SREE

2040 Sheridan Rd
Evanston, IL 60208

Society for Research on Educational Effectiveness
Advancing Education Research

www.sree.org
847.467.4001

2010 SREE Conference Abstract Template

Thank you for your interest in the Society for Research on Educational Effectiveness 2010 Annual Conference. Conference abstracts must be submitted using this template document. The template is based on the recommendations offered by Mosteller, Nave, and Miech (2004, p. 33)[*] for structured abstracts. Abstracts should follow APA Style, as specified in the Sixth Edition of the Publication Manual of the American Psychological Association.

Abstract Preparation and Submission Procedures

Save this document to your computer. Fill in each of the sections below with the relevant information about your proposed poster, paper, or symposium. Make sure to save the document again when completed. When ready, submit your abstract at http://www.sree.org/conferences/2010/submissions/

The template consists of the following sections: title page, abstract body, and appendices (references and tables and figures). Figures and tables included as part of submission should be referred to parenthetically—"(please insert figure 1 here)." The body section of your abstract should be no longer than 5 pages (single spaced, using the Times New Roman 12-point font that has been set for this document). The title page and appendices do not count toward this 5-page limit.

Insert references in appendix A of this document. Insert tables and graphics in appendix B. Do not insert them into the body of the abstract.

For questions, or for help with abstract preparation or submission,
contact us at inquiries@sree.org, or 847-467-4001

[*] Mosteller, F., Nave, B., & Miech, E. (2004). Why we need a structured abstract in education research. *Educational Researcher, 33(1),* 29–34.

Abstract Title Page

Title: Experimental Comparison of Inquiry and Direct Instruction in Science

Author(s): Dr. William Cobern, Department of Biological Sciences & (Director) The Mallinson Institute for Science Education, Western Michigan University

Dr. David Schuster, Department of Physics & The Mallinson Institute for Science Education, Western Michigan University

Betty Adams, Department of Physics & The Mallinson Institute for Science Education, Western Michigan University

Abstract Body.

Background/context:

For many decades there have been both educational and political policy debates over the merits of “inquiry-based” and “direct” approaches to teaching science, with strong opinions on both sides. In broad practice the pendulum has been mostly on the direct side, but in recent years, with the formulation of national and state science education standards, inquiry has become the *sine qua non* for science instruction (1, 2, 3). By inquiry-based teaching of science we mean instruction reflecting the investigative approach, empirical techniques, and reliance on evidence that scientists use in discovering and constructing new knowledge. A related idea is “Scientific Teaching” described by Handelsman et.al. in *Science* (4), an aspect of which is that “the teaching of science should be faithful to the true nature of science by capturing the process of discovery in the classroom.” Many educators feel that inquiry teaching rather than direct is most in keeping with the widely accepted constructivist theory of how people learn; but note that constructivism is a theory of learning, not of instruction. Students must construct their own understanding regardless of the resources, whether laboratory activities, lectures, discussion, or text.

Inquiry teaching of science can be approached in many ways, as noted by Alberts (1). He lists several goals of science education: to know, use, and interpret scientific explanations of the natural world; to generate and evaluate scientific evidence and explanations; to understand the nature and development of scientific knowledge; and to participate productively in scientific practices and discourse (6). Brady (7) notes that most science educators claim that we “know how to teach science,” meaning of course to teach science by inquiry, and that the only question remaining appears to be “when will we do it?” Handelsman and colleagues (4) state that “to the extent that the authors include a pedagogy of inquiry there is a clear claim here that such a pedagogy has been soundly tested and found effective.” However, proponents of both inquiry-based and direct instruction are convinced of their respective positions regarding pedagogical efficacy. But are these claims justified by evidence?

