From Research to Practice: Using Assessment and Early Intervention to Improve Student Success in Introductory Statistics

M. Leigh Lunsford\textsuperscript{a}, Phillip L. Poplin\textsuperscript{a}, and JoEllen G. Pederson\textsuperscript{b}

\textsuperscript{a}Department of Mathematics and Computer Science, Longwood University, Farmville, VA; \textsuperscript{b}Department of Sociology, Anthropology and Criminal Justice Studies, Longwood University, Farmville, VA

\textbf{ABSTRACT}

In this study we used an assessment tool evaluated in a previous study to identify students who were at-risk of not being successful in our introductory statistics course. We then required these students to attend peer tutoring, early in the semester, as an intervention. While we saw a significant increase in student success for all students in this study compared with the previous study, the at-risk students who completed the required tutoring had a significantly higher increase in success than their peers.

\textbf{KEYWORDS}

Assessment; Basic mathematical skills; Intervention; Introductory statistics; Peer tutoring

1. Introduction

In 2006 Johnson and Kuennen published a paper in this journal about a study they conducted to identify student skills and characteristics that may be associated with success in their introductory business statistics course. Using data collected from ten sections of their introductory level statistics course “Economics and Business Statistics” taught by three professors, they found the most important determinates of success (as measured by the grade received in the course) were “…student gender, GPA, ACT science score, and score on a quiz of basic math skills” (Johnson and Kuennen 2006). They also found a “professor effect” which meant when controlling for skill level, students were more likely to be successful with one professor than another. Motivated by that paper and by the high failure rates in our introductory level statistics course (about 40% of students not completing the course with a grade of C or better), we conducted a similar study to better understand our students. Given that our introductory level statistics course (MATH 171 Statistical Decision Making) is algebraically and computationally light, we were particularly interested in Johnson and Kuennen’s result that the score on the quiz of basic mathematics skills was a consistent predictor of success even when controlling for course format and professor. Our study spanned four semesters, included five professors, and 702 students. We deemed a student “successful” if they received a grade of a C or higher in the course and we used a slightly modified version of Johnson and Keunnen’s basic skills mathematics quiz (BSMQ). In 2011 we published a paper in this journal (Lunsford and Poplin 2011) in which we detailed the results of this study, which were similar to those of Johnson and Keunnen’s. Namely, students’ basic mathematical skills, as measured by the BSMQ, were a significant predictor of their success in our course. In addition, we found a professor effect (i.e., when controlling for BSMQ score, students were more likely to succeed with one professor than another).

Based on those findings, we began to use what we had learned to improve student success rates in our introductory statistics class and to alleviate the professor effect (please see the discussion section of Lunsford and Poplin 2011). In particular, we used the BSMQ to identify students who were less likely to be successful in our course. We will refer to these as our “at-risk” students. We then began recommending to these students that they start attending tutoring early in the semester. Evidence from previous research (Topping 1996) and our own anecdotal experience suggests that tutoring within higher education can be a highly effective tool in increasing student grades and achievement.

Our issues with success rates, now defined by passing MATH 171 with a C- or higher (during this time period our University adopted a plus and minus grading system), are not unique. Many students taking introductory level statistics classes in college are doing so because it is a requirement of their major. Studies have shown these students often have high levels of anxiety about taking a math course and low levels of interest in the course material (Gafield and Ben-Ziv 2007; Suanpang, Petocz, and Kalceff 2004). The difficulties associated with keeping nonmathematics majors engaged with course material has led to research on best practices for teaching and reaching statistics students. Several studies suggest that cooperative learning, such as group work or peer-tutoring, is a successful strategy at improving retention and grades in statistics classes (Keeler and Steinhorst 1995; Giraud 1997; Magel 1998). Topping (1996) argues that peer tutoring, “…typically has high focus on curriculum content. Projects usually also outline quite
specific procedures for interaction, in which the participants are likely to have training which is specific or generic or both” (p. 322). While advantages can be understood through a social perspective for both the tutor and the tutee, the advantages for the tutee include, “more active, interactive and participative learning, immediate feedback, swift prompting, lowered anxiety… and greater student ownership of the learning process (Topping 1996, p. 325). Research examining the effectiveness of peer tutoring at the collegiate level is vast. Metaanalyses of this research show that there is, “substantial evidence that peer tutoring is effective in schools” (Topping 1996, p. 326) in terms of better academic outcomes for students, cost, and time efficiency for additional instruction. More specifically, one metanalysis examined 82 studies that researched the effects of peer tutoring in high school and found substantial cognitive gains for both tutees and tutors (Sharples and Sharples 1981). A meta-analysis of studies with control groups found that out of 65 studies, tutored students outperformed control in 45 (Cohen, Kulik, and Kulik 1982). More recently, a metaanalysis was done on articles that used both fixed and mixed effects models for analyzing the effect size of a tutoring intervention, also accounting for other peer learning effects. This metaanalysis also concludes that peer tutoring has a positive impact on academic achievement (Leung 2015).

