Instrumentation In Learning Research

David A. Sears[*] and Daniel L. Schwartz[†]

[*]*Purdue University, BRNG 5130, 100 N. University Ave., West Lafayette, IN 47907*
[†]*School of Education, Stanford University, 485 Lasuen Mall, Stanford, CA 94305-3096*

Abstract. In physics experiments, a great deal of effort is spent calibrating instruments. These include instruments that precipitate some event, and instruments that measure the effects of those events. Design research in the learning sciences often focuses on precipitating learning events, but it does not pay equal attention to designing effective measures. We present the results of a study that compared two types of instruction on students working alone or in pairs. We show how one measure, common to most studies of learning, failed to detect any effects. Then we show how a second measure, called a Preparation for Future Learning measure, detected important differences. Specifically, pairs working to invent solutions to problems in statistics were more prepared to learn about new, related types of statistics than pairs who were shown how to solve the original problems, as well as individuals who invented or were shown how to solve the original problems.

Keywords: Preparation for Future Learning (PFL), transfer, collaborative learning, Innovation and Efficiency
PACS: 01.40.Ha, 01.40.G-, 01.40.gf, 01.50.Kw

INTRODUCTION

One of the major goals of education is preparation for future learning—the ability to use prior knowledge to learn new skills and understand new concepts. Despite this being a primary goal, the tests we give our students often focus exclusively on retrieving and applying prior knowledge rather than learning. For example, a typical test environment requires students to clear their desks and not talk to their peers (or other sources, like the internet). Bransford and Schwartz [1] called these Sequestered Problem Solving (SPS) environments. They attempt to measure mature knowledge, or the skills that students can apply without help. Having mature knowledge may be essential to developing deeper understanding, but, as Vygotsky [2] was careful to note, it is not the full measure of an individual's understanding. Schwartz and his colleagues developed a different type of assessment to measure this less mature form of knowledge, and they called it Preparation for Future Learning (PFL) assessments [1].

A couple of examples can help illustrate how a PFL assessment works. In an early use of a PFL test, Schwartz and Bransford [3] compared the learning effects of graphing data sets from psychological experiments versus writing summaries of a chapter about the same studies. Given an SPS test in the form of a multiple-choice test, the college students who had done the graphing activity did not look better than the students who had summarized the chapter. For the PFL assessment, students in both conditions then heard a lecture, and afterwards, they had to make predictions about a novel experiment. Students who had completed the graphing activity learned much more from the lecture as indicated by their superior abilities to predict the outcomes of the novel experiment.

In a later study, Schwartz and Martin [4] examined the benefits of asking high-school students to invent their own solutions to statistics problems. The authors did not think the students would invent correct statistics. Instead, the invention activity was designed to prepare them to learn the conventional solutions more effectively. In the invention condition, the students invented mathematical procedures for how to compare scores from different distributions (i.e., normalize data). In the direct instruction condition, students received the same data set, and they were taught and then practiced the graphical equivalent of finding Z-scores (a way to normalize data).

The authors measured student learning by creating a posttest that covered several weeks of instruction. To experimentally evaluate which condition better prepared students for future learning, the authors made two posttest forms. Both forms of posttest included a difficult transfer problem near the end of the test relevant to computing normalized scores. The difference was that the PFL form included an embedded worked example in the middle of the test. The SPS form of the test did not include the worked example. The worked example held the keys to solving the difficult transfer question. The goal of this target transfer question was to see which condition would show a greater benefit from the worked example. The direct instruction led to the same performance on the transfer problem regardless of whether the students received the SPS or PFL version of the test. These students had not been prepared to learn, and they did not learn from the worked example. In contrast, the invention students who received the PFL version of the test did roughly twice as well as the invention students who received the SPS version, and all the students from the direct instruction condition.

Thus, both studies showed that the PFL measure could detect effects of instruction missed by standard SPS assessments. The field of the learning sciences has spent a good deal of time creating instructional treatments that cause learning, but less time developing measures of those treatments. Given the effectiveness of PFL measures, we decided to design a study that put a PFL measure to work to identify which aspects of those treatments cause learning. In particular, we wanted to determine whether collaborative learning, which is common to many reform-minded pedagogies, improves preparation for future learning, and if so, under what conditions.

