, they would not have shown a similar advantage over the T&P condition for the 1- and 4-case versions. Notably, the T&P students also benefited from the 4-case version, which makes sense, given that the T&P students were confronting a set of contrasting cases without being told a formula beforehand, similar to the ICC's original instruction.

As in Experiment 1, regardless of treatment, students who recalled the deep structure of the clown worksheet were more likely to transfer than those who did not. A 3-step regression analysis used the performance on the transfer problem (trampoline) as the dependent measure. In the first step, structural recall of the clown task was regressed against transfer performance; $F(1,111) = 14.37$, $p < .001$, $r = .34$. Students who recalled the deep structure of the clowns were more likely to transfer. In the second step of the regression analysis, experimental condition was added to the regression equation. It improved the fit of the model, $F_{change}(1, 110) = 5.13$, $p = .025$, $r_{model} = .39$. Overall, students in the ICC condition were more likely to transfer, presumably because they noticed the ratio structure on the subsequent inventing tasks. Finally, in the third step, the interaction term of condition by structural recall was added to the equation. It did not improve the model's fit; $F_{change}(1, 109) = 0.76$, $p = .385$, $r_{model} = .40$. The benefit of finding the structure of the clowns for transfer did not differ by condition. It was just that the ICC students learned that structure more often.

One possible source of the ICC advantage is that students searched the contrasting cases to find a deep structure on which to base their index. To find out if the ICC students searched across the cases more, we coded the frequency each videotaped pair transitioned their attention from one clown company (row) to another. Pointing to, writing on, or discussing a particular company was coded as a "reference" to that case. A transition occurred when at least one student of the pair shifted from referencing one company to another. The minimum number of transitions is two, one between each of the three companies. The ICC pairs transitioned an average of 20.1 times (SD = 10.3) compared to 6.0 (SD = 4.7) for the T&P pairs; $F(1, 22) = 18.7$, $p < .001$, $d = .92$. As with the full sample, the videotaped ICC pairs recalled the ratio structure of the clowns more than the videotaped T&P pairs (average of pair recall $M_{ICC} = 81.3\%$, SD = 33.0% ; $M_{T\&P} = 40.3\%$, SD = 29.7%; $F(1, 22) = 4.71$, $p = .041$).

T&P students already had the formula, and they tended to apply it to each case separately, which reduced their chances of noticing the invariant of ratio across the cases. Within the T&P condition nine of ten pairs exhibited ten or less transitions, and there was no correlation between number of transitions and a pair's average structural recall, $r(12) = .05$; $p = .87$. For the ICC condition, eight of the ten pairs showed more than ten transitions. This suggests that these eight ICC pairs were searching the cases to find the invariant property on which to base their index. For ICC, there was a non-significant negative correlation between number of transitions and structural recall; $r(12) = -.25$, $p = .38$. Past a certain threshold, more transitions did not yield better learning. For example, the ICC pair with the most transitions (42) kept searching the cases because they never discovered the invariant structure. At recall, they only recreated one ratio.

The videotapes were collected to test a specific behavioral prediction (number of transitions) rather than generate a corpus for protocol analysis. However, to enrich the picture of the process in action, we provide protocol excerpts from a pair of students from each treatment as they found their first answer for one of the clown companies. It should be noted that these excerpts were specifically chosen to highlight treatment differences. The preceding statistical analysis shows they reflect general differences, though there was natural variation across the pairs. With each transcript, we provide a brief commentary pointing out the key features from our perspective. For clarity, companies are referred to as Company A, B, or C, from top to bottom on the Figure 1 worksheet.

The following transcript reflects the T&P students' tendency to focus on the formula such that the task becomes a problem of mapping the variables of the formula to the discrete features of a specific case. Notice that these students begin by explicitly stating the formula "D equals M over V" and then work from there. These students, like others, spend a good deal of time figuring out that each compartment (rather than a complete bus) represents one unit of volume. Moreover, these students make only one transition – from company C to company B – throughout the entire segment.

<u>A T&P Pair.</u>

Student 1: Okay, this one kinda makes more sense [points to company C]. Okay, this one, I think is 2 objects. There's 6 things in here, so…

Student 2: Or are there 3?

Student 1: Oh, ok, I see, I see… I'm gonna start with this one first [company C].

Student 2: Let me write it. D equals M over V. [writes on sheet]

Student 1: OK. Well… I'm just going off how… what she [the instructor] did up there. So there's 2 objects there [points to company C]… and add.

Student 2: Do we put this in there? [points to company C answer line]

Student 1: No, I'm just doing it [the compartments] together, rather than separate. So, hold on here.

Student 2: The mass is how many things are inside of it.

Student 1: So together there's 6, individually there's 4. So I think… OK, it's for *each* bus [compartment].

Student 2: So this [compartment] is separate?

Student 1: Think so? So each one… are separate ones, yeah.

Student 2: But why is it dashed line?

Student 1: I have no clue. The same thing right here. Hmm, well. So I don't think…

Student 2: [points to company B] This one 1, 2, 3, 4, 5, 6, 7, 8, 9, and 3 busses [compartments] plus one bus [compartment] is 4, so 12, so 12 divided by… 4, is 3.

Student 1: There you go! So 3 here.

Student 2: Three objects per cube.

The following transcript, of about the same length, from an ICC pair illustrates how they made far more transitions as they looked across cases to notice critical features and tried to develop an indexing scheme.

An ICC Pair.

Student 3: So we're writing down a number?

Student 4: Yeah we're writing down a number, a value.

Student 3: Do we count the clowns or do we count the buses? 'Cause this one's separate from the other ones, I think…

Student 4: I think… lemme read the… [flips to instructions then back to worksheet]. OK, so it's not about the number of clowns, it's about how crowded it is. Which one do you think is the least crowded?

Student 3: [points to company A] It's a bigger bus.

Student 4: The most crowded?

Student 3: [points to company C, then company B] One of these two?

Student 4: I think it's this one. [points to company C]

Student 3: Because, if these two get added in, they're really crowded [points to company C] and this one too… [points to company B]

Student 4: Let's see. 1, 2, 3, 4 -- 4 in 2 sections [pointing at company C]. 1, 2, 3, 4, 5, 6, 7, 8, 9. [pointing at company B]

Student 3: There's 3. [points at company B]

Student 4: There's 3 in each section [points to company B], this is 2 in each section [points to company C], so this seems more crowded [points to company B]. So what do you think we should do?

Student 3: Should we add these guys in here? [points at company C]

Student 4: Um, no I think that's good. I think this is less [points to company A], this is most [points to company B], this is middle [points to company C].

The group has a total of 12 transitions within this segment of conversation, reflecting how they moved back and forth from one case to another. It is useful to note that the transitions were in the service of finding a single account. They were not simply comparing and contrasting the cases, but rather, they were driving towards a single mathematical explanation of crowdedness. In comparing companies to figure out which was the “most crowded,” the students realized that they could not discriminate the crowdedness level between companies B and C without counting. It is only after counting the clowns and bus compartments that they decide company B is the most crowded. Finally, they checked their indexing procedure by their intuitive sense of crowdedness. Their quantitative activity and their intuitive sense of crowdedness worked together to sharpen what they learned from the cases.