While a large amount of research examines peer tutoring in the college setting, very little research looks at required tutoring for students in statistics classes. Leppink et al. (2013) found tutor guidance in problem-based learning for statistics an effective addition to the course. Effectiveness in this study was based on students learning a conceptual understanding of statistics. Dancer et al. (2015) examined the effects of peer-assisted study sessions (PASS) in a first-year business statistics course and found that PASS significantly helped all students but the effect was more pronounced for international students and lower-achieving students. Giraud (1997) found that students who took a statistics class that required cooperative learning did better on tests than a class that relied solely on a lecture method of instruction, with both classes were taught by the same professor.

In the Fall of 2011 we started a new study which we completed in the Spring of 2014. In this study we implemented a cooperative learning intervention in the form of peer tutoring for all of our at-risk students (explained in more detail in Section 3). The results were encouraging: We had an overall increase in success compared to our first study. More interestingly, the increase in success for our at-risk students was significantly higher than for our other students. We also saw a continued professor effect, but it had decreased among the professors who were in both studies. In this article we will give some background information about our introductory statistics course and the results of our first study. We will then describe our methodology and results from this new study and finally discuss ramifications and recommendations from our work.

### 2. Background Information

#### 2.1. Course Description and Issues

Statistical Decision Making (MATH 171) at LU is a three semester-hour non-calculus-based freshman-level introduction to statistics course. The course is part of LU’s General Education Curriculum and thus has no prerequisites. Course content and emphasis has followed the American Statistical Association (ASA) endorsed Guidelines for Assessment and Instruction in Statistics Education (GAISE) (http://www.amstat.org/education/gaise/GAISECollege.htm). The course is algebraically and computationally light with all professors teaching the course using technology (e.g., the TI-84 series calculator, on-line applets, etc.), in-class activities, and real data sets. During the years 2011–2014, there were two textbooks used by the different faculty members. They were *Essential Statistics*, by David Moore, and *The Basic Practice of Statistics*, by David Moore.

Over the years we have seen a substantial increase in the number of students taking the course due to other disciplines, particularly psychology and business, requiring the course for their students (Table 1).

| Academic year | Fall semester | Spring semester | Total (# sections) |
|---------------|---------------|-----------------|-------------------|
| 2006/07       | 202           | 95              | 297 (9)           |
| 2007/08       | 265           | 198             | 463 (15)          |
| 2008/09       | 155           | 179             | 334 (11)          |
| 2009/10       | 194           | 125             | 319 (11)          |
| 2010/11       | 257           | 166             | 423 (13)          |
| 2011/12       | 285           | 165             | 455 (14)          |
| 2012/13       | 340           | 161             | 496 (17)          |
| 2013/14       | 412           | 189             | 601 (21)          |
| 2014/15       | 479           | 172             | 651 (21)          |
| 2015/16       | 452           | 228             | 680 (23)          |

With this increasing enrollment in MATH 171 and our administration’s emphasis on retention and graduation rates, success rates have become an issue. Students who do not receive a C- or higher in this course cannot take our second non-calculus-based course in statistics, MATH 301 Applied Statistics, which is also being required by more majors including all business majors. A low grade in MATH 171 can also affect a student’s overall GPA to a point where they may be put on academic probation and hence be in danger of not being retained. In addition, with more sections and thus more professors teaching the course, we believe it is important to maintain consistency, quality, and fairness in the course. Introductory statistics is a deceptively difficult course to teach (Conners, Mccown, and Roskos-Ewoldson 1998; Garfield and Ben-Zvi 2007). While none of the faculty in our department has degrees in statistics, almost all of the faculty teaching the course have been engaged in some form of professional development aimed at improving introductory level statistics teaching. This includes minicourses and workshops offered at professional meetings, summer stand-alone workshops, and serving as readers for the Advanced Placement (AP) Statistics Exam.