COLLABORATIVE LEARNING

Collaborative learning has been successful at promoting student achievement, as measured by SPS outcomes, especially when certain scaffolds are in place such as scripts and roles, individual accountability and interdependence, and training [5,6,7,8,9,10]. When collaborative learning is not scaffolded, it tends not to promote educational gains beyond what individuals accomplish studying alone [10]. This suggests at least two possibilities. One is to create collaborative arrangements that do not depend on instructional designs that are difficult for teachers to implement. The second is to determine whether there are specific situations where collaborative benefits are especially effective and merit the effort needed for implementation. We examine these two possibilities.

One way to state our question is to ask which types of tasks naturally benefit from group work, and for what types of outcomes. A broad framework for characterizing different types of tasks comes from recent developments in the field of transfer. Transfer is the generalization of knowledge to new contexts. A longstanding debate in the field has involved whether tasks featuring direct instruction, incremental mastery, and immediate feedback were better for transfer than those that featured discovery, multiple-solutions, and delayed feedback. Schwartz, Bransford, and Sears [11] suggested a resolution to this debate by proposing that rather than seeing these tasks as being on opposite ends of a single dimension, they could be seen as on two complementary dimensions. They called these dimensions Efficiency and Innovation. Efficiency tasks often involve direct instruction followed by repeated practice. The goal is speed and accuracy, or routine expertise [12]. Examples include times-tables recitations and end-of-chapter review problems. Innovation activities involve opportunities for students to explore, invent, or discover solutions to problems rather than being told what to do from the outset. The outcomes of the innovation dimension of instruction have not been as clearly defined as the outcomes of instruction that emphasizes efficiency.

Schwartz et al. [11] hypothesized that one benefit of tasks aligned to the Innovation dimension is that they would prepare students for future learning. For example, by allowing students to attempt to find a solution before receiving the expert approach, they will be more likely to connect their prior knowledge to the new material and to notice key features that characterize the problem. When they receive the expert solution, they should be more prepared to appreciate how it works and how they can apply it flexibly in the future. These authors proposed a need to balance innovation and efficiency experiences in instruction, because each experience when appropriately combined contributes to a trajectory of adaptive expertise [12].

In the current study, we compared Efficiency only instruction with Innovation plus Efficiency instruction. We assessed the outcomes using both SPS measures of efficiency and PFL measures of students' abilities to learn given new resources during the test.

Of particular interest was whether initially working alone or collaboratively would affect performance on the SPS and PFL measures. Under what conditions are two heads better than one, and for what outcomes? For Efficiency tasks, one can imagine that knowing the procedure to use would allow groups to monitor and correct mistakes. If so, we should expect collaborators on efficiency tasks to do better on SPS assessments than those working alone or those collaborating on innovation tasks. Alternatively, one might imagine that innovation tasks would permit students to gain access to the benefits of different points of view so they can cull out the deep structure common to both their perspectives [13]. If true, then collaborating over innovation tasks should better prepare students to learn in the future, because they would develop a more structured understanding of the domain.

In the experiment, the Innovation and Efficiency framework was tested for its ability to characterize naturally productive tasks for collaborative learning. University students, mostly undergraduates, worked alone or in same-gender pairs to learn about the chi-square formula. On the posttest, all participants worked alone and answered three types of questions: 1) SPS calculations of the chi-square, 2) SPS comprehension measures that tapped the depth of understanding about how the formula worked, and 3) PFL measures of students' abilities to transfer their lessons to learn a related, but complicated statistical concept (i.e., computing inter-rater reliability).

METHODS

Participants: Seventy-six university students were randomly assigned to one of two conditions: Innovation or Efficiency. Participants either worked in same-sex pairs (36) or alone (40). Forty-eight women (24 in dyads) and 28 men (12 in dyads) with little or no background in statistics participated.

Materials—A nine-page learning packet about the chi-square formula and a seven-item posttest comprised the materials for this experiment. The learning packet consisted of three units about different aspects of the chi-square formula. Each unit contained three pages: Lesson, Problems, and Final Practice Example with solution provided. (The order of these pages was shuffled to implement the Efficiency v. Innovation plus Efficiency contrast, as described below.)