Discussion

As in Experiment 1, the T&P and ICC groups performed similarly on the word problem test and remembered similar numbers of surface features on the clown recall task. The ICC group recalled the density structure better and exhibited higher rates of transfer. Performance on the clown recall task was a good predictor of success on the delayed transfer task, regardless of condition. The ICC group was just more likely to understand the structure by the first day of instruction. Interestingly, the ICC group was slightly less successful than the T&P group in solving the contrasting cases, indicating that initial rates of success at the learning activity do not necessarily predict later transfer rates (cf. Kapur, 2008).

What transferred? The ICC advantage over T&P did not increase for the four-case problem. Instead, the ICC students maintained the same relative advantage for the single-case problem. This makes it likely that they transferred an understanding of the ratio structure, rather than being dependent on a problem with multiple cases.

How did students come to understand the underlying structure in the first place? The video analysis demonstrated that students in the ICC condition made far more transitions between the cases as they searched for common structure. The T&P students, on the other hand, went through the cases in a linear fashion, mapping the formula to each company in turn.

The relative effects of the ICC instruction compared to T&P did not differ systematically for individuals of different levels of achievement. Low-achieving students benefitted from ICC compared to T&P as much as high-achieving students. It is useful to note that students were working in pairs, so it is still possible that working individually would lead to a prior achievement by treatment interaction.

General Discussion

Standard tell-and-practice instruction is important because it delivers the explanations and solutions invented by experts, and students need opportunities to hear and practice these ideas. To gain this benefit without undermining transfer, the current studies suggest expositions should happen <u>after</u> students have explored novel-to-them deep structures. The inventing activity can serve as "preparation for future learning" by readying students to more fully appreciate the expert solutions and deep structures when they are explained (Bransford & Schwartz, 1999). Giving students the end-product of expertise too soon short-cuts the need to find the deep structure the expertise describes. Students in the T&P condition focused on what they had been told, and they applied the formulas sequentially to the problems, which reduced their chances of finding the deep structure. Without an appreciation of deep structure, students are less likely to see the structure in new situations that differ on the surface, and they will fail to transfer.

The current studies tried to address the challenge of maintaining instructional fidelity to different pedagogies while isolating a causal variable. The difficulty of doing both simultaneously may be one reason why much of the transfer research is predicated on comparing T&P instruction to itself. This makes it possible to vary only one thing, while maintaining fidelity to the dominant model of instruction. We sampled transfer articles of math and/or science learning published between 2003 and 2008.[2] Among the 70 articles (136 unique authors) that resulted from our search, 75% of the studies used T&P instruction for both treatment and control conditions while investigators manipulated other variables. T&P instruction was not a variable in these studies, and therefore not suspected as a potential contributor to transfer success or failure.

As the current studies show, the model of instruction is a large contributor to transfer, and therefore, many of the psychological claims based on transfer studies may not generalize beyond T&P instruction. For example, consider the relation between abstract and concrete elements of instruction. Some scholars favor abstract presentations to avoid obscuring problem structure (e.g., Bassok & Holyoak, 1989; Harp & Mayer, 1998; Kaminski, Sloutsky, & Heckler, 2008), whereas others favor concrete instances to connect to prior knowledge (Goldstone & Sakamoto, 2003; McNeil, Uttal, Jarvin, & Sternberg, 2009). Many of the relevant studies, however, have used T&P instruction for both conditions. By hypothesis, T&P instruction does not help students pick up the deep structure, so it makes sense that surface features would appear to be an issue in many studies. In the current studies, if surface features interfered with picking up deep structure, then there should have been a negative correlation between surface feature and deep structure recall (e.g., Rothkopf & Billington, 1979), but there was not. Moreover, Schwartz, Chase, and Bransford (in press) had students work with either the crowded clown worksheet or an analogous abstract worksheet that used dots in cubes. For T&P instruction, the abstract worksheet led to better structural recall, consistent with other studies that have used T&P instruction (e.g., Bassok & Holyoak, 1989). However, for the ICC instruction, the students performed the same for either version of the worksheet, and the ICC students doubled the rate of structural recall found for the T&P students who received the abstract worksheet. These examples demonstrate the risk of generalizing psychological claims, for example about the value of abstract examples, without taking into consideration the broader instructional context.

Alternative Hypotheses

On the relatively low performance of T&P.

Our execution of the "telling" portion of T&P was consistent with prevalent classroom practice (Hiebert & Stigler, 2004), though we fortified the approach with multiple worked examples, practice across far analogies, and a lecture that explicitly described the importance of ratio and further pointed out the ratio analogy across separate physical domains. Our goal was to maximize the reach of the findings across current educational practices by demonstrating a natural psychological response to being told the form of an answer before working on problems. Nevertheless, it is possible that our results will not generalize across all variants of direct instruction (see Atkinson, Derry, Renkl, & Wortham, 2000 for a review). For instance, direct instruction may be relatively effective for teaching discrete facts or procedural steps, whereas density involves a principled relation between dimensions (volume/weight). More pertinent to the current studies, it seems reasonable that if the teachers had shown the general ratio structure in the clown worksheet before students had worked on it, the students would have exhibited better recall of the deep structure.

Showing students the ratio structure may improve structural encoding for the specific problems, but it may not transfer as well as having students induce the structure on their own. In the current studies, T&P students were explicitly told and shown the ratio structures in the lecture after the clown worksheet. However, they did not transfer very well, despite having three more lessons that depended on ratios. Perhaps T&P students were less able to recognize the ratio structure on the transfer problems because they never had the experience of actively recognizing the ratio structure on their own (Roll, 2009). Another possibility is that students in typical classrooms, compared to laboratory settings, may not allocate critical lessons on structure the special attention they deserve. They may be more interested in just learning how to solve the problem.

With regard to the "practice" portion of T&P, the task of computing answers to problems is fairly typical, and students had further practice on subsequent word problems. The formulas for density and speed are relatively straightforward equations. The T&P students were able to apply them effectively to the T&P cases and posttest. This leads to the possibility that the practice was so easy for the T&P students that they did not have to do very deep processing, and therefore they learned less. Perhaps T&P becomes more effective when tasks become more difficult. Before embracing this hypothesis fully, it is important to note that the T&P students engaged in a good deal of germane cognitive effort as they tried to map the quantities to the variables of the equation (see transcript in Exp. 2).

A second issue is that telling first may be more effective for learning complex procedures, although less effective for complex structures. The current studies do not bear on this hypothesis except in demonstrating that early explicit instruction in science, which is designed to help students handle procedural complexities, runs the risk of focusing student attention on the procedures themselves at the expense of the structures the procedures were designed to encapsulate. There are alternatives to T&P for teaching complex procedures. Schwartz and Martin (2004), for example, demonstrated that an ICC sequence better prepared students to understand the complex formulas for statistical variance compared to procedural instruction at the outset.

On the relatively high performance of ICC.

These studies were not meant to isolate the causal effect of each ingredient in ICC instruction. Here, we offer some speculation about these ingredients, plus predictions about possible studies to test those ingredients. We assume that the contrasting cases are one crucial component because they can help students find structure in phenomena. Without these specially designed cases, which highlight critical similarities and differences, it would be difficult to find the underlying structure. A simple study might compare the outcomes of inventing for the current crowded clown worksheet (Fig. 1) versus a work sheet where all the companies use the exact same ratio (with no built-in contrasts). Our prediction is that the latter condition would not fare as well.