#### 2.2. Results from Previous Study

Our first study (Lunsford and Poplin 2011) was conducted over two academic years (Fall 2006 through Spring 2008), involved 741 students, and five professors. Students were given the BSMQ on the first day of class. The quiz measures student facility with basic mathematics skills including performing simple
algebra and working with ratios. The original 15 question quiz developed by Johnson and Kuennen (2006) was used for the Fall 2006 semester and in the Spring 2007 semester we slightly modified this quiz by adding five questions (see Appendix A in the online supplemental files). In our 2011 paper we only described the results using the score on the 15 question quiz, though the results were similar for the three semesters in which we used the score on the 20 question quiz. We also measured other independent variables including the professor the student had for the course, whether or not the student attended tutoring in the Learning Center (LC), and if they did attend, how many hours they attended during the semester. At the end of the semester, we recorded the student’s final grade (ignoring plus and minus grades) and deemed a student to be successful if they earned a C or higher in the course.

Using binary logistic regression, we found the score on the BSMQ was a significant predictor of success. Since the logistic regression model uses the logit function, which is just the odds of success for the binary response variable, we saw that an increase of one point on the 15-point BSMQ increased the odds of success in the class by 23%. There was also a significant professor effect, meaning when controlling for BSMQ score, students were more likely to succeed with one professor than another. In Figure 1 we see graphs of the predictive logistic models for the probability of success for students overall (on the left) and separated by professor (on the right). Overall we see a positive relationship between the score on the basic skills test and success in the course with the extent of this relationship varying across professors. For instance, among students who had a score of 10 (50%) on the BSMQ, those who had Professor 2 had about a 31% chance, on average, of being successful in the class whereas those who had Professor 3 had about a 66% chance.

We were also surprised that tutoring was not a significant predictor of success but noted both how few students attended tutoring and whether a student attended tutoring was not independent of professor. Please see our 2011 paper for more details.

This first study was very informative. We found that even when teaching an algebraically and computationally light statistics course, basic mathematics skills, as measured by the BSMQ, are still important. We also became much more aware of the mathematical level of our students and no longer assumed our students had basic knowledge of ratios, algebra, and area. This affected our teaching as we began to take the time to review these mathematical concepts as we covered the statistics where they were used. We also began to question our assessment methods and saw a need for regular daily assessment in addition to in-class tests. Many of us now use on-line homework systems to help students keep up with the material (Lunsford and Pendergrass 2016). Lastly, once we became aware of the professor effect we made an effort to share more of our course materials as well as improve a set of required LU General Education common final exam questions to better reflect the content and conceptual emphasis in the course.

3. Our New Study

One of the goals of our first study was to find “an instrument that would be a quick indicator of student success in the class and would give us an easy and early way to identify students who were likely to have problems in the class” (Lunsford and Poplin 2011). Another was to use the results of the study to “provide information on possible ways to improve our teaching of introductory statistics and ultimately, to increase student success in the course.” Certainly being aware of our students’ mathematical level and the professor effect caused us to change some of our teaching practices. However, while we were disappointed that attending tutoring was not associated with success in the first study, anecdotally we felt that tutoring did indeed help our weaker students. Current and previous research on statistics education supports our reasoning. Peer tutoring in statistics has been shown to improve passing rates in mathematics placement exams (Garcia, Morales, and Rivera 2014), improve retention in mathematics courses (Roberts and Baugher 2015), and increase final exam grades (Bude et al. 2008; Giraud 1997). Now that we had a way to identify our at-risk students, we decided to try an early intervention that we hoped would help them succeed.

The early intervention we eventually settled upon was requiring all of our at-risk students to obtain peer tutoring in LU’s Center for Academic Success (CAS) (a new name for the...
For each student we recorded their score on the 20-question BSMQ. We also recorded the semester the student was enrolled, which professor the student had for the course, and how many hours each student spent getting tutoring in the CAS. If available, we included SAT scores (SATM, SATR) and/or ACT scores for each student. As stated above, we wished to find an instrument that would give a quick indicator of student success which the BSMQ was shown to achieve in the first study. However, we also wanted to measure other possible indicators and potentially compare them to the BSMQ, in particular the SAT scores. Unfortunately it is difficult to get this information at LU, and many students were missing these scores, especially students that transferred to LU from another institution.

