Implementing a Supplemental Instruction Program for Introductory Statistics at a Regional Campus

Dywayne Nicely
Ohio University Chillicothe

Introductory Statistics has been categorized as a high-risk course on our campus due to the fact that students have difficulty completing this course. Thus, a Supplemental Instruction (SI) program was instituted to help students become more successful with this traditionally difficult course and an experiment was conducted to test the effectiveness of the SI program. Students enrolled in Introductory Statistics for the 2012-13 academic year served as the control group (non-SI group) and the experimental group (SI group) was populated with students enrolled in Introductory Statistics for the 2013-14 academic year. This paper shows that the SI group performed better than those from the non-SI group. Furthermore, this paper shows how students who regularly participated in the SI sessions performed better than students who occasionally participated or did not participate in SI sessions. A t-test was performed ($p = 0.1116$) and the effect size ($d = 0.43$) was calculated to validate data.

Introduction

Many students struggle with and are unsuccessful in Introductory Statistics. We attempted to help remedy this situation by implementing a SI program. SI is a voluntary program that was developed in 1973 at the University of Missouri–Kansas City (UMKC), and UMKC is home to the International Center of Supplemental Instruction (ICSI).

ICSI (2015) continues to define SI as, “a non-remedial approach that targets ‘high-risk courses’ rather than ‘high-risk students.’” Research has shown that students are more successful academically and withdraw less when participating in SI (Congos & Schoeps, 1993; ICSI, 2003; Oja, 2012; Simpson et al., 1997). However, McCarthy and Smuts (1997) point out that the more academically capable students participate in SI. This is an issue because it stands to reason that the less academically talented students should participate in SI and would benefit more from it.

For our campus, SI is used as a series of weekly review sessions offered to students enrolled in historically difficult classes. These regularly scheduled free study sessions, called SI sessions, are conducted outside of the classroom by an SI Leader. SI Leaders are undergraduate students who have previously performed well in the course and who attend all class lectures to take notes and act as a model for the students who are enrolled in the course. Students are told that the SI sessions can be thought of as a time to get together with classmates to compare notes, review key concepts, and develop effective strategies for studying and preparing for quizzes and tests. In addition, we challenge the students to ask themselves if they have the drive and motivation to work hard and excel in these challenging courses.

In this paper, we will give details on the SI sessions in an introductory statistics course. Also, data will support the notion that students who had the opportunity to utilize the SI sessions perform better than those who did not. Moreover, we will show how those students who regularly attended SI sessions outperformed their classmates who attended SI sessions sporadically or not at all. Data was collected from the 2012-2013 and 2013-14 academic years. Students from 2012-13 academic year represent the control group (Non-SI group), and students from the 2013-2014 academic year represent the experimental group (SI group).

High-Risk Courses and Some Difficulties in Finding SI Leaders

As mentioned above, SI focuses on high-risk courses instead of high-risk students. Currently there are six courses (listed alphabetically) that are identified as high-risk for our campus.

- Elementary Statistical Reasoning
- Foundations of Accounting
- Fundamentals of Chemistry I
- Introductory Statistics
- Principles of Chemistry I
- Statistics for the Behavioral Sciences.

All six of the high-risk courses involve some mathematics. So all of our SI Leaders need to be highly competent in various areas of mathematics, especially statistics.

This highlights one major difficulty that instructors of high-risk courses can have when at a small regional campus. Since enrollments are small, the talent pool from which to select a competent SI Leader is also small. In addition, regional campus student populations are composed of many non-traditional students and students with very different priorities and responsibilities than the typical college student. Some of these students might be ideal candidates for SI Leaders, but their personal obligations might not allow for time needed to serve as an SI Leader.

SI Sessions in Introductory Statistics

During the first week of the term, students are introduced to the SI Leader. The SI Leader surveys the students to determine the best times to conduct SI sessions. It is explained that attendance for the SI sessions is voluntary. Voluntary attendance aligns with the philosophy of the SI program, but, we offer, that others have made SI sessions mandatory with mixed results (Porter, 2010).

SI sessions tended to be held immediately after the end of classroom lecture. Generally, there were two SI sessions per week. Though, for weeks that involved an exam, there was an increase to three or four SI sessions. SI sessions were held in a room adjacent to the regular classroom or in our Academic Success Center, which can be thought of as a tutoring center. In the SI sessions, the SI Leader acts as a facilitator to motivate discussion but not as a pseudo-lecturer. Hence SI sessions, by design, are less formal and more relaxed than in-class lecture.

Students did review concepts in the SI sessions but, the SI Leader reported, that the majority of the time spent was devoted to reviewing examples discussed in the lectures and, more so, to homework exercises. Since the majority of what was discussed in the SI session was devoted to homework exercises, we point out that homework was not an element of the students' final grade. Therefore those students who attended the SI sessions did not have an unfair advantage over the students who did not attend the SI sessions when grades were calculated.

