

Comparing Traditional and Studio Courses through Gains and Losses

Jacquelyn J. Chini and Jarrad W. T. Pond

Department of Physics, University of Central Florida, 4000 Central Florida Boulevard, Orlando, FL 32816

Abstract: We analyze data from over 5500 students, spanning multiple semesters and instructors within a single institution, to explore how Gains and Losses in Force Concept Inventory scores vary between instructional approaches, including traditional lecture/laboratory courses, lecture-supported “mini-studio” and studio courses. We observe that the trend for Gain to increase and Loss to decrease with increasing pre-test score, initially observed across institutions, extends to individual student effects within an institution. The correlations between pre-test score and Gain/Loss were least pronounced for more minimally guided instructional strategies, such as studio. Additionally, we find similar or improved Gains and Losses for studio-type courses compared to traditional courses. These trends argue against the claim that novices require strongly guided instruction.

Keywords: physics education research, instructional strategies, conceptual understanding, introductory physics

PACS: 01.40.Fk, 01.40.Di

INTRODUCTION

In a recent paper, Lasry, Guillemette and Mazur [1] highlighted an intriguing trend of students switching correct pre-test Force Concept Inventory (FCI) [2] responses to incorrect responses on the post-test. This switching, or Loss, occurred at a higher rate than noise due to guessing [1]. The rate of Loss increased with decreasing pre-test score, while Gain decreased. The authors proposed several explanations for the high rate of Loss observed, such as confusion induced by instruction, fragility and context-dependence of the knowledge building process, and dependence of the ability to add new concepts on existing concepts. The authors suggested that Gain and Loss be explored for various instructional strategies to test these explanations and identify efficient pedagogies with high Gains and low Losses.

At the University of Central Florida (UCF), we began collecting FCI data in all first semester introductory courses in Spring 2011, across a variety of instructional methodologies (described below). This dataset allows us to explore how the trends observed by Lasry *et al.* differ by course type within a single institution and test hypotheses about the efficiency of teaching methods. For example, Kirschner, Sweller and Clark argue that novices benefit from strongly guided instruction rather than being required to construct concepts and procedures for themselves. [3]. Thus, we might expect to see a higher correlation between pre-test score and Gain or Loss for more minimally guided courses. In this paper, we explore the correlation of pre-test score with Gain and Loss and the variation in Gain and Loss across instructional strategies.

SAMPLE AND METHODOLOGY

We explore FCI results from the first course of the two-course introductory sequence at UCF, which covers typical mechanics topics, for both algebra- and calculus-based physics courses. These courses have been offered in traditional, honors, studio and lecture-supported mini-studio (mini-studio) modes. Traditional courses met for three hours of lecture (with 50 to 300 students) and three hours of laboratory per week (with approximately 30 students); in most cases, one hour of lab was used for recitation. The calculus-based course was also taught as a traditional honors section, with classes of 20 students from the university’s honors program. Both courses were additionally taught as more minimally guided studio courses (with 50 to 100 students). Studio courses are marked by flexibility to move between lecture, laboratory and recitation activities, with an emphasis on group work. At UCF, the algebra-based studio emphasized student-designed video analysis projects and most sections of the calculus-based course closely followed the SCALE-UP curriculum for *Matter and Interactions* with little lecturing [4]. The algebra-based course was additionally taught as a mini-studio, which combined three hours of lecture with three hours of student-centered engagement with concepts, problems and experiments, rather than a traditional recitation and laboratory. Table 1 summarizes the number of sections, instructors and students representing each mode.