Our independent variable was the binary variable “success.” As in our first study, student success was determined by their letter grade in the course. We determined a student to be successful if they earned a grade of C (ignoring plus and minus grades) or higher in the course. Students who took the basic skills test but dropped from the course before the course drop/add deadline (usually around the beginning of the second week of classes) do not appear on the final class rosters at LU and thus these students were not included in the study. As in our first study, students who withdrew from the course after this deadline, and thus received a grade of “W,” were included in the analysis but were not considered to be successful in the course. We argued then, and now, that these students had invested more time in the course yet did not complete the course with a C- or better.

3.3. Issues With the Data

In a study involving this many students and professors, it is not surprising that several issues arose. First, while the required intervention was a departmental policy, we did have one professor who did not participate in the study (essentially by not enforcing the policy). That professor and their students were not included in the analysis. There were also 81 student records that showed up as repeats, i.e., students who were not successful in the course, retook the class, and thus appear in our data set more than once. There were 77 students that retook the class once and two students that retook the class twice. As in our first study we argue that since these students took the BSMQ each time they took the class, these could be considered as independent attempts at passing the course, and hence each of those attempts was included in the study. Next, there were 42 students (46 student records, including repeats) who scored 50% or lower on the BSMQ who did not complete the required tutoring but did receive a grade in the course. Since the purpose of this study is to examine if there was any association between success and required tutoring, these students were not included in the study. We will discuss these students in more detail in Section 5.1. Finally, two students received a grade in the course but did not take the BSMQ. Those students have also been removed from our analysis.

After cleaning the data, there were seven professors with a total of 1508 student records (including the 81 repeats) in the study during the 3-year period. The number of students taught by each professor during the study is given in Table 2. We note that Professors 1, 2, and 3 taught the most students and were...
also in the first study. We will refer to these three as our “more experienced” statistics professors. Also, note that Professor 6, an adjunct, only taught one section during the study.

4. Results from Second Study

In comparing our second study to the first study it is important to reemphasize that we used the 20 question BSMQ. Thus, unlike our 2011 paper, we will use results from the Spring 2007 to the Spring 2008 semesters of the first study as we only used the 15 question BSMQ in the Fall of 2006. In Table 3 you see a comparison of the two studies in terms of number of semesters, student records, and professors in each.

4.1. The Basic Skills Test, Early Intervention, and Student Success

Using binary logistic regression, we found that the score on the BSMQ was still a predictor of success in the course (please see details in Appendix B in the online supplemental files). The average score on the BSMQ did not change much from the first study to the second. In the first study our students’ mean (standard deviation) score on the BSMQ was 12.5 (3.9) and in the second it was 12.9 (2.9). While this difference is statistically significant ($t = 1.96, p = 0.05$) it is not practically significant (note the Cohen’s $d$ statistic is 0.12 which also indicates a small effect size). Next we note that even after removing the 42 students (46 records) who did not complete the intervention, the percent of at-risk students was not significantly different for the two studies (24.4% in the first study and 20.9% in the second ($z = 1.649, p = 0.10$). However, there was an overall increase in the percent of students who were successful in the course. In the first study 269 of 500 (53.8%) of the students were successful and in the second study 1018 of 1508 (67.5%) were successful. This increase of 13.7 percentage points was significant ($z = 5.54, p = 0.001$).

Certainly there are several possible explanations for the overall increase we saw in our students’ success. Informed by our first study, we were more cognizant of our students’ mathematical level and adjusted our teaching and assessment methods to take this into account. Also, motivated by our new awareness of the professor effect, we have become more collaborative in sharing our course materials and ideas for teaching. In addition, we have improved the required General Education common core questions on our final exam to better reflect the material covered in the course. We believe all of these efforts could be contributing to the increased success of our students.

The more interesting question was what impact, if any, the early intervention had on the success of our at-risk students? First, we note that our second study alone cannot directly address this question as we did not conduct a full factorial randomized experiment using the binary independent variables whether the student was at-risk or not, and whether the student completed required tutoring or not, and the binary response whether the student was successful or not. However, we can compare our results from our second study to those of our first study. Based on our anecdotal observations, and supported by the literature cited in the Introduction, we had reason to believe that early tutoring would be helpful for our at-risk students. Thus we felt both practical and ethical constraints if we identified these students and did not provide them an opportunity to be more successful than expected.