Participants and Results

Students who traditionally enroll in Introductory Statistics on our campus are either pursing a two-year or a four-year degree. Students from the two-year degrees are typically majoring in Computer Technology, Environmental Engineering Technology, or Nursing. For the four-year degrees, we see students from the academic programs of Nursing or Middle Childhood Education.

The non-SI group served as the control group for our experiment and this group was populated with students from the 2012-13 academic year. Students from the SI group, which served as the experimental group, were populated from those enrolled during the 2013-14 academic year. Our enrollments allow for only one section of statistics during each term of the academic year but, this turned out to be an advantage in the experimental design. Since the control and experimental groups did not run concurrently, this removed the chance of intermingling between the two groups and helped to reduce bias.

The next two tables give final grade data (in terms of a weighted average and separated by semester) for the two groups. Grades were calculated based on the following weights.

- Quizzes 25%
- Three Hourly Exams 45%
- Final Exam 30%

Table 1 displays the data for the non-SI group and Table 2 displays data for the SI group. For both tables, the variable *n* denotes the sample size.

Table 1. Class averages for the non-SI groups

Group	Average-Fall 2013 (%)	Number-Fall 2013	Average-Spring 2014 (%)	Number-Spring 2014
Entire Class	59.03	20	63.36	16
Completed Course	65.47	16	73.51	12

The information in Table 1 shows the overall averages for the entire class, and the overall average of those students who completed the course from the non-SI group. A student who completed the course is defined as a student who took the final exam. The overall averages in Table 1 show that the Fall 2012 students, as a whole, earned a grade of F, but improved somewhat to a grade of D when reviewing the students who completed the course. Better results were seen in the spring term. Students from Spring 2013 have a grade of D, collectively, but a solid C grade when looking at the students who completed the course.

Table 2. Class averages for the SI groups

Group	Average- Fall 2013 (%)	Number- Fall 2013	Average- Spring 2014 (%)	Number- Spring 2014
Entire Class	68.62	18	69.16	14
Completed Course	74.52	16	77.02	12
R-SI Subgroup	78.82	5	86.25	3
NR-SI Subgroup	64.69	13	64.50	11
NR-SI Subgroup (Completed Course)	72.20	11	73.94	9

Similar data displayed in Table 2 for the SI group shows that the Fall 2013 students received a grade of high-D, as a whole, and a solid grade of C from those students who completed the course. Like the non-SI group, the SI group saw higher scores in the spring term. Collectively the Spring 2014 students received a slightly higher grade of D than the Fall 2013 students and, for those students who completed the course, there is an increase of a mid-C to a high-C when comparing Fall 2013 and Spring 2014. Also noteworthy, is the near 8% increase and high-C seen from the students who completed the course in Spring 2014 when comparing to the overall average of the Spring 2014 students.

Additional data in Table 2, not given in Table 1, shows the averages of regular SI participants (R-SI), all of whom completed the course, and non-regular SI participants (NR-SI). A regular SI participant was defined as a student who attended at least one SI session per week. Mainly the students who populated the NR-SI subgroup attended SI sessions only before exams. All attendance records for the SI sessions were recorded and verified by the SI Leader.

Notable improvements in the overall averages are shown in Table 2 when comparing the R-SI and NR-SI subgroups. For Fall 2013, an increase of over 14% in the overall average is seen from the R-SI subgroup versus the NR-SI subgroup. The trend of realizing better results in the spring terms continues here and we see an increase of nearly 22% between the R-SI and NR-SI subgroups. Also results are still favorable (respective increases of over 6% and 12% in the fall and spring terms) for regular SI attendance when we compare the R-SI subgroup to the NR-SI subgroup of students who completed the course.

Further analysis of the R-SI and NR-SI subgroups does lead to a concern. Only a quarter of the students (8 out of 32) that had the opportunity to take advantage of the SI sessions did so. This corresponds to the findings of McCarthy and Smuts (1997) who found that more academically proficient students participate show greater participation in SI than their counterparts.

A *t*-test ($\alpha = 0.05$) was performed to determine whether the increase in the overall average seen from the SI group was statistically significant. The results of the *t*-test are shown in Table 3. The means for the non-SI (2013-14) and SI groups (2012-13) were calculated cumulatively by combining fall and spring data from the academic years for the *t*-test. Also only those students who had completed the course were used in calculating the cumulative means and this resulted in identical sample sizes (*n* = 28) for the two groups.