Another key ingredient is the behavior triggered by the directive to come up with a single account, or index, for all the cases. The protocols of student problem solving showed that ICC students revisited each company about seven times on average. In the process, the students were actively comparing and contrasting the cases as they transitioned back and forth between them. The benefits of comparison for transfer have been well-documented (e.g., Gentner, Lowenstein, and Thompson, 2003), and it is presumably necessary for the effect of ICC. But, it may not be sufficient. A useful study would compare how students perform in a compare and contrast condition versus an ICC condition. Our prediction is that if students were simply instructed to compare and contrast the cases without driving towards a single parametric explanation, the students would compare and contrast familiar features rather than finding the underlying structure that makes apparent differences the same.

A final ingredient for the current instance of ICC is the quantitative information that helped students to learn the ratio structure with greater precision. Without countable quantities, the students would probably make comparisons with terms like “more” and “less,” which lack the precision required to find the ratio structure. Schwartz, Martin, and Pfaffman (2005), for example, found that children who were asked to give verbal explanations to balance scale problems did not learn to relate the dimensions of weight and distance. In contrast, children who were asked to invent mathematics to explain their answers were more likely to discover that proportionate ratios determine whether the scale balances. Without the precision and demand of quantitative reasoning, students may gloss over important specifics and possible relations. For instance, we would predict that students would not learn as well if they were asked to explain the companies instead of inventing a quantitative index.

The brief inductive activities of the ICC physics lessons did not cover many important aspects of density and speed. One concern is that the mathematical focus of inventing may have been at the expense of qualitative understanding. Our assumption, however, is that helping students learn the deep structure prepares them to transfer to learn more about these concepts later. For instance, understanding density as a ratio of mass to volume should help students understand buoyancy, where density, mass, and volume are frequently conflated. Thus, a useful study would determine whether this is true.

The ICC instruction did not improve strategic knowledge for finding patterns across multiple instances, as demonstrated by the lack of a special benefit for the 4-case transfer problem over the 1-case problem. However, the current implementation did not have the time to emphasize the development of strategic knowledge and scientific dispositions (Gresalfi & Cobb, 2006). Taylor, Smith, van Stolk, and Spiegelman (2010) found that students who received a full course of inventing activities in college-level cell biology were more able to generate explanations for novel cell phenomena compared to T&P students.

Conclusions

Practical Applications of ICC

There are different types of learning that range from skill acquisition to identity formation, and it seems unlikely that a single pedagogy or psychological mechanism will prove optimal for all types or learning. ICC is one among many possible ways to support students in learning deep structure. Other alternatives to T&P include problem-, project-, and inquiry-based instruction, as well as approaches that attempt to "problematize" tasks so that students will not take the eventual solutions for granted (Bransford, Franks, Vye, & Sherwood, 1989; Fensham & Kass, 1988; Hiebert et al., 1996; Limon, 2001; Needham & Begg, 1991). All these approaches will support productive transfer to the extent they help students find the deep structure that generalizes across situations. The ICC activities differ from many of these pedagogies, however, because ICC strategically precedes standard T&P pedagogy rather than replacing it. Moreover, the ICC activities do not sacrifice overall classroom efficiency for the sake of transfer. The ICC treatment required about 10 extra minutes of instruction across four days, which was mostly spent on explaining the novel task of 'inventing an index.'

The inventing task differs from "discovery" tasks because students do not have to re-discover the "answer" discovered by experts. Based on Experiment 2, students actually did reinvent versions of the true formulas over 80% of the time. But, other work has demonstrated that if students notice a subset of the relevant structure, this is sufficient to help them make sense of the rest, given a formal exposition (Schwartz & Bransford, 1998, Kapur 2008). The contrasting cases are designed to support incremental, even partial, induction rather than one-shot insight. Thus, when used in classrooms, it is important to help students tolerate the short-term ambiguity of not being told the right answer. The effort to find and characterize the structure can improve learning and test performance in the long run.

Towards a Theoretical Account

Cognitive psychology theories provide one way to describe the mechanisms behind the effects of ICC. For example, one might propose that the inventing activities helped students use analogical processes to abstract a ratio schema, and this abstract schema enabled the transfer (e.g., Gick & Holyoak, 1983). An alternative class of theories starts with perception (e.g., Greeno, Smith, & Moore, 1993; Hofstader, 1995). In particular, Gibson's (1979) ecological theories of perception and perceptual learning provided useful guidance for developing the predictions, measures, and contrasting cases. While the ICC tasks engage cognitive mechanisms, and therefore should not be reduced to purely perceptual mechanisms, many of the insights from perceptual learning apply.

With respect to predictions and measures, a perceptual learning account proposes that students need to learn to pick up or notice information in the environment (E. Gibson, 1969). For example, expert radiologists can see diagnostic details in x-rays overlooked by residents (Myles-Worsley, Johnston, & Simmons, 1988), and sommeliers can differentiate wines that simply taste "red" to the uninitiated. If these experts had not learned to perceive the relevant information in the stimulus array, their problem solving, judgment, and abilities to learn from future experiences would be seriously hampered.

We included the redrawing test on the prediction that T&P would not help students notice the ratio structure in the clown worksheet, and therefore, they would have no structures to remember when redrawing the worksheet. In addition, because they had never learned to perceive ratio structures, they would not see them embedded in the transfer tasks, and therefore, they would fail to transfer. Our prediction was based on the assumption that telling-first would lead T&P students to pay attention to what they had just learned (the formulas) rather than searching for the new-to-them information in the worksheet (the physical ratio structure). In this regard, T&P can sometimes exacerbate a more general phenomenon, where prior knowledge filters out perceptual information, in part, because people presume they have seen all that is necessary to complete the task at hand (Nickerson & Adams, 1979; Neisser & Becklen, 1975; Simons & Levin, 1998; Stevens & Hall, 1998).

With respect to the materials, the explicit use of contrasting cases for learning found initial expression in Gibson and Gibson's (1955) demonstration that people learn to pick up information in the environment by progressively noticing structure across systematic variation. According to Gibson (1979), a major task for psychology involves identifying the physical information that makes perception possible, which differs from the cognitive enterprise of identifying the mental structures that enable people to “enrich” or go beyond the available information (also see Gibson & Gibson, 1955). Identifying the information that makes perception possible is highly relevant to the design of guided discovery lessons, where the goal is to have students induce the structure of phenomena. If the cases used for instruction do not include sufficient information, students will not be able to perceive the underlying structure. Perceptual theory provides guidance for deciding what information to include in the examples.

Perception and induction depend on finding structure within variability (e.g., Pass & van Merrienboer, 1994). For the topic of physics learning, the theory of a perceptual gradient helps specify what variability to include in the examples. According to Gibson (1979), gradients make perception possible. A simple example of a gradient is how objects in the distance appear smaller than nearby objects. Paintings capitalize on the receding-into-the-distance gradient so people can see relative distances and object size on flat canvas. The gradient makes it possible to see depth. Painting a single object would not suffice. Creating instructional examples is similar to creating a painting. Examples need to include sufficient information for students to perceive the desired property. If we take the example of density, a single instance does not provide sufficient information for students to perceive density; there needs to be a gradient of densities.