In Table 4 we see a comparison of the two studies. While the overall increase in percent successful from the first to second study was 13.7 points, we found that the increase in percent successful for our at-risk students in the second study was 21.2 points compared with an increase in for non-at-risk students of 10.8 points. This was significant (Cochran-Mantel-Haenszel Test; Agresti 2002, p. 231ff).

The increase in the percent of successful at-risk students can also be seen in the graphs of the predictive logistic models for the probability of success in each study in Figure 2. Both curves show the BSMQ Score is a fairly good predictor of success, however there is a clear shift up in the lower half of the success curve for the second study. While we cannot establish causation, we do believe the required intervention had some impact on the improvement in success for our at-risk students. In Section 5 we discuss several reasons why we believe the intervention may have helped.

4.2. The Professor Effect

Once again we found a significant professor effect. First we note that we did not see any dramatic disparities, in terms of BSMQ and SATM scores, among the students by professor. In Figure 3 we see the distribution of students’ BSMQ and SATM scores by

---

**Table 2. Number of students taught by each professor for each semester.**

|                  | Fall ‘11 | Spring ‘12 | Fall ‘12 | Spring ‘13 | Fall ‘13 | Spring ‘14 | Total |
|------------------|---------|------------|---------|-----------|---------|------------|-------|
| Professor 1      | 68      | 67         | 98      | 0         | 68      | 56         | 357   |
| Professor 2      | 62      | 68         | 58      | 53        | 68      | 58         | 367   |
| Professor 3      | 88      | 0          | 28      | 0         | 95      | 54         | 265   |
| Professor 4      | 0       | 4          | 34      | 57        | 65      | 29         | 185   |
| Professor 5      | 0       | 35         | 67      | 57        | 0       | 0          | 159   |
| Professor 6      | 0       | 0          | 0       | 34        | 34      | 0          | 34    |
| Professor 7      | 58      | 0          | 33      | 0         | 50      | 0          | 141   |
| Total            | 276     | 170        | 318     | 167       | 380     | 197        | 1508  |

**Table 3. Quick comparison of first and second study.** In our 2011 JSE publication we included the Fall 2006 semester since we only reported results from the 15 point BSMQ.

|                  | First study | Second study |
|------------------|-------------|--------------|
|                  | 3 Semesters Spring 2007—Spring 2008 | 6 Semesters Fall 2011—Spring 2014 |
| 4 Professors     |                |              |
| 500 Student records |                |              |
| No intervention  |                |              |
| The same 20 point BSMQ given to all students in both studies. |

**Table 4. A comparison of the two studies.**

|                  | Students who scored 50% or lower on BSMQ (Required intervention in second study) | Students who scored above 50% (No required intervention in either study) |
|------------------|-----------------------------------------------------------------------------------|---------------------------------------------------------------------|
| Success Failure  | Total                                                                             | Success Failure Total                                               |
| First study      | 47 (38.5%) 75 122                                                                 | 222 (58.7%) 156 378                                                 |
| Second study     | 188 (59.6%) 127 315                                                              | 830 (69.6%) 363 1193                                                 |
| There was a 21.2 point increase in percent successful in this group. | There was a 10.8 point increase in percent successful in this group. |
professor. While the difference in mean BSMQ scores for the professors was statistically significant \([F(6,1501) = 4.600, p = 0.00012]\) the distributions of the scores do not appear to be practically different. We note the means of the SATM scores were not statistically different \([F(6,1360) = 1.713, p = 0.114]\). Based on these results we argue that no professor was particularly advantaged or disadvantaged in terms of their students’ mathematical ability as measured by these two instruments.

However, there was a noticeable difference in the probability of success curves of the professors. In Figure 4 we see the overall probability of success curve and the curves delineated by professors. Note that among students who scored 50% on the BSMQ, the students who had Professor 5 had about a 43% chance, on average, of being successful in the class whereas those who had Professor 6 had about a 77% chance, on average. Also note that the success curves for Professors 2 and 7 are virtually indistinguishable.

For the three professors who were included in both the first and second study we saw a definite reduction in the professor effect. In Figure 5 we see the probability of success curves from both studies for these three professors. What is noteworthy is not only how the lower end of the curves has shifted up, but also how those curves have moved closer together. In fact, the model shows no significant difference in the coefficients for these three professors (see Appendix B, Model 4). In Section 5.1 we further discuss the professor effect.