A recent meta-analysis conducted by the Educational Development Center (EDC) of research studies into inquiry instruction does not find them to yield sufficiently unconfounded inferences to adequately address the central issue (8). Indeed, the EDC reports that research rigor had in fact declined from 1984 to 2002. Various confounding threats to the validity of inquiry research since the 1960s include:

- Comparative controlled studies, pitting inquiry against worthy alternative instruction, are few. Controls are often absent or represented by poor or nebulous “traditional” teaching.
- Few studies use randomized assignment of subjects to treatment groups or quasi-experimental efforts to control for differences between subject groups.
- Evaluations not independent of the developers and researchers can be problematic
- Replication may not be feasible due to insufficiently detailed specification of treatments.
- Rarely is implementation compared to the intended instruction, so that tacit assumptions of fidelity may be unwarranted.

We seek to avoid each of these threats to validity, considering first the issue of a significant “alternative” to inquiry, a model designed to qualify as high quality direct instruction. Science education specialists invariably talk about moving teachers away from direct instruction toward inquiry, a understandable commitment for those who love both science and the teaching of science. A less commendable argument for inquiry, however, is to contrast it with straw man caricatures, such as direct instruction cast as exposition, memorization, and cookbook laboratory work. David Ausubel addressed such misrepresentation more than 40 years ago (9). He argued that the true issue was rote learning vis-à-vis meaningful learning, and that neither inquiry instruction nor direct instruction *automatically* lead to meaningful learning.

Since considerable educational and political support exists for various forms of direct instruction, countering the emphasis on inquiry in the standards, the question remains: is some expert form of direct instruction more effective than inquiry instruction? David Klahr (10) thinks so, claiming his research findings "challenge predictions derived from the presumed superiority of discovery approaches for deeper, longer lasting, and 'more authentic' understanding of scientific reasoning processes…" Such views are not isolated; in 2006 Kirschner, Sweller and Clark (11) published a paper entitled "Why minimal guidance during instruction does not work," and Sweller (12) critiqued constructivism in 2009. However, one may note that Klahr's research involves open "discovery," the most unstructured form of inquiry, rather than the guided approach advocated by the NRC and AAAS, and his "direct" mode arguably involves aspects of guided inquiry. Moreover, the study focuses on science process skills rather than speaking to the core question regarding inquiry instruction for concept development. Nevertheless, Klahr and Sweller draw attention to the precarious evidentiary support for inquiry instruction. Convincing evidence is similarly lacking for the historical status quo of direct instruction. Hence, inquiry instruction vis-à-vis direct instruction has not been subjected to experimental controlled studies comparable to that of Klahr or the work referred to by Sweller.

Purpose / objective / research question / focus of study:

It is evident that 'experientially-based' instruction and 'active student engagement' are advantageous for effective science learning. However, "hands-on" and "minds-on" (5) aspects can occur in both inquiry and direct science instruction, and convincing comparative evidence for the superiority of either mode remains rare. Thus, the pertinent question we seek to address is not whether active, experiential learning of science is more effective than passive, non-experiential learning. Our research question is whether an inquiry approach or a direct approach to experientially-based instruction is more effective for science concept development, when both approaches are expertly designed and well executed. This research undertook a controlled experimental study comparing the efficacy of carefully designed inquiry instruction and carefully designed direct instruction in realistic science classroom situations at the middle school grades.

Setting:

Research trials were held in science classrooms on the Western Michigan University campus in Kalamazoo, Michigan. Identical equipment and supplies, specifically suited to the chosen instructional topics, were on hand for each of five classrooms.

Population / Participants / Subjects:

The subjects were 180 incoming 8th grade students from several Midwest school districts, urban, suburban and rural. Districts sent out advance program announcements to parents, hence participation was a family decision. Our five instructors were veteran middle school science teachers. Their judgment was that the students were not noticeably different overall from those in their regular school courses. A voluntary summer program, however, has limitations. There are no grade incentives, and to include homework or reading assignments would be unrealistic. Learning is thus essentially dependent upon in-class student engagement.