To allow comparison to the Lasry data, we analyze performance on the FCI through question-level fractional Gain and Loss. (We note that the FCI may not be well matched to the curriculum used in the

TABLE 1. Summary of seven semesters' data

| Course Type | Level | Semesters | Sections | Instructors | Students |
|-------------|----------|-----------|----------|-------------|----------|
| Traditional | algebra | 5 | 8 | 6 | 786 |
| Studio | algebra | 7 | 10 | 3 | 553 |
| Mini-studio | algebra | 6 | 8 | 6 | 1306 |
| Traditional | calculus | 7 | 17 | 5 | 2066 |
| Studio | calculus | 6 | 11 | 4 | 697 |
| Honors | calculus | 6 | 7 | 9 | 116 |

calculus-based studio courses [5].) Following Lasry, Gain is defined as the ratio of questions corrected from pre-test to post-test to the total number incorrect on the pre-test [1]. Conversely, Loss is defined as the ratio of questions switched from correct on pre-test to incorrect on post-test to the total number correct on the pre-test.

Initial visual inspection of the data revealed that scores were not normally distributed, and standard transformations for percentage data did not improve normality. Non-parametric statistical tests and bootstrapped standard errors, effect sizes, and confidence intervals and robust Cohen's d_R were calculated with the bootES package for R [6, 7]. These methods and the specific statistical tests used below are described in Ref. 7 and 8.

DATA AND ANALYSIS

Pre-test Correlation with Gain and Loss

Lasry *et al.* found that, across institutions, students with higher pre-test scores gained more and lost fewer concepts from pre- to post-test [1]. We use our dataset to explore this trend within an institution and among different instructional strategies. Figures 1 and 2 display Gain (a) and Loss (b) versus pre-test score for the algebra- and calculus-based courses by course type. Due to the large sample size, these scatterplots are heavily populated. The R package ggplot2 [9] was used to layer semi-transparent data points to emphasize more and less heavily populated regions. The best fit line and 95% confidence interval (shaded region) were also added. These plots indicate that the previous trend observed extends to within institution effects and holds across course types.

We explore the correlations with a non-parametric version of Pearson's r , Kendall's τ ; results, displayed in Table 2, indicate all correlations are significant, with effect sizes ranging from small to large (small: $\tau \sim .06$; medium: $\tau \sim .19$; large: $\tau > .33$) [10]. The pre-test score is more strongly correlated to Loss than Gain across all courses, and all correlations were stronger for the calculus-based than the algebra-based course. In most cases, correlations are slightly smaller with similar slopes for the less strongly guided courses, such as studio, indicating novices fare well in these courses.

Gain and Loss across Course Structures

Lasry *et al.* suggested that Gain and Loss analysis could be used to assess the efficiency of instructional approaches. We use our dataset to compare course

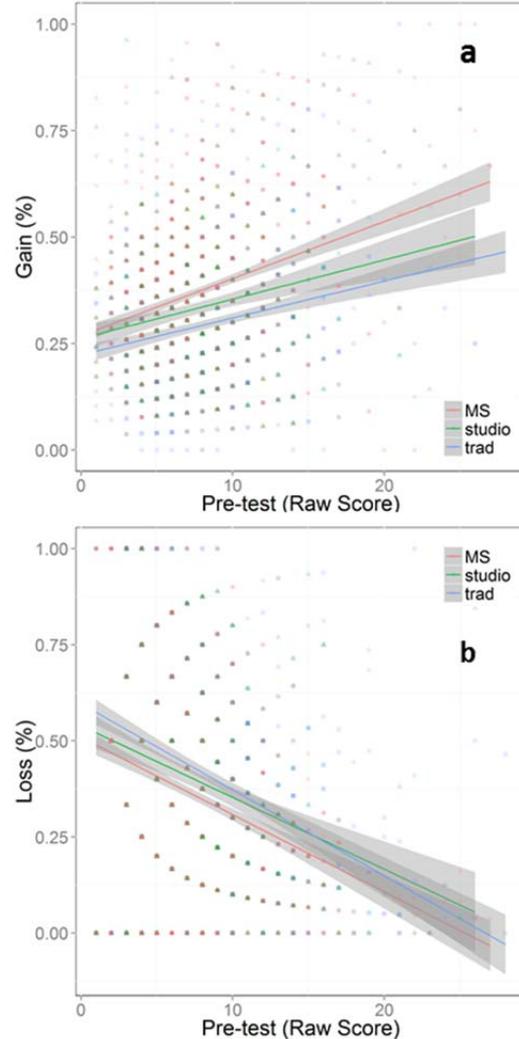


Figure 1. Pre-test correlation with Gain(a) and Loss (b) for algebra-based courses. Shaded area represents 95% confidence interval.