5. Discussion, Recommendations, and Future Work

This study reinforces a result from previous studies (Johnson and Kuennan 2006; Lunsford and Poplin 2011). Namely, even in an algebraically and computationally light introductory statistics class, students with better basic mathematics skills, as measured by the BSMQ, are more likely to be successful. It also adds to the body of literature regarding the benefits of peer tutoring in college statistics classes. While many research studies find advantages to using peer tutoring as a supplement to instruction, very few studies describe a successful model in introductory statistics classes (see Magel 1998, Keeler et al. 1995, and Garfield and Ben-Zvi 2007 for exceptions). We were happy to see an overall increase in student success in our new study and, as discussed above, we believe many of the changes
we implemented, informed by the first study, contributed to that. However, we think the statistically larger increase in success from the first to second study among our at-risk students (who completed the required intervention) versus our other students is the more interesting result. Clearly we cannot establish causation in this study but we do have some ideas as to why we saw this result.

5.1. Discussion of Study Results

First, we should mention that in a study of this size, involving a large number of professors and students, there are bound to be students and possibly professors who do not comply with the departmental policy. Indeed, we had one professor who did not properly implement the required tutoring so was not included in this study. We found 46 records (42 students with four repeats) that were categorized as at-risk but did not complete the required tutoring. Of these records, we had 32 that received a grade of F(14) or W(18), but 14 received a grade of in the course. So, even among the professors who did enforce the departmental policy, there were students that slipped through the cracks.

Second, the CAS tutoring for statistics is not managed by our department. And while the CAS was cooperative, we did have some logistical problems. Because there was no procedure in place which required students to sign up for when they would be going to tutoring and also to obtain their tutoring hours in an evenly spaced manner, the walk-in model led to excessively large tutoring sections in the last 2-week period before the deadline for completion of the 6 h. This was a frustrating experience for both the students and the tutors. Also, while all professors now use the TI-84 series calculator for the class, the professors do not follow the same order of topics or cover the topics at the same rate. This made the job more difficult for the tutors.

Given the issues we had with the required tutoring, we are not sure about the quality of the tutoring our students received or how much it contributed to their success. It is possible a reason for the increase in success for our at-risk students may be that the required tutoring forced them to be engaged, even minimally, with the material early in the semester. Certainly, a limitation of our study was that the tutoring used was not assessed. Currently there is no effective way for us to measure the level of tutoring received or to distinguish the students who actually used the tutoring from those who were minimally engaged. Some qualitative research on peer tutoring suggests that the simple act of going to tutoring spaces creates a unique experience for undergraduates. This experience creates a culture where tutees engage with tutors in a way that encourages studying which may not otherwise occur. The different form of interaction and the programmatic structures in which it occurs may influence the outcomes tutees achieve (Breslin 2014) Also, the early material in the class (descriptive statistics) is generally easier than the later material (inferential statistics). If at-risk students can do well enough on the early material, we believe it gives them an advantage in the second half of the class (to either better learn the later material or have a high enough grade by midterm to have a better chance of passing the entire class).

We also found it interesting that the professor effect was greatly reduced among the three professors who were in both studies. We essentially chalk this up to “experience” and awareness of the effect. We have continued to share materials among all professors teaching the course, including slide presentations, review materials, and tests (both old and current). Something we implemented after our first study was improving the
required LU General Education common final exam questions for MATH 171. These eight multiple questions cover basic topics in the course. For the 673 students in our study for which we had the data, we found that the mean of the number of correct answers to these eight questions was significantly higher for the successful students than for the students who were not successful (see Figure 6). The graph shows the third quartile of the number of correct answers for the unsuccessful students is the same as the first quartile of the successful students. Thus, these eight questions give a coarse indicator that students who are successful in our course are learning more statistical concepts than those who are not.

5.2. Recommendations and Future Work

This study highlighted three issues we have faced in trying to improve student success rates in our introductory statistics class via an early peer tutoring intervention. First is finding a way to identify our at-risk students. Second is providing a quality intervention that helps those students to succeed in the course by engaging them with the material. Lastly, we feel the professor effect is fundamentally unfair and we should work together to reduce it. Below we elaborate on each of these.

Some of the issues we have with administering the BSMQ on the first day of class are many of the professors feel it sets the wrong tone and, even though it is done via Scantron, there are some logistical challenges involved in getting it graded with results back to the students and professors in a timely manner. Also, as assessment has become more important for other university and state-required initiatives, we feel there is an “assessment burden” on our students. Since the completion of our second study, we have been experimenting with other assessments the university is using, such as the QLRA (Gaze et al. 2012), and which we can administer on-line during the summer at new student orientation. Ideally we hope we can find an assessment tool, or perhaps combine two or more tools into one, that will not only be easy to administer and meet other assessment needs but also enable us to identify our at-risk students in MATH 171.

