THE IMPACT OF A STATISTICAL COLLABORATION LABORATORY ON THE STATISTICS STUDENTS WORKING IN IT

ABSTRACT

Graduate level statistics education curricula often emphasize technical instruction in theory and methodology, but can fail to provide adequate practical training in applications and collaboration skills. We argue that a statistical collaboration center ("stat lab") structured in the style of the University of Colorado Boulder's Laboratory for Interdisciplinary Statistical Analysis (LISA) is an effective mechanism for providing graduate students with necessary training in technical, nontechnical, and job-related skills. We summarize the operating structure of LISA, and then provide evidence of its positive impact on students via analyses of a survey completed by 123 collaborators who worked in LISA between 2008–15 while it was housed at Virginia Tech. Students described their work in LISA as having had a positive impact on acquiring technical (94%) and non-technical (95%) statistics skills. Five-sixths (83%) of the students reported that these skills will or have helped them advance in their careers. We call for the integration of stat labs into statistics and data science programs as part of a comprehensive and modern statistics education, and for further research on students' experience in these labs and their impacts on student outcomes.

Keywords: graduate curriculum; collaboration; consulting; communication; technical skills; nontechnical skills; career skills

1. INTRODUCTION

1.1. HISTORICAL BACKGROUND

The call to educate and train statistics students in effective collaboration skills dates back to the early 20th century and the post-war increased interest in the science of statistics. Hotelling (1988) noted that statistics education began to increase after World War I and reached "large proportions" by 1949, while noting discrepancies among departments regarding the presence and quality of statistical laboratories. He claimed that "the teaching of statistics should be accompanied by considerable work in applied statistical problems… best conducted in a laboratory," and noted that research work should span many fields. Kimball (1957) recommended that statistics graduate students attend consulting meetings with faculty members, engaging in active participation and discussion in the meeting to avoid giving the right answer to the wrong problem, which he defined as an "error of the third kind." Writing about the future of data analysis, Tukey (1962) noted that "all sciences must teach their apprentices how to think … and what are its current beliefs and practices." He warned that compared to scientists in other disciplines, who often have practical experience in laboratory and field settings, data analysts and statisticians may lack similar practical experience in a professional context.

More recently, Jeske et al. (2007) emphasized students' need for a balance of technical and nontechnical skills to ensure that statistically sound methodology is used in appropriate scientific and research circumstances. The American Statistical Association (ASA; 2014) explicitly called for the incorporation of collaboration skills into a statistics curriculum, highlighting the need for skills enabling communication and collaboration with clients and collaborators from a variety of fields. An ASA panel discussion on "Challenges and Opportunities for Statistics in the Next 25 Years" called for collaboration with computer science and other disciplines and improved teamwork (Kettenring et al., 2015). The National Association of Colleges and Employers (NACE) 2015 Job Outlook Survey found that the ability to work as a team and written communication were more sought-after on a candidate's resume than technical and problem-solving skills (National Association of Colleges and Employers, 2014). Börner et al. (2018) noted that data science and data engineering work requires strong interpersonal communication, writing, and collaboration.

These necessary communication and collaboration skills can be augmented by students' experience working in a statistical advising, consulting, or collaboration laboratory or center, which we call a "stat lab". Stat labs can take many forms (LeBlanc et al., 2022) and have the potential to transform evidence (data) into action for the benefit of society (Olubusoye et al., 2021; Vance & Love, 2021). Vance and Pruitt (2022) wrote that when a stat lab's administration, personnel, services, communication strategy, and budget are aligned with a mission to train students and support data-driven decision-making in the community, a stat lab can achieve excellent outcomes and impacts.

Although many have called for these skills to be explicitly included in a statistics curriculum, Blei and Smyth (2017) noted that students do not typically practice these skills through direct classroom experience, but through experience and collaboration with others. Vance and Smith (2019) developed their ASCCR framework to teach five essential components of effective collaboration (Attitude, Structure, Content, Communication, and Relationship) explicitly in the classroom. This framework incorporates Zahn's POWER process (2019) to structure meetings among clients, consultants, and collaborators; how to ask great questions (Vance et al., 2022); and how to create shared understanding (Vance et al., 2022), among other collaboration skills that can be taught and learned through direct classroom experience. Sharp et. al. (2021) produced ten videos to help students learn statistical collaboration skills through observation and discussion.

1.2. PURPOSE OF THIS WORK

The purpose of this work is to provide quantitative and qualitative evidence that experience working in a stat lab answers the literature's persistent call to prepare statisticians better for rigorous and impactful collaboration with professionals in other fields. We address this goal via the results of a survey instrument completed in 2015 by 123 (mostly) graduate students who worked in the Laboratory for Interdisciplinary Statistical Analysis (LISA), housed at the time of the study at Virginia Polytechnic Institute and State University (Virginia Tech) in Blacksburg, Virginia. Specifically, we report on the positive impacts this work has on several components of students' experiences:

1) Technical skills, such as statistical theory, methods, applications, and computation;
2) Nontechnical skills, such as communication, collaboration, structuring meetings, and explaining statistics to non-statisticians;
3) Professional advancement, such as finding and succeeding in meaningful work;
4) Professional curiosity and fulfillment, such as an increased desire to apply statistics, solve problems, answer questions in other fields, and a greater sense of confidence in making a positive impact on one's workplace.