Table 3. T-test comparing cumulative mean scores between the two groups

Group	Mean (%)	Standard Deviation	Standard Error	Sample size
non-SI group	68.92	18.53	3.50	28
SI-group	75.59	11.54	2.18	28

Notes: *t* = 1.6177, *df* = 54, and *standard error of difference* = 4.125

From Table 3 we see that the SI group's overall average was over 6.5% greater than the overall average of the non-SI group, but the resulting *p*-value of 0.1116 from the *t*-test means that this increase is not statistically significant. This could be due to the small sample sizes that we are using in this study. The *t*-test was developed to deal with small sample sizes, but a small sample size can still affect the statistical significance (a higher *p*-value) of a result because small sample sizes increase uncertainty. On the other hand, a larger sample size has more information so uncertainty is decreased and, consequently, the confidence in your estimate of the particular parameter is increased. Basically, smaller sample sizes decrease the confidence in the estimate of a certain parameter (the mean in our experiment) and can result in a greater likelihood that randomness or chance could have played a role in the outcome of the experiment (the improvement seen in the mean scores in our experiment).

However, statistical significance is not the only way of validating results from an experiment. Practical significance is another measure used for validation, especially in educational studies. Therefore, we calculated the effect size to determine if the improvement seen from the SI group was practically significant. Specifically, Cohen's *d* was used to calculate the effect size. Cohen's *d* is calculated from a formula that standardizes (using the standard deviations from the experimental and control groups, SD_2 and SD_1 respectively) the difference between the two means (M_2 from the experimental group and M_1 from the control group),

$$d = \frac{M_2 - M_1}{SD_{pooled}}, \text{ where } SD_{pooled} = \sqrt{\frac{SD_1^2 + SD_2^2}{2}}$$

Using the means and standard deviations from Table 3 resulted in an effect size of *d* = 0.43. A *d* of 0.43 denotes a near-medium effect size and the improvement in overall average from the SI group is visibly seen and practically significant, but a larger sample size could be necessary for further validation.

Even with acknowledging the above, our experiment was worthwhile and corroborated from prior research. Hattie (2012) argued that for an educational intervention to be considered successful an effect size needs to be larger than 0.4, the hinge-point (or h-point). Our calculated effect size of *d* = 0.43 is larger than Hattie's h-point. Therefore we can conclude that the implementation of an S

Table 4. Grade distributions for both groups

Group	A	B	C	D	F	W
non-SI group	3	8	6	5	14	10
SI group	4	10	7	3	6	15

The final data displayed, shown in Table 4 gives the grade distributions for the two groups. In Table 4 we see that As, Bs, and Cs increased and the Ds and Fs decreased (Fs more considerably) in the SI group when compared to the non-SI group. These results are positive and are the types of results that we had hoped to see. Yet, we do note that the number of withdrawals increased in the SI group. This is antithetical to previous research and is a major cause of concern. Further study is required into this phenomenon, but one explanation for the increase in withdrawals could be a result of the small sample sizes used in our experiment.

Conclusions and Areas for Improvement

We instituted a SI program for our introductory statistics course in hopes that our students would become more successful in this high-risk course. Data shown implies that students perform better when SI is available and those students who take advantage of the SI sessions perform much better than students who do not take advantage of SI. Also we saw increases in the grade distributions for the grades of A, B, and C and a substantial decrease in the number of Fs from students in the SI group.

Areas for improvement include measures to decrease the number of withdrawals in introductory statistics and to better promote the advantages of the SI program to the students so that there are more students regularly participating in the SI sessions. In addition, we concede that the effectiveness of the SI program may require further study for increased validation and to decrease the number of withdrawals. Nevertheless, we saw positive and practically significant results, as suggested by our calculated effect size of $d = 0.43$, from implementing an SI program for introductory statistics. We hope that others will implement SI programs for their targeted “high-risk courses” as a way to improve student success.

References

Congos, D. H., & Schoeps, N. (1993). Does supplemental instruction really work and what is it? *Studies in Higher Education*, *18(2)*, 165-182.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.

International Center for Supplemental Instruction (ICSI). (2015). Overview of Supplemental Instruction. Retrieved March 20, 2015, from http://www.umkc.edu/asm/si/overview.shtml.

McCarthy, A., & Smuts, B. (1997). Assessing the effectiveness of supplemental instruction: A critique and a case study. *Studies in Higher Education*, *22(2)*, 221-235.

Oja, M. (2012). Supplemental instruction improves grades but not persistence. *College Student Journal*, *46(2)*, 344-349.

Porter, R. C. (2010). The effects of supplemental instruction on student achievement in College Algebra. *Georgia Journal of Science*, *68(3)*, 124-131.

Simpson, M. L., Hynd, C. R., Nist, S. L., & Burrell, K. I. (1997). College academic assistance programs and practices. *Educational Psychology Review*, *9(1)*, 39-87.

Dywayne Nicely, Ohio University Chillicothe: nicely@ohio.edu