We note that in our new study, unlike our first study, we collected SAT scores and found that SAT Math (SATM) score was a significant predictor of success. For the BSMQ a one point increase yields a 13% increase in odds of success and a 10 point increase in SATM score yields a 6.3% increase in odds of success. Unfortunately, as mentioned above, it is difficult to get SATM scores for the large number of students enrolled in MATH 171. Also, in order to use the results from our first study and to compare results in our new study to the first study, we needed to use the BSMQ scores. However, in the future, we hope to work with our administration to find a way to more easily obtain SATM scores. For the students for which we have data, their BSMQ and SATM scores are positively correlated. The graph shown in Figure 7 is a scatterplot of student SATM and BSMQ scores color coded by whether the students were successful in the course or not with a vertical line drawn at the first quartile of the SATM scores. Since the BSMQ cutoff score

![Figure 6. Boxplots of the number of common final exam questions correct by student success.](image6.png)

![Figure 7. BSMQ scores versus SATM scores for the new study.](image7.png)
of less than 10 essentially gives us about 20% of our students enrolled in MATH 171, we would initially use the first quartile of the SATM scores to identify at-risk students. Note that while there is some overlap in the students categorized as at-risk by the BSMQ and SATM, there would be a different set of at-risk students if we used SATM scores instead. Also note our at-risk students, as determined by the BSMQ, did receive tutoring whereas this was not the case with all at-risk students determined by the SATM scores. Including SATM scores or some combination of assessments, such as the BSMQ and SATM, to identify at-risk students is clearly a route for further study.

Based on our observations and student feedback, we also believe our university tutoring services could definitely be improved. This possibly includes better logistical planning by having students sign up on-line for tutoring times and having more flexibility in assigning the number of tutors per session and the times of and number of tutoring sessions, especially during peak times in the semester. In the Fall of 2014 (after our second study was completed and as part of a student’s senior project), a survey was administered to MATH 171 students at the end of the semester to gauge their attitudes about the required tutoring. The preliminary results from this survey echoed the need for better logistical planning while simultaneously noting that the tutors were “helpful.” In future studies we will include evaluation of student attitudes regarding any interventions we may try. In addition to improving logistics, we also wonder if specific training in tutoring for both remedial mathematics and statistics as well as better pay for the tutors would be helpful in improving tutoring quality. We would like to see our tutoring program have a professional training component that would be résumé worthy for our tutors. Finally, we believe if professors tried to synchronize the order of and length of time to cover the topics in the class, this would improve the tutoring experience for both students and tutors. Before requiring a tutoring intervention in any future studies, we would like to see resources allocated so that these, or other, best practice procedures can be implemented to improve tutoring quality. This includes assessing the tutoring program and any efforts made to improve it. In particular, we would like to find a way to measure the quality of the tutoring received and gauge the level of a student’s participation in the required tutoring.

Lastly, to help to alleviate the professor effect, we have discussed implementing a departmental common (or 80% common) final exam for MATH 171. We can certainly understand why some of our colleagues bristle at this idea with appeals to their academic freedom. However, we argue that by having a common exam, and by grading it together (using a model somewhat like the AP Statistics reading), we could provide in-house professional development for and improve the communication among the mathematics professors at our university, both full time and adjunct, who teach introductory statistics.

Supplemental Materials

The appendices are available online as supplementary material at the publisher’s website.

Acknowledgments

We would also like to thank the editors and anonymous referees whose comments and suggestions greatly improved this article. We would also like to thank our colleagues for agreeing to administer and count the basic skills test in their introductory statistics classes, the LU CAS administrators and tutors for working with our department to implement the required tutoring, the Longwood Assessment Minigrant Program (LAMP) for partially supporting this research, our student Sarah Kessler for her work on student attitudes about the required tutoring, and most importantly, our MATH 171 students for being patient with us as we work to find ways to improve their outcomes.