Intervention / Program / Practice:
Research trials were held for two weeks of June in both 2007 and 2008. Two science unit topics with substantial conceptual demand were chosen for instruction:

It's Dynamic! **The concepts of force, motion and mass, and their interrelationship in Newton's first and second laws of motion.**

It's Illuminating! **Basic science (light energy depends on angle, distance, time), and application to temperature variation on Earth due to location (latitude) and time of year (seasons)**

Four days each week, five classes met in the mornings to cover one lesson from each of these two science units, with an intermission snack break. Students were embedded in classes, and each class had a dedicated teacher embedded in a particular mode of instruction. Three teachers taught by "inquiry" mode, two by "direct" mode (teachers will switch roles for the subsequent two years of trials, so as not to cross-contaminate modes within teachers). Identical pre- and post-tests were administered on each unit topic.

The inquiry and direct versions of each science unit ("Dynamics" and "Light") were designed in parallel, to ensure equivalence in science content and approximate teaching time, while differing in the sequencing and epistemological bases of experiential activities. The essential aspect distinguishing inquiry and direct modes was boiled down to an "active agent" of "how students come to the concept." In one mode, students develop the ideas and principles from their own exploration, in the other they are told before they confirm. Common lesson ingredients reflected generic features of good instruction, including interactivity with people and materials.

Research Design:
For our controlled experimental study, we created a special summer program to enable random assignment of students to treatment and control groups, often difficult in a regular school setting. This also facilitated control of time-on-task. To minimize threats to validity, we built four features into our research: Specificity, Fidelity, Objectivity, and Transparency. We *specify* explicit *models* for our two instructional modes, with a Guided Inquiry instruction model based on the Karplus Learning Cycle (13), Exploration, Concept Formation, and Application. Students are guided toward "inventing" relevant scientific concepts and "discovering" the relationships and laws, guided by the instructor. In contrast, Presentation, Explanation, Confirmation, and Application, are the phases of our direct instruction model, with the teacher presenting and explaining concepts, relationships and laws directly to the students, as finished products to be learned and understood. We call our direct model "direct active" since it includes hands-on practical work, though of a confirmatory nature and with prescribed steps.

Fidelity to mode and to curriculum are important features of the research. Teachers practiced in advance, and were evaluated for fidelity during teaching in three ways. First, independent specialist observers from the Science and Mathematics Program Improvement (SAMPI) group (14), initially "blind" to teacher mode, visited two lessons per unit for each teacher, two observers seeing each teacher and scoring them on fidelity to method. They successfully identified modes for each teacher, and were subsequently able to score teachers specifically on fidelity to the intended lessons. The second fidelity check was that teachers kept journal notes on each day's teaching, and third, lessons were videotaped for review.

Our research design had several areas of *blindness* to enhance objectivity and minimize bias; teachers were blind to assessments, SAMPI was initially blind to teacher mode, and SAMPI coded and analyzed the data without knowledge of group. In the interest of *transparency*, we have placed all critical study information on our project web site (15), including complete unit descriptions, learning objectives, student materials, teacher guides, and assessments.

Data Collection and Analysis:
The assessment instruments are sets of 24 conceptual multiple-choice questions, each with four choice options. The questions embody our criterion for conceptual understanding and the ability to *apply* it (16), rather than recall of factual knowledge. Pre- and post-tests were administered and recorded by independent project evaluators.

Student performance data was analyzed to determine gains (*t*-test and ANOVA, via SPSS), and to compare across various factors, including instructional mode, science topic, teacher, and program year. On the way toward answering our primary research question, we also gleaned information about our attempts at randomization, and teacher consistency over the two trials.

Findings / Results:
Our teacher fidelity-to-mode median rating of 86% is arguably adequate for our research purposes while remaining realistic with respect to inevitable variation in actual science classrooms. Fidelity scores were somewhat higher for direct instruction than inquiry, which is not unexpected since direct is easier to 'execute' than guided inquiry. Student scores on the pre-tests indicated that randomization of students across classrooms was effective, i.e. any variation in pre-scores between classes was consistent with that expected by chance for class sizes of around 20 students. Differences in results between years were not statistically significant, thus could be viewed as replication data with different students, and aggregated to study other factors.