The theory of gradients led to the specific contrasting cases used in this study. For the crowded clowns worksheet (Fig. 1), we created a density gradient across three companies. We chose three companies, instead of two, to specify the linear aspect of the gradient. We also used the same context (clowns within busses), so students could more readily discern the gradient than if we had used three different contexts (e.g., an example of clowns in busses, an example of atoms in a sphere, etc.). For the students in the current studies, it was also necessary to include information specifying a second gradient. Students of this age also needed to notice the proportionality of mass-to-volume for a given density. Therefore, within each company, we created a gradient by varying the number of busses and clowns for its given density.

Our methodology for designing the contrasting cases involved identifying and creating the relevant gradients. It is a method for deciding what similarities and differences to include in the cases. It differs from asking students to explore or freely experiment in the hopes they will generate the relevant gradient information. It also differs from picking instructional examples based on other considerations, such as everyday familiarity.

While the materials were created to include the sufficient and relevant gradients, the inventing directive was included so the students would search for the property we intended them to perceive. The T&P condition demonstrated that the presence of information does not guarantee its uptake. Critically, the inventing task was also designed so students would develop a rule that accounts for the gradients. If we return to the painting example, observers can see depth because artists have included the relevant gradient information. The artists, however, differ from casual observers, because they also have a set of rules that enable them to produce the relevant gradients. They understand – rather than only experience – the relations among objects that produce the perception of depth (e.g., how much to change relative sizes, shapes, and positions to indicate the desired depth). With the inventing task, we simultaneously wanted students to pick up the gradient of density and to develop an understanding of the invariant that produces the gradient (ratios of mass to volume).

It would take more refined research to partial out the relative value of perceptual and cognitive theories for describing the learning in the current studies. For example, it would be useful to determine whether poor recall on the clown redrawing task is due to a failure to perceive the ratios in the first place or due to a cognitive problem such as missing the mental schema of ratio. It would also be an interesting exercise to determine whether cognitive and perceptual theories would make the same prescriptions for creating the cases from which students learn. In the meantime, a major take away from the current research is that it is not sufficient for students to only hear and practice symbolic explanations, lest the structure of the phenomenon gets lost in plain sight.

References

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked example research. Review of Educational Research, 70(2), 181-214.

Barron, B. J., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., Bransford, J. D., & CTGV. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. Journal of the Learning Sciences, 7, 271-312.

Barrows, H. S. & Tamblyn, R. M. (1976) An evaluation of problem-based learning in small groups utilizing a simulated patient. Journal of Medical Education, 51, 52–54.

Bassok, M., & K. J. Holyoak (1989). Interdomain transfer between isomorphic topics in algebra and physics. Journal of Experimental Psychology: Learning, Memory, & Cognition, 15, 153-166.

Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13(4), 640–645.

Bowers, J., Cobb, P., & McClain, K. (1999). The evolution of mathematical practices: A case study. Cognition and Instruction, 17, 25–64.

Bransford, J. D., Franks, J. J., Vye, N. J. & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 470-497). New York, NY: Cambridge University Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), Review of Research in Education , 24, 61-101. Washington DC: American Educational Research Association.

Brown, A. L., & Kane, M. J. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. Cognitive Psychology, 20, 493-523.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. Journal of Experimental Psychology: General, 127(4), 355-376.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. Cognitive Psychology 4, 55-81.

Edelson, D. C, Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. Journal of the Learning Sciences, 8(3/4), 391-450.

Eisner, E.W. (1972). Educating artistic vision. New York: The Macmillan Company.

Fensham, P.J. & Kass, H. (1988). Inconsistent or discrepant events in science instruction. Studies in Science Education, 15, 1-16.

Forbus, K. (1984). Qualitative process theory. Artificial Intelligence, 24 (1), 85-168.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. Journal of Educational Psychology, 95, 393-408.

Gentner, D., & Markman, A. (1997). Structure mapping in analogy and similarity. American Psychologist, 52(1), 45-56.

Gibson, E. J. (1969). Principles of perceptual learning and development. New York: Meredith.

Gibson, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment.

Psychological Review, 62, 32-51.

Gick, M. L., & Holyoak, K. J. (1983). Schema Induction and Analogical Transfer. Cognitive Psychology, 15, 1-38.

Gick, M. L., & Paterson, K. (1992). Do contrasting examples facilitate schema acquisition and analogical transfer? Canadian Journal of Psychology, 46(4), 539-550.

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. Cognitive Psychology, 46, 414-466.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. Journal of the Learning Sciences, 14, 69-110.

Greeno, J.G., Smith, D.R. & Moore, J.L. (1993). Transfer of situated learning. In D. K. Detterman and R. J. Sternberg (Eds.), Transfer on trial: Intelligence, cognition, and instruction (pp. 99-167). Norwood, NJ: Ablex.

Gresalfi, M.S., & Cobb, P. (2006). Cultivating students' discipline-specific dispositions as a critical goal for pedagogy and equity. Pedagogies: An International Journal, 1, 49-58.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. Journal of Educational Psychology, 90, 414-434.

Hiebert, J., & Stigler, J. W. (2004). A world of difference. National Staff Development Council, 25(4), 10-15.

Hiebert, J., Thomas, C.P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. Educational Researcher, 25(4), 12-21.

Hiebert, J. & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. Cognition & Instruction, 14(3), 251-283.

Hofstader, D. H. (1995). Fluid concepts and creative analogies. New York, Basic Books.

Jee, B. D., Gentner, D., Forbus, K., Sageman, B. & Uttal, D. H. (2009). Drawing on experience: Use of sketching to evaluate knowledge of spatial scientific concepts. In N. A. Taatgen & H. van Rijn (Eds.), Proceedings of the 31st Annual Conference of the Cognitive Science Society (pp 2499-2504). Amsterdam, The Netherlands.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (April, 2008). The advantage of abstract examples in learning math. Science, 320, no. 5875, pp. 454-455.

Kapur, M. (2008.) Productive failure. Cognition and Instruction, 26, 379-424.

Kapur, M. (2010). Productive failure in mathematical problem solving. Instructional Science, 38(6), 523-550.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educational Psychologist, 41(2), 75-86.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. Psychological Science, 15, 661-667.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. Child Development, 67, 2797–2822.

Lehrer, R, & Schauble, L. (2004). Modeling natural variation through distribution. American Educational Research Journal, 41(3), 635-679.

Lesh, R. A., & Doerr, H. M. (Eds.) (2003). Beyond constructivism: Models and modeling perspectives on mathematics problem solving, learning, and teaching. Mahwah, NJ: Erlbaum.

Mayer, R. E. (2009). Constructivism as a theory of learning versus constructivism as a prescription for instruction. In S. Tobias & T. Duffy, (Eds.), Constructivist instruction: Success or failure (pp. 184-200). New York: Routledge, Taylor & Francis.

McNeil, N. M., Uttal, D. H., Jarvin, L., & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. Learning and Instruction, 19, 171 – 184.

Michael, A. L., Klee, T., Bransford, J. D., & Warren, S. (1993). The transition from theory to therapy: Test of two instructional methods. Applied Cognitive Psychology, 7, 139- 154.

Myles-Worsley, M., Johnston, W.A., & Simons, M.A. (1988). The influence of expertise on X-ray image processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 553-557.

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. Memory & Cognition, 19, 543–557.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. Cognitive Psychology, 7, 480-494.

Nickerson, R.S., and Adams, J.J. (1979). Long-term memory for a common object. Cognitive Psychology, 11, 287-307.

Pass, F. G., & Van Merrienboer, J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. Journal of Educational Psychology, 86(1), 122-133.