Time taken for the learning-packet and the posttest was recorded. The seven-item posttest included two problems requiring chi-square calculations, three comprehension questions about where and how the formula worked, and a difficult PFL assessment with two parts related to the statistics topic of inter-rater reliability. For the PFL assessment, a resource question introduced the topic of inter-rater reliability, and the target question built upon those principles and required participants to transform their chi-square formula into a measure analogous to Cohen's Kappa. This is a very far transfer task, because students had not been taught Cohen's Kappa during the initial instruction. The goal was to see which version of the task, Innovation or Efficiency, would better prepare students for learning how to solve the PFL target problem.

Procedures—The key difference between conditions was that participants in the Innovation condition had to invent solutions to the Problems *before* seeing the Lesson while those in Efficiency saw the Lesson and then applied what they learned to the Problems. Table 1 summarizes the procedures.

TABLE 1. Procedures

Step	Context	Innovation	Efficiency	Time
1	Alone / Dyads	9-page Learning Packet on the Chi-Square Formula		35 to 65min
		1) Problems (INVENT a formula) 2) Lesson (instruction) 3) Final Example (reinforce)	1) Lesson (instruction) 2) Problems (APPLY the formula) 3) Final Example (reinforce)	
		Short Break		5min
2	Alone	Posttest (7 problems: 2 calculations, 3 comprehension, PFL—resource & target)		25min

RESULTS

The posttest was consistent internally (alpha = .81) and across different raters (all Cohen's Kappas > .81). Time-on-task showed no significant differences between conditions as shown in Figure 1.

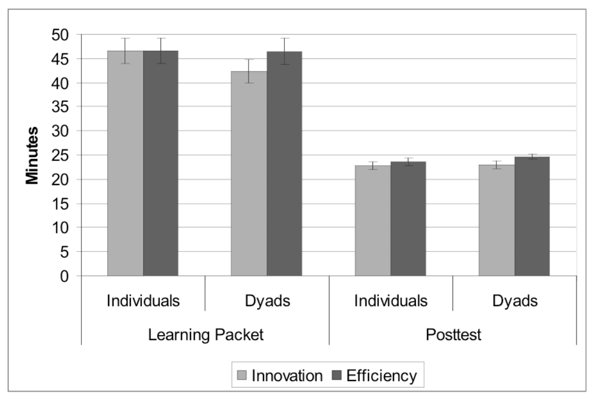

FIGURE 1. Time-On-Task. (SEM Bars in All Graphs).

As shown in Figure 2, studying in groups showed no benefits on the SPS calculation and comprehension measures for either condition. Thus, working in dyads did not improve efficiency outcomes relative to working alone, and working on innovation tasks (with subsequent efficient instruction) did not harm either individuals or dyads.[1] Thus innovation-based instruction can lead to good SPS outcomes.

---

[1] Dyads scored lower on the comprehension questions than individuals. This was unexpected and likely due to them having to share the Lessons page rather than each having a copy. Post-hoc analyses indicated that dyads, particularly those in the Innovation condition, spent less time on the Lessons pages than individuals (Means: 9.2, 6.3, 9.9, and 9.2 minutes for Innovation individuals and dyads, and Efficiency individuals and dyads, respectively; SEMs: ± 0.7).

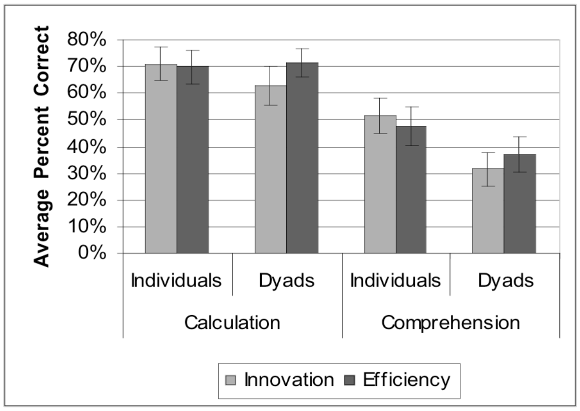

FIGURE 2. Innovation Outperformed Efficiency on the Transfer Measures.