Average pre-test scores on the multiple choice assessment instruments were around 50%, with standard deviations around 20%. In every sub-category we analyzed, there were statistically significant though modest gains. Standard deviations on post-tests were similar to those on the pre-tests. Figure 1 shows representative pre- and post-test score data, for the case of the Light unit in Inquiry mode. **(please insert figure 1 here).**

Performance gains between pre- and post-tests were expressed as both raw and normalized gain scores, the latter being the ratio of gain to maximum possible gain given the pre-score. To avoid distortions that can occur when a gain is negative, we used the concept of normalized *change* in calculations (17). Normalized change is the gain or loss over the maximum possible gain or loss respectively, expressed as a percentage.

Mean normalized gains were of the order of 20% for the Dynamics unit and 30% for the Light unit, for both treatments, comparable to results on the FCI (Force Concept Inventory) of physics education fame (18). The effect sizes (Cohen's d) for raw gain were .69 for the Light unit and .54 for the Dynamics unit. Effect sizes for normalized gain were 0.99 for Dynamics and 1.4 for Light. Raw gains showed consistent (negative) correlation with pretest scores, but normalized gains did not, indicating that normalization was working in this regard.

Comparative results for the Light/Climate and Dynamics units are displayed as bar charts in Figures 2 and 3, with standard error bars overlaid, and accompanied by tabulated values for scores, gains, and standard deviations. The two charts display results by topic, teacher, and instructional mode, data aggregated for both years. **(please insert figure 2 and figure 3 here)**

One might possibly expect a *correlation* between student gains in two topics taught by the same teacher. We find that raw percent gain does not correlate between the Light and Dynamics unit (Pearson's, $r=.077$, $p=.308$), but normalized gain/change shows a statistically significant correlation (Pearson's, $r=.238$, $p=.001$). No statistically significant differences in normalized gain were found between different teachers within the same instructional mode. Observationally, however, "natural variations" in personal teaching style and practice were clearly evident.

Regarding the central research question, comparing inquiry and direct instruction, the center portions of the bar charts in Figures 2 and 3 represent these results. For the Light unit, over two trials, the difference between direct and inquiry groups on normalized gain/change was not statistically significant ($t(178)=.755$, $p=.451$) (mean diff. 3.8, std. error diff. 5.1, effect size Cohen's d=.12). Raw gain had a small but statistically significant difference between one direct teacher (Ann) and one Inquiry teacher (Tom) ($t(73)=2.132$, $p=.036$); upon normalization the difference was not statistically significant ($t(73)=1.857$, $p=.067$). Similarly, for the Dynamics unit, differences between direct and inquiry groups on normalized gain/change were not statistically significant. ($t(178)=.717$, $p=.474$) (mean diff. 3.1, std. error diff. 4.4, effect size Cohen's d=.11). Standard deviations reflected the wide range in scores in a voluntary program.

Conclusions:
Results indicate that inquiry and direct methods led to comparable science conceptual understanding in roughly equal instructional times. Gain differences between instructional modes were not statistically significant within the observed natural variation of students and teachers. Given such score spreads, both gains and gain differences would need to be larger than those observed in order to show statistical or practical classroom significance. A single larger-scale study would of course provide larger N size, but at the cost of precision, since it becomes more difficult to prepare, control and monitor instructional and classroom situations, thus increasing variation. Following Cronbach (19) a number of separate local studies in various environments would be more informative, to see whether and how the findings generalize to other situations and to refine and study the effect of various parameters.