Rittle-Johnson, B, & Star, J. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge: An experimental study on learning to solve equations. Journal of Educational Psychology, 99, 561-574.

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. Journal of Educational Psychology, 101(3), 529-544.

Roll, I. (2009). Structured invention activities to prepare students for future learning: Means, mechanisms, and cognitive processes (Doctoral dissertation). Carnegie Mellon University, Pittsburgh, PA.

Rothkopf, E.Z., & Billington, M. J. (1979). Goal guided learning for text: Inferring a descriptive processing model from inspections times and eye movements. Journal of Educational Psychology, 71, 310-327.

Schmidt, H. G., De Volder, M. L., De Grave, W. S., Moust, J. H. C., & Patel, V. L. (1989). Explanatory models in the processing of science text: The role of prior knowledge activation through small-group discussion. Journal of Educational Psychology, 81(4), 610-619.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. Cognition & Instruction, 16, 475-522.

Schwartz, D. L., Chase, C. C., & Bransford, J. D. (in press). Adaptive transfer: The tension of routines and novelty. Educational Psychologist.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for learning: The hidden efficiency of original student production in statistics instruction. Cognition & Instruction, 22, 129-184.

Schwartz, D. L., Martin, T., & Pfaffman, J. (2005). How mathematics propels the development of physical knowledge. Journal of Cognition and Development, 6, 65-88.

Siegler, R. S. (1981). Developmental sequences within and between concepts. Monographs of the Society for Research in Child Development, 46.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people in a real-world interaction. Psychonomic Bulletin and Review, 5, 644-649.

Stevens, R., & Hall, R. (1998). Disciplined perceptions: Learning to see in technoscience. In M. Lampert and M. L. Blunk (Eds.), Talking mathematics in school: Studies of teaching and learning (pp. 107-148). Cambridge, UK: Cambridge University Press.

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

Sweller, J. (2009). Response to Schwartz, Lindgren, and Lewis. In S. Tobias & T. Duffy, (Eds.), Constructivist instruction: Success or failure (p. 53). New York: Routledge, Taylor & Francis.

Taylor, J. L., Smith, K. M., van Stolk, A. P., & Spiegelman, G. B. (2010). Using invention to change how students tackle problems. CBE – Life Sciences Education, 8, 504-512.

Tobias, S. & Duffy, T. (Eds.) (2009). Constructivist instruction: Success or failure. New York: Routledge, Taylor & Francis.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. Journal of Educational Psychology, 91(2), 334-341.

Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. Cognitive Science, 34, 776-806.

Acknowledgments

This material is based upon work supported by the National Science Foundation under grants SMA-0835854 & DRL-1020362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors thank Kimi Schmidt for help in executing the empirical study; Dylan Arena for his help in the survey of the transfer literature; Melissa Gresalfi for her help in designing and implementing the video protocol; and, Carl Wieman for encouraging us to demonstrate the point.

Footnotes

1. A reasonable question is why the overarching lecture on ratio was given after the T&P students had worked on the clown worksheet and not before. We did not want to use the lecture before students completed phase 1 because the lecture would cloud the strict comparison of T&P instruction versus ICC instruction on the redrawing measure. If the T&P students had received a lecture and then an introductory lesson specific to density before the memory measure, one might argue that a broad orienting lecture is not typical of T&P instruction. Alternatively, one might argue that the T&P students were overloaded by the extra instruction, which might hinder their learning from the subsequent density lesson with the clowns. Therefore, we avoided using the lecture in phase 1. To determine if the broader ratio lecture overcame problems with T&P, we can look at the transfer measures which also test whether students had learned the deep structure of ratio, though less directly than the redrawing measure.

2. An *ISI Web-of-Knowledge* search used “transfer” conjoined with “thinking,” “procedure,” “learning,” or “psychology.” The resulting 350 articles were winnowed to those in English that reported original data on the transfer of math and/or science learning. For each study, two coders separately evaluated whether a solution procedure or concept was provided prior to students’ independent problem-solving. The 3% disagreements were resolved by discussion.

Tables

Table 1. Averages of outcome measures by experiment and treatment, including a further breakout of Experiment 2 based on a median split of achievement using students' prior class performance.

	EXPERIMENT 1				EXPERIMENT 2							
									Broken Out By Prior Achievement			
	ICC		T&P		ICC		T&P		ICC		T&P*	
									High	Low	High	Low
MEASURES	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	M	M	M
Deep Structures	75.5%‡	(35.3)	52.3%	(41.9)	58.6%‡	(40.6)	38.4%	(36.0)	73.1%	46.9%	48.1%	23.2%
Surface Features	27.1%	(18.9)	31.3%	(19.0)	33.1%	(14.8)	28.5%	(15.2)	35.3%	31.2%	29.2%	27.5%
Ratios at Transfer Surface Pressure	35.2%+	(4.6)	13.8%	(3.3)	n/a		n/a		n/a	n/a	n/a	n/a
Ratios at Transfer Spring Constant	53.5%‡	(50.0)	23.0%	(42.4)	50.9%‡	(49.9)	23.3%	(41.8)	64.0%	40.9%	32.9%	8.7%
Word Problems	70.1%	(23.1)	75.2%	(20.3)	64.8%	(19.6)	66.4%	(19.5)	75.0%	57.0%	72.9%	56.5%

+ Comparison of treatment means, $p < .05$. ‡ Comparison of treatment means, $p < .005$. * Two T&P students did not have achievement scores.

Figure Captions

Figure 1. Crowded Clowns Worksheet. Students in the Tell-and-Practice (T&P) condition were told the formula for density and then practiced applying it to determine the crowdedness of the clowns for each company. Students in the Invent-with-Contrasting-Cases (ICC) condition were not told about density, but instead, had to invent their own crowdedness index for each company.

Figure 2. Design of Experiment 1. The Tell-and-Practice (T&P) students received identical practice cases and tests as the Invent-with-Contrasting-Cases (ICC) students. T&P students were told key formulas and concepts, and then practiced on the cases. ICC students had to invent an "index" for each set of cases without being told the appropriate formulas. The images provide snippets of the materials students received. Picture recall, word problem, transfer, and delayed transfer tests served as measures of learning and transfer.

Figure 3. Instructional Pages for Crowded Clown Worksheet. Instructional information was read aloud by students. (a) T&P instructions contained an introduction to the concept with everyday examples, the relevant formula, and a worked example. Students were then told, "On the next page, compute the density for each company that busses clowns." (b) ICC instructions explained the concept of an "index" with everyday examples, provided a narrative for finding a crowdedness index, and introduced rules for the index. Students were then told, "Invent a procedure for computing a crowded clown index for each company."

Figure 4. Transfer Task on Spring Constant. Students were asked to determine the stiffness of the trampoline fabrics, which is the equivalent of finding their spring constants. Though a far semantic transfer from speed and density, it shares the deep structure of ratio.

Figure 5. Samples of Clown Recall. Samples of students' worksheet recreations, partitioned by the inclusion of surface and deep features.

Figure 6. Sample Solutions for the Trampoline Problem. Samples of students' solutions for delayed transfer on the spring constant. Going clockwise from upper left, students noted (a) the total number of people, (b) the total number of rungs, (c) a global impression of stretchiness, or (d) the ratio structure of rungs by people. Only the latter was considered an instance of transfer.