Participants in the Innovation condition surpassed their Efficiency counterparts on the PFL measure. On the PFL target question, Innovation dyads exceeded all others (see Figure 3). Thus, they showed the greatest preparation for future learning, and we can tentatively conclude that innovation tasks are naturally suited to the benefits of working in groups. For SPS outcomes, it does not matter much how students are taught if it is a well-designed curriculum. However, for PFL outcomes, innovation activities using dyads provide a special benefit.

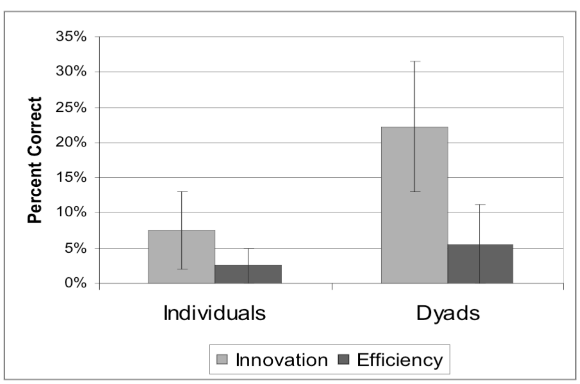

FIGURE 3. Performance on the PFL Target Question.

CONCLUSION

By using PFL assessments, we found that opportunities to invent solutions to problems followed by direct instruction on the canonical solution can facilitate transfer and collaborative learning. In so doing, we found that the Innovation and Efficiency framework provides a means of characterizing naturally productive tasks for group learning. If only traditional measures of performance were used, we would not have seen any differences between Innovation and Efficiency individuals or dyads. This would have led to the erroneous conclusion that the novel Innovation method of instruction produced no benefits compared to the more traditional Efficiency instruction, and that working in groups has no natural benefits relative to working alone.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. SLC-0354453. 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.

REFERENCES

1. J. D. Bransford, and D. L. Schwartz, "Rethinking Transfer: A Simple Proposal with Multiple Implications," in *Review of Research in Education*, edited by A. Iran-Nejad and P. D. Pearson, Washington, D.C.: American Educational Research Association, 1999, pp. 61-100.
2. L. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes,* edited by M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, Cambridge, MA: Harvard University Press, 1978/1932.
3. D. L. Schwartz and J. D. Bransford, *Cognition and Instruction*, **16**, 475-522, (1998).
4. D. Schwartz and T. Martin, *Cognition & Instruction,* **22**, 129-184, (2004).
5. E. Coleman, *The Journal of the Learning Sciences*, **7**, 387-427, (1998).
6. M. R. Gillies and A. F. Ashman, *Journal of Educational Psychology*, **90**, 746-757, (1998).
7. D. Johnson and R. Johnson, *Theory into Practice*, **38**, 67-73, (1999).
8. A. King, "Discourse Patterns for Mediating Peer Learning," in *Cognitive Perspectives on Peer Learning,* edited by A. O'Donnell and A. King, Mahwah, NJ: Lawrence Erlbaum Associates, Inc, 1999, pp. 87-115.
9. A. O'Donnell, "Structuring Dyadic Interaction Through Scripted Cooperation," in *Cognitive Perspectives on Peer Learning,* edited by A. O'Donnell and A. King, Mahwah, NJ: Lawrence Erlbaum Associates, Inc, 1999, pp. 179-196.
10. R. Slavin, *Contemporary Educational Psychology,* **21**, 43-69, (1996).
11. D. Schwartz, J. Bransford and D. Sears, "Efficiency and Innovation in Transfer," in *Transfer of Learning from a Modern Multidisciplinary Perspective*, edited by J. Mestre, CT: Information Age Publishing, 2005, pp. 1-51.
12. G. Hatano, and K. Inagaki, "Two Courses of Expertise," in *Child Development and Education in Japan*, edited by H. Stevenson, H. Azuma, and K. Hakuta, New York, NY: Freeman, 1986, pp. 262-272.
13. M. L. Gick and K. J. Holyoak, *Cognitive Psychology,* **15**, 1-38, (1983).