References

Agresti, A. (2002), Categorical Data Analysis (2nd ed.), Hoboken, NJ: Wiley-Interscience.
Breslin, J. D. (2014), “Social Learning in the Co-curriculum: Exploring Group Peer tutoring in College,” Doctoral Dissertation from University of Kentucky. Available online: http://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1024&context=sepe_etds
Bude, L. et al. (2008), “The Effect of Directive Tutor Guidance in Problem-based Learning of Statistics on Students’ Perceptions and Achievement,” Higher Education, 57, 23–36. doi:10.1007/s10734-008-9130-8
Cohen, P. A., Kulik, J. A., and Kulik, C. C. (1982), “Educational Outcomes of Tutoring: A Meta-analysis of Findings,” American Educational Research Journal, 19, 237–248.
Conners, F. A., Mccown, S. M., and Roskos-Ewoldson, B. (1998), “Unique Challenges in Teaching Undergraduate Statistics,” Teaching of Psychology 25, 40–42.
Dancer, D., Morrison, K., and Tarr, G. (2015), “Measuring the Effects of Peer Learning on Students’ Academic Achievement in First-year Business Statistics.” Studies in Higher Education, 40, 1808–1828. Available online: http://srhe.tandfonline.com/doi/abs/10.1080/03075079.2014.916671
Garcia, R., Morales, J. C., and Rivera, G. (2014), “The Use of Peer Tutoring to Improve the Passing rates in Mathematics Placement Exams of Engineering students: A Success Story,” American Journal of Engineering Education, 5, 61–72. doi:10.19030/ajee.v5i2.8952
Garfield, J. and Ben-Zvi, D. (2007), “How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics,” International Statistical Review, 75, 372–396. doi:10.1111/j.1751-5823.2007.00029.x
Gaze E. et al. (2012), Quantitative Learning and Reasoning Assessment Project. Available at http://scre.carleton.edu/qral/index.html
Giraud, G. (1997), “Cooperative Learning and Statistics Instruction,” The Journal of Statistics Education, 5.
Johnson, M. and Kuennen, E. (2006), “Basic Math Skills and Performance in an Introductory Statistics Course,” Journal of Statistics Education [on-line], 142, 1–16.
Keeler, C.M. & Steinhorst, R.K. (1995), “Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A Report from the Field,” J. Stat. Educ., 3. doi:10.1080/10691898.1995.11910485
Leppink, J. et al. (2013), “The Effect of Guidance in Problem-Based Learning of Statistics,” The Journal of Experimental Education, 82, 391–407. Available at http://www.tandfonline.com/doi/full/10.1080/00220973.2013.813365
Leung, K. C. (2015), “Preliminary Empirical Model of Crucial Determinants of Best Practice for Peer Tutoring on Academic Achievement,” Journal of Educational Psychology, 107, 558–579.
Lunsford, M. L. and Pendergrass, M. H (2016), “Making Online Homework Work,” PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, Special Issue: Special Issue on Teaching with Technology, 26, 531–544.
Lunsford, M. L. and Poplin, P. (2011), “From Research to Practice: Basic Mathematics Skills and Success in Introductory Statistics,” Journal of Statistics Education [Online], 19, 1–22.
Magel, R. C. (1998), “Using Cooperative Learning in a Large Introductory Statistics Class,” The Journal of Statistics Education, 6.
Roberts, C. H., and Baugher, G. A. (2015), “Academic Enhancement and Retention Efforts in Mathematics Designed to Support and Promote Enhanced STEM Education for Adult Learners at Mercer University,” Interdisciplinary STEM Teaching & Learning Conference. 5. Available at http://digitalcommons.georgiasouthern.edu/stem/2015/2015/5
Sharpley, A. M., and Sharpley, C. F. (1981), “Peer Tutoring: A Review of the Literature,” *Collected Original Resources in Education*, 5, 7–C11
Suanpang, P., Petocz, P., and Kalceff, W. (2004), “Student Attitudes to Learning Business Statistics: Comparison of Online and Traditional Methods,” *Educational Technology & Society*, 7, 9–20.

Topping, K. J. (1996), “The Effectiveness of Peer Tutoring in Further and Higher Education: A Typology and Review of the Literature,” *Higher Education*, 32, 321–345. Available at [http://www.fau.edu/CLASS/CRLA/Level_Three/The_effectiveness_of_peer_tutoring_in_further_and_higher_education-a_typology_and_review_of_the_literature.pdf](http://www.fau.edu/CLASS/CRLA/Level_Three/The_effectiveness_of_peer_tutoring_in_further_and_higher_education-a_typology_and_review_of_the_literature.pdf)