Mastery of science content in the alternative modes was our central research question, but the inquiry-versus-direct debate is not just about content: it is also about the nature of science and about efficiency. Most science educators feel that Inquiry Instruction, by its very nature, provides crucial added value, in having students 'do' science for themselves. For Direct Instruction, given our finding that it does not lead to a better grasp of the basics, it is not as clear what other grounds there might be on which to argue superiority. It may be easier or less time consuming for the teacher, or less demanding for weaker students, at least initially. However, direct instruction risks sending the message that science is simply a body of knowledge to be learned.

Given the composite nature of all good lessons, and the realities of implementation in classrooms, we see that common claims for superior concept acquisition by either direct or inquiry instruction may be viewed as overstated. Our study shows that good direct and inquiry instruction led to similar understanding of science concepts and principles in comparable times. It may well be that under more tightly controlled and rehearsed conditions one could better distinguish the performance effects of mode of instruction, which would be of significant theoretical interest; but this study gives a practical indication of what is likely to happen in the field. Thus, the promotion of one mode of instruction over the other, where both are based on sound models of expert instruction, should not be based on content acquisition *alone*.

Inquiry-based instruction clearly offers significant potential advantages for science education, by modeling scientific inquiry during concept learning; these concomitant benefits would need to be studied in research designed for that purpose. However, as far as science concept understanding is concerned, our conclusion is that expertly designed instructional units, sound active-engagement lessons, and good teaching are as important as whether a lesson is cast as inquiry or direct (20).

Appendices

Appendix A. References

1. National Research Council. (2000). *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning.* Washington, D.C.: National Academies Press.
2. American Association for the Advancement of Science. (1990). *Science for all Americans: Project 2061.* New York: Oxford University Press.
3. Alberts, B. (2008). Considering Science Education. *Science,* 319, 1589.
4. Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S. M., Wood, W. B. (2004). Scientific Teaching. *Science.* 304, 521-522.
5. National Research Council. (2000). *How People Learn.* Washington, D.C.: National Academies Press.
6. Alberts, B. (2009). Redefining Science Education. *Science,* 323, 437.
7. Brady, T. E. (2008). *PDK*, 89(8), 605-607.
8. Education Development Center. *Inquiry Synthesis Project: Executive Summary*. [Web Page]. URL: http://cse.edc.org/products/inquirysynth/ [2009, July 12].
9. Ausubel, D. P. (1961). In Defense of Verbal Learning. *Educational Theory.* 11(1), 15-25.
10. Klahr, D. *Paths of Learning and Their Consequences: Discovery Learning versus Direct Instruction in Elementary School Science Teaching* [Web Page]. URL: http://www.lrdc.pitt.edu/supergroup/ [2009, April 10].
11. Kirschner, P. A., Sweller, J., Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work. *Educational Psychologist.* 41(1), 75-86.
12. Sweller, J. (2009). In *Constructivist Instruction: Success or Failure?* S. Tobias, T. M. Duffy, Eds. New York: Routledge. pp. 127-143.
13. Lawson, A. E., Abraham, M. R., Renner, J. W. (1989). *A Theory of Instruction: Using the Learning Cycle to Teach Science Concepts and Thinking Skills.* NARST, Cincinnati, OH.
14. SAMPI [Web Page]. URL http://www.wmich.edu/sampi/ [2009, June 4]
15. Project materials are available at: www.wmich.edu/way2go/.
16. Anderson, L. W., Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* New York: Longman.
17. Marx, J., Cummings, K. (2007). Normalized Change. *American Journal of Physics*, 75(1), 87-91.
18. Hake, R. (1998). Interactive-Engagement Versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses. *American Journal of Physics*, 66, 64-74.

19. Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116-127.

20. Funded by the National Science Foundation's Interagency Education Research Initiative (IERI/NSF 04-553) Award #0437655. Any opinions, findings, conclusions or recommendations in this paper are those of the authors and do not necessarily reflect the views of the NSF.