Figure 7. Transfer Performance by Treatment and Number of Cases Included in the Problem. The advantage for ICC over T&P on both 1- and 4-case problem types indicates that students were transferring on the basis of picking up the ratio structure and not a strategy of searching for structure across multiple cases.

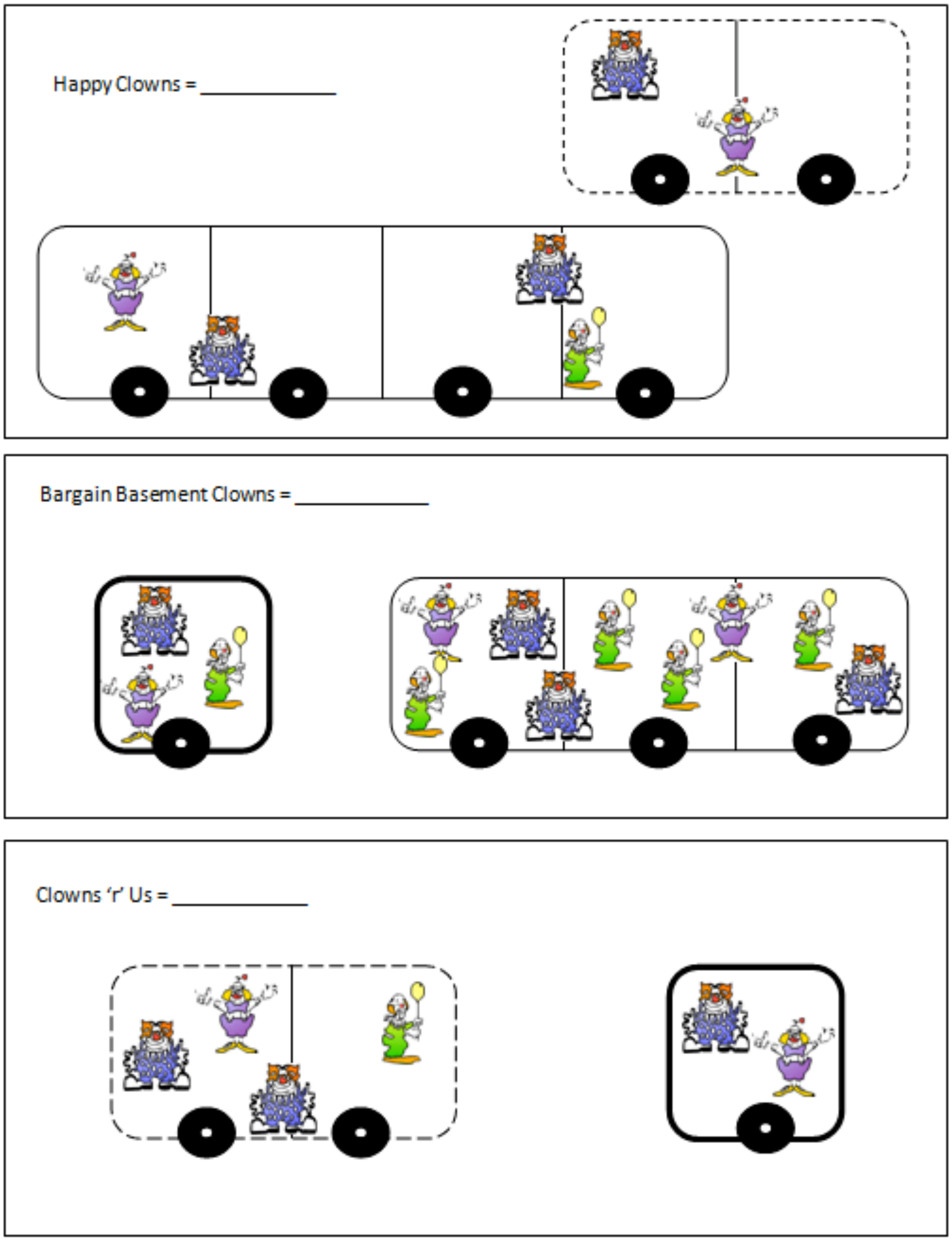


Figure 1. **Crowded Clowns Worksheet.**

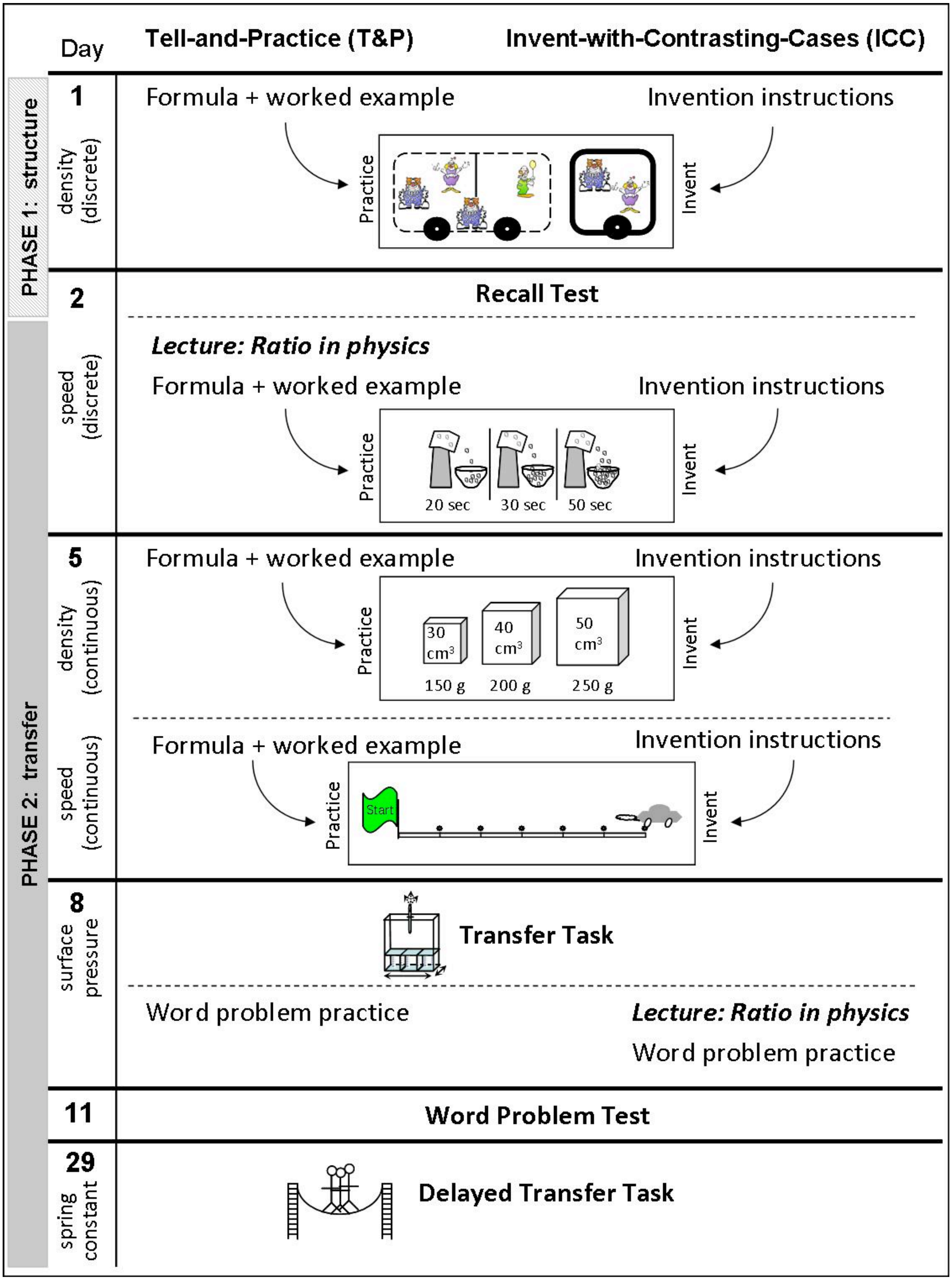


Figure 2. **Design of Experiment 1.**

a.