TABLE 2. Summary of correlation analysis.

| Course Type | df | Pre-test & Gain | | Pre-test & Loss | |
|-------------------------------|------|-----------------|-------|-----------------|-------|
| | | τ | p | τ | p |
| <i>Algebra-based courses</i> | | | | | |
| Traditional | 784 | .15 | <.001 | -.29 | <.001 |
| Studio | 551 | .09 | .003 | -.21 | <.001 |
| Mini-studio | 1304 | .15 | <.001 | -.23 | <.001 |
| <i>Calculus-based courses</i> | | | | | |
| Traditional | 2055 | .30 | <.001 | -.37 | <.001 |
| Studio | 695 | .28 | <.001 | -.41 | <.001 |
| Honors | 114 | .30 | <.001 | -.41 | <.001 |

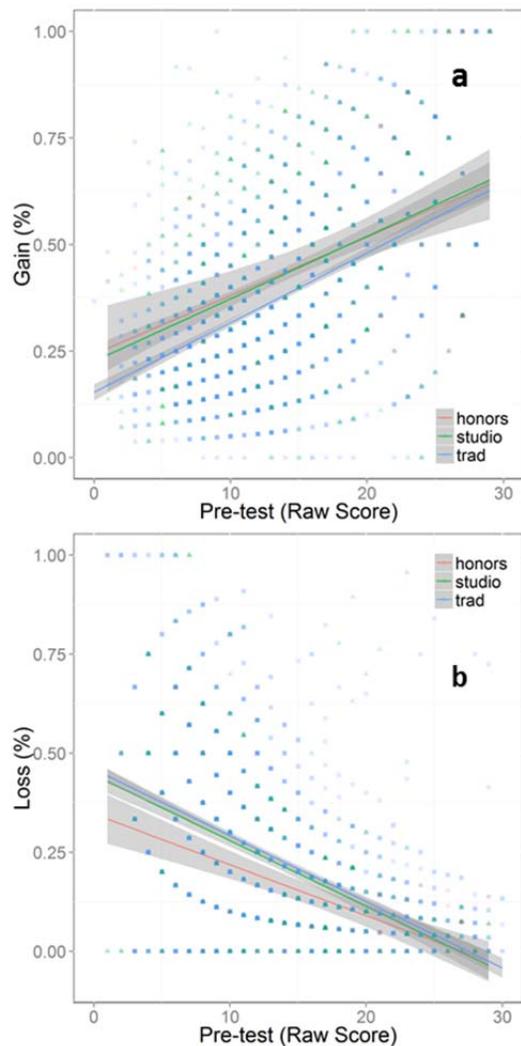


Figure 2. Pre-test correlation with Gain(a) and Loss (b) for calculus-based courses.

types and explore whether the Gain and Loss analyses yield similar insight into course effectiveness. Figures 3 and 4 display the differential Gain (a) and Loss (b) by instructional method for each section of the algebra- and calculus-based courses. The overall average for each was calculated across all sections (weighted evenly for each section), and individual section averages are compared to the grand average. Exploring Fig. 3, we observe that in the algebra-based course, the mini-studio sections typically performed better than average, most traditional sections performed worse than average, and there was a wide variety in the performance of studio sections on Gain and Loss. For the calculus-based course, the honors sections typically performed above average, most traditional sections performed below average and studio sections were scattered about the average. Both measures lead to similar conclusions about the effectiveness of the