Appendix B. Tables and Figures

Fig. 1. Prescore/Postscore distribution

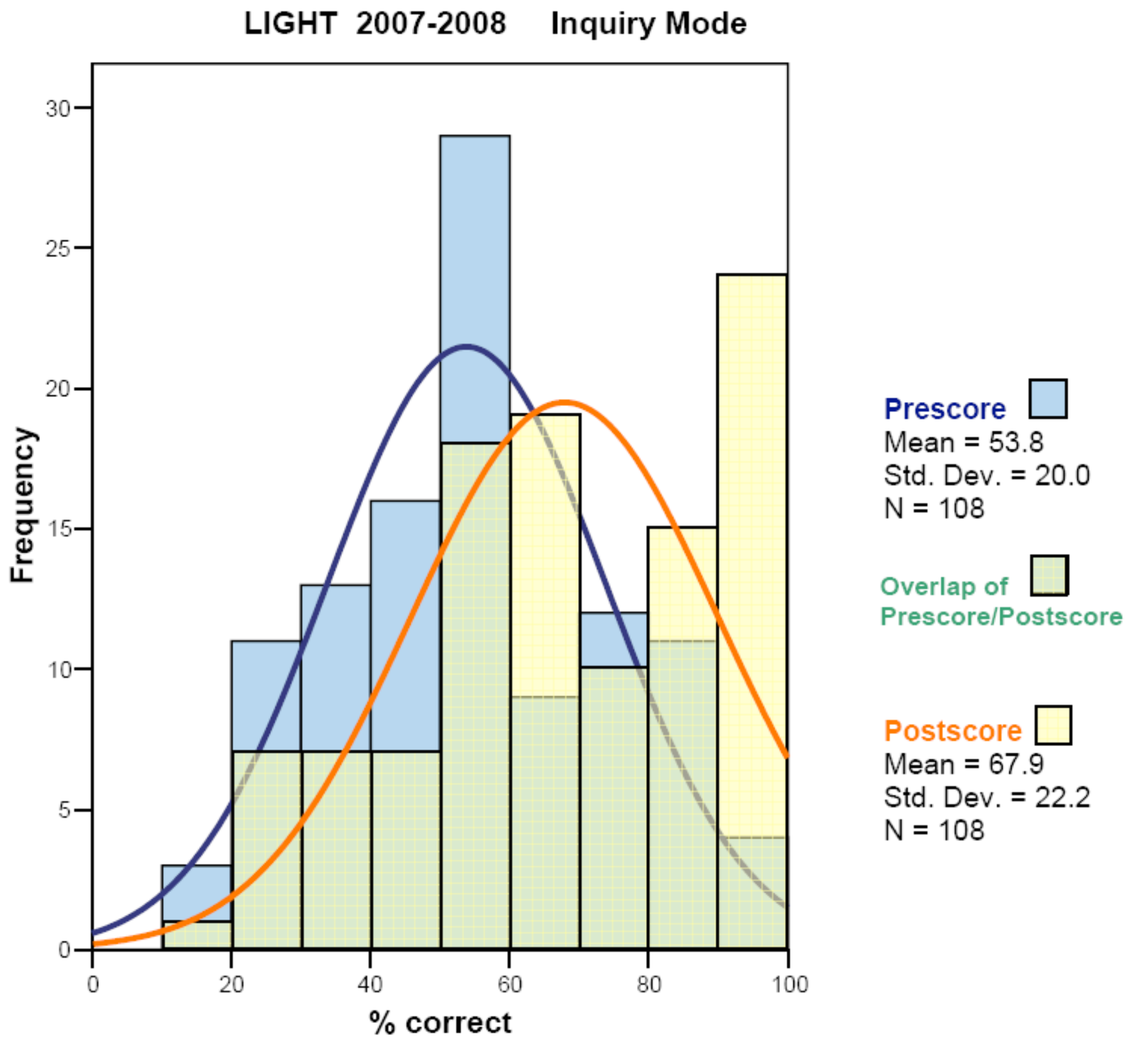