FINDING DENSITY

Density is how much stuff is packed into a space. Density can be the number of people in a room, the density of feathers in a pillow, and many other things.

Density is very important in chemistry. Density is a property of ***matter***. Gold is denser than carbon, because more matter is packed into each atom of gold compared to each atom of carbon.

When working with density, the trick is to use the simple equation:

$$D = \frac{M}{V} \quad \text{or} \quad Density = \frac{Mass}{Volume}$$

Density is a measure of the mass of a substance per unit of volume.

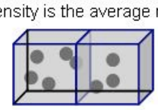
Sometimes we find the ***Mass*** by counting the number of objects.

Volume is the amount of space. Volume is harder to find, because a volume can take many shapes – a sphere, a balloon, a bottle.

To make it easier, we will tell you the volume. We will measure it in cubes.

In the example below, there are two cubes. There are 8 objects spread across the cubes.

Density is the average number of objects per unit of volume.



Density = # objects / volume
= 8 objects / 2 cubes
= 4 objects / cube

On the next page, compute the density for each company that busses clowns.

b.

INVENTING AN INDEX

An **index** is a number that helps people compare things.

Miles per gallon is an index of how well a car uses gas.
Batting average is an index of how well a baseball player hits.
Grades are an index of how well you are doing in school.
Star rating is an index of how efficient an electrical appliance is.

We want you to invent a procedure for computing one kind of index.

THE CROWDED CLOWNS INDEX

- Companies send clowns to parties, circuses, amusement parks, sporting events, and so on.
- To get the clowns to the event, each company packs the clowns into a bus. Some companies make the clowns more crowded than other companies.
- The more crowded the clowns are, the grumpier they will be.
- People who order clowns want to know a company's crowded clown index.
- Invent a procedure for computing a crowded clown index for each company.

RULES FOR THE INDEX

1. The same company always crowds the clowns the same amount, no matter how many clowns get ordered. So a company only gets a single crowded clown index.
2. You have to use the exact same procedure for each company to find its index.
3. A big index value should mean that the clowns are more crowded. A small index number should mean that the clowns are less crowded.

Good luck!