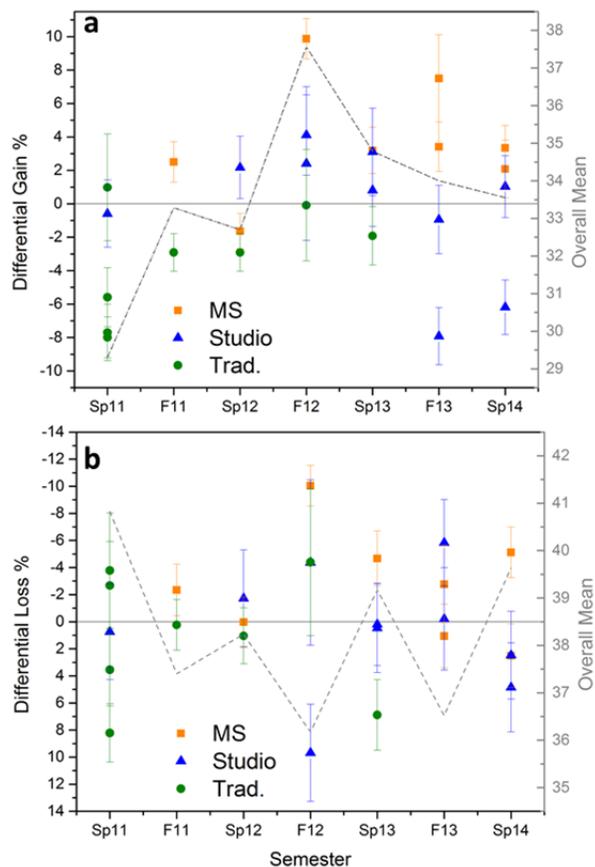


Figure 3. Differential Gain (a) and Loss (b) for algebra-based courses. Average Gain = 33.5%. Average Loss = 38.4%. Error bars represent bootstrapped standard error. Dashed line indicates individual semester averages

course structures. However, the Loss analysis shows the greatest stability across semesters and instructors for the calculus-based courses. Individual sections show some variation across measures; for example, the Fall 2012 algebra-based traditional section had average Gain but better than average Loss.

Non-parametric tests were used as the data violate normality assumptions. However, density plots revealed that certain groups of interest had very similar distributions of pre-test scores. Thus, we proceed with the Kruskal-Wallis test, a non-parametric ANOVA-type test, while noting this does not account for differences in pre-test. For the algebra-based course, there were significant differences by instructional method for Gain ($\chi^2(2)=101$, $p<.001$) and Loss ($\chi^2(2)=28.5$, $p<.001$). Wilcoxon-Mann-Whitney tests with Bonferroni corrections and calculations of robust Cohen's d_R and confidence intervals were used to identify differences between sections. The results are summarized in Table 3, and indicate significant differences between methods for all three measures, except for Loss between studio and traditional courses.

Table 3. Summary of Wilcoxon-Mann-Whitney and robust Cohen's d_R results for algebra-based course. MS=mini-studio. * $p < .05$, † $p < .01$, ‡ $p < .001$

| | Gain | | Loss | |
|--------------------|------------------|---------|------------------|----------|
| | d_R | CI | d_R | CI |
| Studio-Traditional | .21 [‡] | .10-.32 | .06 | -.06-.18 |
| MS-Traditional | .47 [‡] | .37-.56 | .25 [‡] | .16-.34 |
| MS-Studio | .26 [‡] | .15-.37 | .19 [*] | .08-.30 |