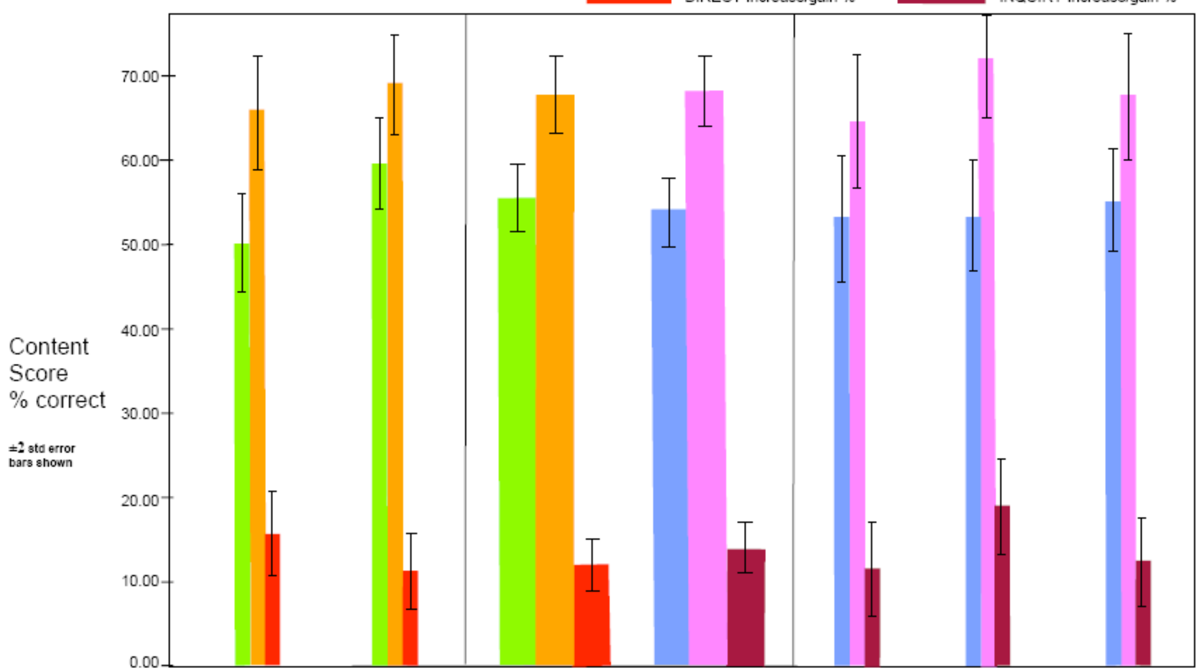
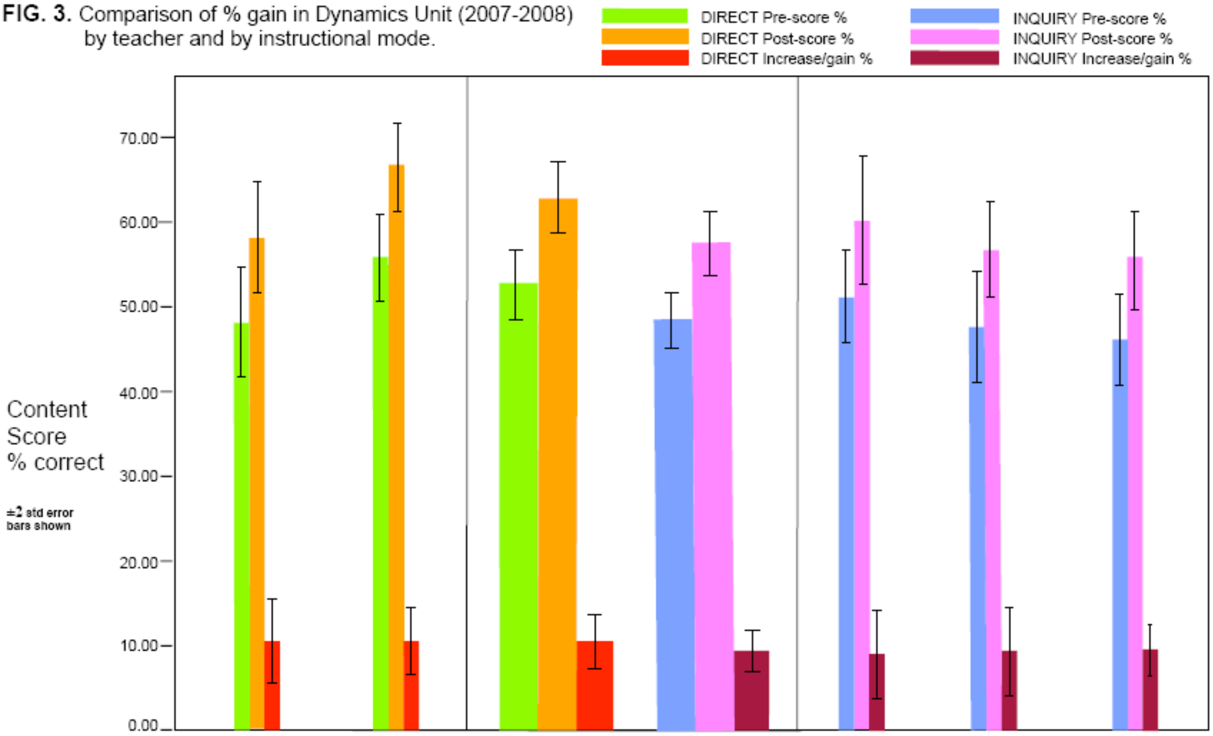