Figure 3. **Instructional Pages Preceding the Crowded Clowns Worksheet.**

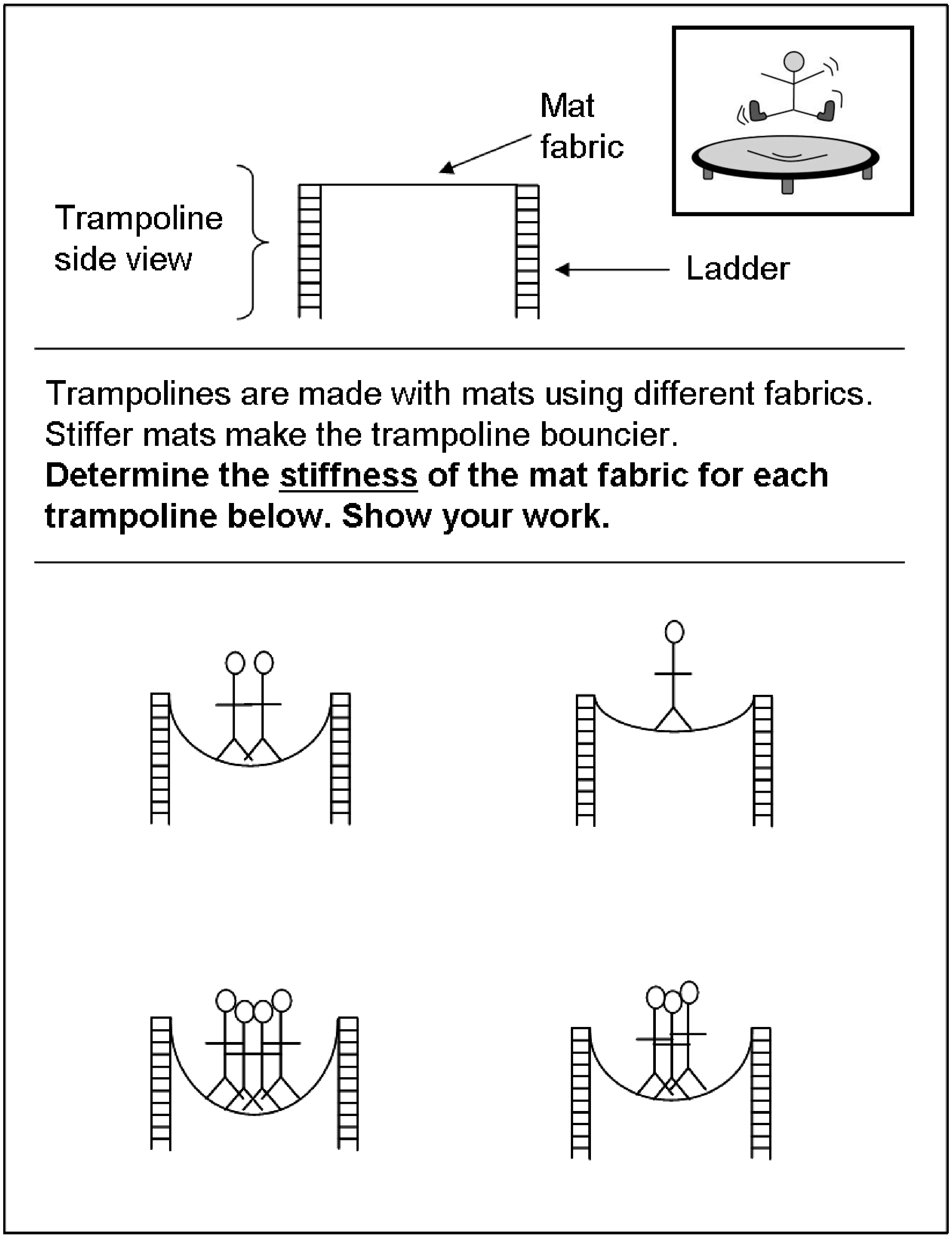


Figure 4. **Transfer Task on Spring Constant.**

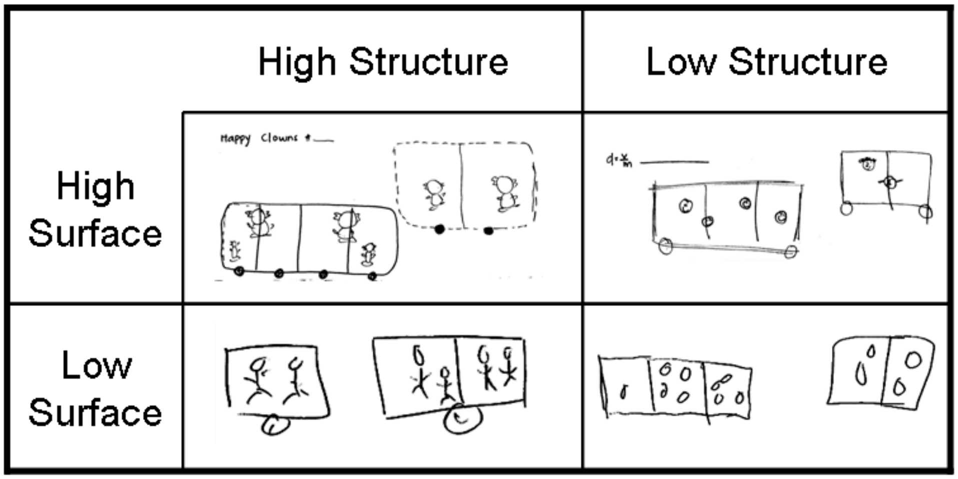


Figure 5. **Samples of Clown Recall**

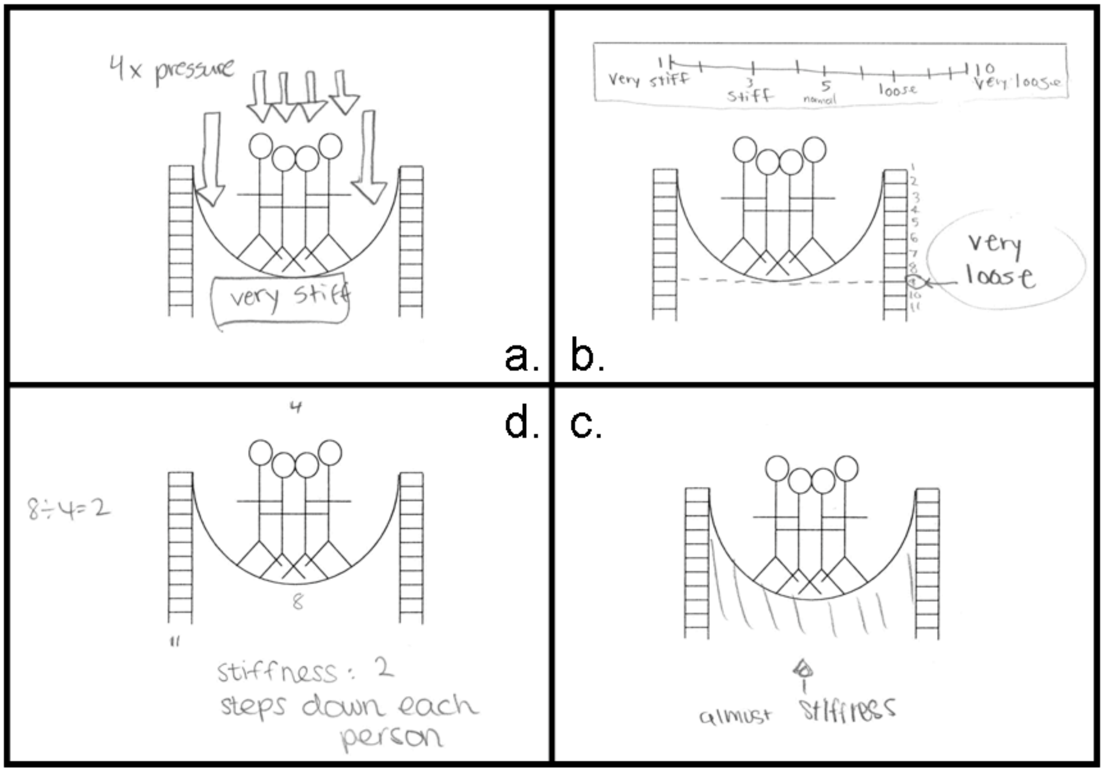


***Figure 6*. Sample Solutions for Trampoline Problem.**

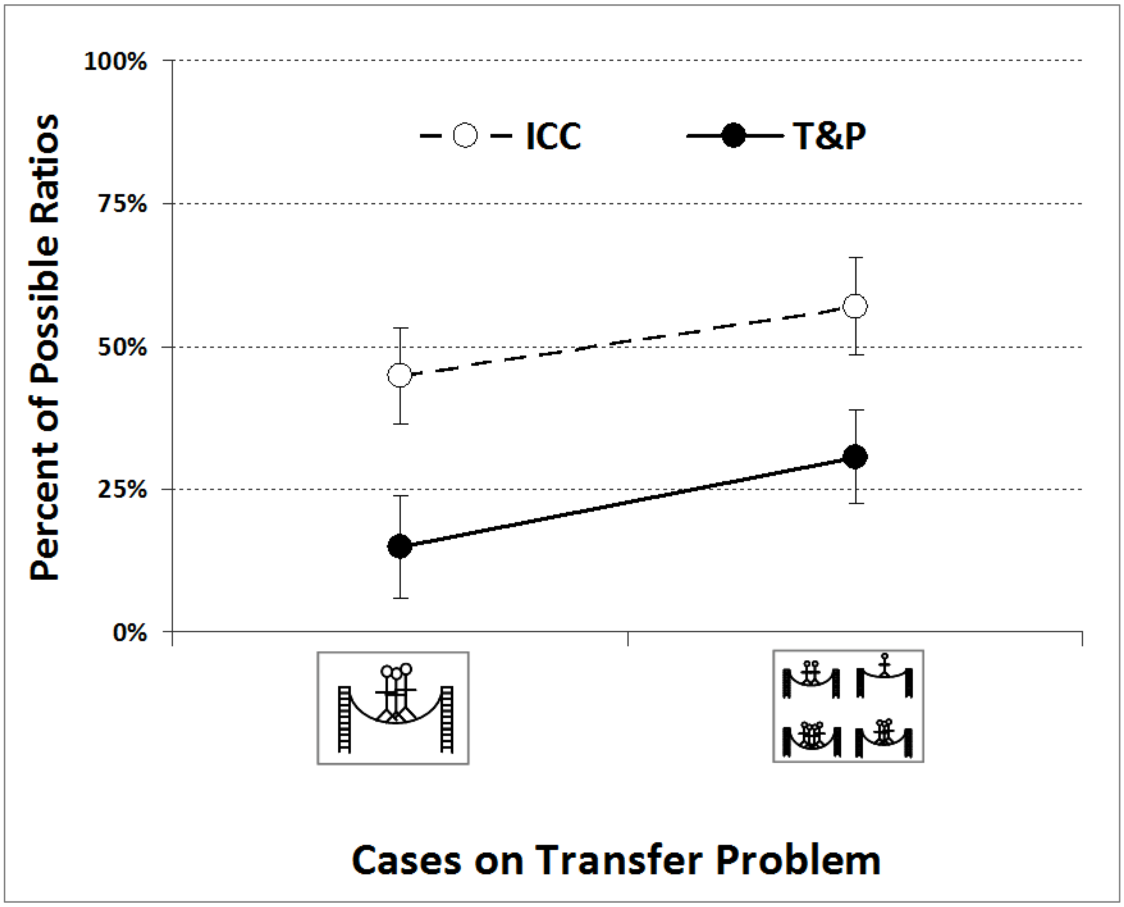


Figure 7. **Transfer Performance by Treatment and Number of Cases Included in the Problem.**