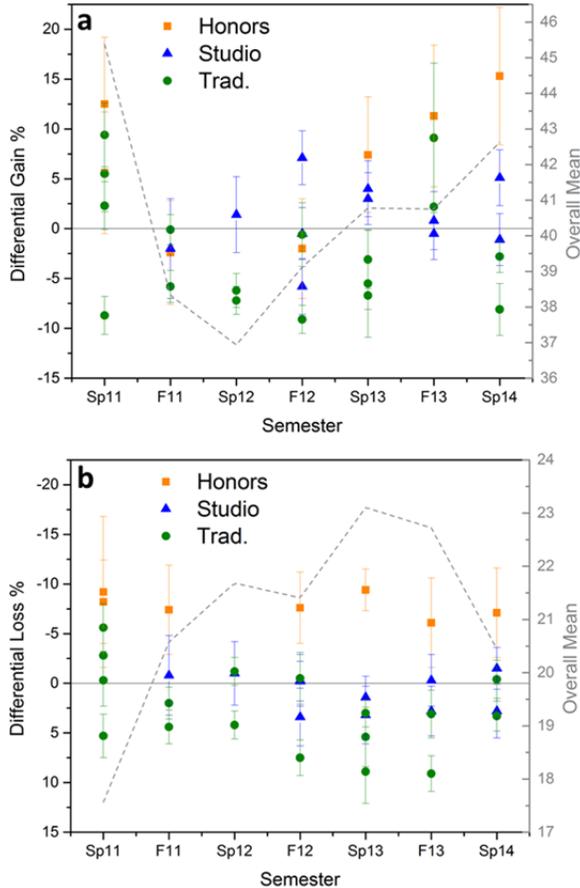


Figure 4. Differential Gain (a) and Loss (b) for calculus-based courses. Average gain = 40.9%. Average loss = 21.0%. Error bars represent bootstrapped standard error. Dashed line indicates individual semester averages.

We observe the strongest differences between mini-studio and traditional sections, and only weak differences in Loss scores across all three comparisons.

For the calculus-based course, the honors students had higher pre-test scores than students in traditional and studio courses. Thus, we focus on only the traditional and studio courses here. The Wilcoxon-Mann-Whitney test indicated a significant difference but small effect size between studio and traditional sections' Gain ($W=833735$, $p < .001$, $d_R=.14$ [.03, .26]) and Loss ($W=662961$, $p < .002$, $d_R=-.06$ [-.18, .06]). As with the algebra-based course, Gain and Loss scores give us similar information about teaching methods, with smaller differences in Loss than Gain.

DISCUSSION AND SUMMARY

We observed that the correlation of pre-test with Gain/Loss observed by Lasry *et al.* across institutions held at the level of individual students within an institution and extended across course types. Students in the calculus-based courses had higher Gains and lower Losses than students in the algebra-based course on average (40.9% vs. 33.5% and -21.0% vs. -38.4%), which is expected given the typical preparation of students in these courses. However, it may be somewhat surprising that the correlations were stronger and steeper between pre-test and Gain/Loss in the calculus-based courses than the algebra-based courses.

We also observed differences in Gain and Loss between course structures, with stronger differences in Gains than Losses. In the algebra-based course, the mini-studio performed better than traditional and studio courses on both Gain and Loss and the studio performed better than the traditional course on Gain. In the calculus-based course, the studio performed somewhat better than the traditional course on Gain and Loss. These findings refute Kirschner *et al.*'s contention that novices benefit from more strongly guided instruction [3]; pedagogies that put more emphasis on student guided learning performed better than the strongly guided traditional course.

ACKNOWLEDGMENTS

This work is supported in part by National Science Foundation grants DUE1347515 and DUE1246024 and a PhysTEC grant.

REFERENCES

- [1] N. Lasry, J. Guillemette, and E. Mazur, *Nat. Phys.*, **10**, 402-403 (2014).
- [2] D. Hestenes, M. Wells and G. Swackhamer, *Phys. Teach.*, **30**, 141-158 (1992).
- [3] P.A. Kirschner, J. Sweller and R. E. Clark, *Educ. Psychol.*, **41**, 75-86 (2006).
- [4] <http://scaleup.ncsu.edu/>
- [5] M. Caballero *et al.*, *Am. J. Phys.*, **80**, 638-644 (2012).
- [6] R Core Team. R: A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria, 2014).
- [7] D. Gerlanc and K. Kirby, bootES: Bootstrap Effect Sizes (2013).
- [8] M. Hollander, D. A. Wolfe and E. Chicken, *Nonparametric statistical methods*, Vol 751 (Wiley, Hoboken, NJ, 2014).
- [9] H. Wickham. ggplot2: elegant graphics for data analysis. (Springer, New York, 2009).
- [10] D. Walker, *J. Mod. Appl. Stat. Methods*, **2**, 525-530 (2003).