FIG. 2. Comparison of % gain in Light Unit (2007-2008) by teacher and by instructional mode.



Means (SD)	Joe	Ann	DIRECT	INQUIRY	Sam	Tom	Liz
	(N=32)	(N=40)	(N=72)	(N=108)	(N=37)	(N=35)	(N=36)
Pre-score	50.1 (16.5)	59.4 (17.9)	55.3 (17.8)	53.8 (20.0)	52.9 (22.7)	53.2 (19.3)	55.2 (18.3)
Post-score	65.5 (18.6)	70.5 (16.3)	68.2 (17.4)	67.9 (22.2)	64.5 (23.6)	71.8 (20.5)	67.6 (22.2)
% Increase	15.3 (13.2)	11.0 (13.8)	12.9 (13.6)	14.1 (16.4)	11.5 (16.6)	18.6 (16.8)	12.4 (15.3)
Normalized % change	32.5 (28.3)	29.9 (30.8)	31.1 (29.5)	34.9 (35.9)	31.5 (38.4)	43.6 (32.9)	30.0 (35.4)

FIG. 3. Comparison of % gain in Dynamics Unit (2007-2008) by teacher and by instructional mode.



Means (SD)	Joe (N=32)	Ann (N=40)	DIRECT (N=72)	INQUIRY (N=108)	Sam (N=37)	Tom (N=36)	Liz (N=35)
Pre-score	48.1 (18.0)	54.6 (14.9)	51.7 (16.6)	48.2 (17.3)	51.2 (16.7)	47.4 (19.3)	46.0 (15.8)
Post-score	58.2 (18.5)	65.5 (16.2)	62.2 (17.5)	57.4 (19.4)	60.1 (23.2)	56.6 (17.4)	55.4 (17.1)
% Increase	10.1 (13.9)	10.8 (13.2)	**10.5** (13.4)	**9.2** (13.6)	8.9 (15.5)	9.3 (15.5)	9.4 (8.8)
Normalized % change	19.6 (28.8)	24.1 (27.8)	22.1 (28.2)	19.0 (29.2)	21.8 (37.9)	16.3 (27.6)	18.7 (19.1)