Additionally, we analyzed the factors that moderate the impacts that students report, such as the degree to which they engaged on collaborative projects and the leadership and supplementary roles they served.

By quantifying the direct impacts work in a statistical collaboration lab has on student outcomes, we also expand the existing literature that has mainly focused on broader impacts that research has on science, culture, and society. For instance, Penfield et. al. (2014) provided a review of how university research affects external bodies and how these impacts are quantified. Morton (2005) provided evidence that statistical collaboration can impact policy decisions, and Lee et. al. (2005) found that as the number of scientific collaborators increased, the resultant research became more widely cited. Aldieri et. al. (2018) found positive impacts from scientific collaboration on academic performance at the university scale. While it is clearly important to quantify and describe these impacts on external stakeholders, these works did not describe or quantify the impacts this research has on the researchers themselves. In our case, we will propose novel metrics to measure these impacts directly on the graduate student collaborators involved in the research projects, and quantify the impacts.

We refer to the graduate student study participants as collaborators, in keeping with the recommendations of Love et al. (2017) and Vance (2015) to encourage more scientifically, statistically, and professionally cooperative relationships between collaborators and their clients.

Similarly, we will subsequently refer to clients as domain experts to further emphasize that this working relationship should not have a hierarchy of importance or role.

The remainder of this paper is organized as follows. In Section 2, we describe the experiences of our graduate student collaborator respondents involved in LISA and its training program at Virginia Tech. In Section 3, we describe the survey instrument used to assess students' perceptions of their experiences in LISA and describe our study design and methods. In Section 4 we present our quantitative results on these impacts. In Section 5 we present qualitative results and discuss the potential impact of all results on the statistics and data science education community. We conclude in Section 6 by calling for more universities to support robust training programs for statistics and data science students to gain practical experience applying statistics and data science to solve real problems.

2. LISA AND ITS TRAINING PROGRAM

LISA was created in 2008 and was housed within the Virginia Tech Department of Statistics until 2016 (Vance & Pruitt, 2016) when it moved to the Department of Applied Mathematics at the University of Colorado Boulder. The stat lab remaining at Virginia Tech transitioned into the new Statistical Applications & Innovations Group (SAIG) in 2017. LISA was built on the foundations of the Statistical Laboratory (1948–1972) and the Statistical Consulting Center (1973–2007; Arnold J. C., 2000; Arnold et al., 2013). At Virginia Tech, LISA's mission was to train statisticians to become interdisciplinary collaborators, provide research infrastructure to enable and accelerate high impact research, and engage with the community in outreach activities to improve statistical skills and literacy.

Participants in our study worked as statistical collaborators in LISA from 2008–2015 to serve one or more purposes: as part of their funded assistantship duties, to satisfy the M.S. degree requirement to work in LISA for one semester, as post-doc or pre-doc scholars, or as volunteers to gain practical experience. They all, to varying degrees, partook in LISA's five-part education and training program consisting of preparation, practice, doing, reflecting, and mentoring (see Vance et. al., 2020 and LeBlanc et. al., 2022). Collaborators prepared to work in LISA by taking technical statistics courses and the nontechnical "Communication in Statistical Collaborations" course—typically during their second semester in the statistics M.S. or Ph.D. program—starting in 2009. They practiced collaboration skills in that course and opportunistically during weekly staff meetings and Video Coaching and Feedback Sessions (VCFS). They engaged in LISA collaboration projects and then reflected on their experience during VCFS. Collaborators completed the LISA training program by mentoring novice collaborators. Some contributed to LISA in additional ways, such as by teaching Short Courses, serving as a Lead Collaborator and/or Pod Leader, or hosting Walk-in Consulting hours. To provide additional context for our study, we detail these program components below.

2.1. LISA PROGRAM COMPONENTS

Foundational Technical Education

Graduate M.S. and Ph.D. students completed 20 credits of technical courses during their first year, including a two-credit computing course; a three-credit design of experiments course; a three-credit probability course; a two-course, six-credit sequence on statistical inference; and a two-course, six-credit sequence on regression and linear models. These courses provided LISA collaborators an essential foundation in the theory and methods of statistics.

Communication in Statistical Collaborations Course

Students typically began their LISA experience by enrolling in a three-credit, one-semester formal collaboration course. Course content included the following topics:

- Adopting an attitude of collaboration and managing effective statistical collaboration meetings with domain experts.
- Writing about, presenting, and verbally communicating statistical concepts, analyses, and results to non-statistical audiences.

- Using peer feedback, self-reflection, and video analysis as a process for improving communication and collaboration skills.
- Effectively collaborating with group members; creating reproducible statistical workflows.
- Data ethics and ethical collaborations.

From 2012–2016, the course employed Team-Based Learning (Vance, 2021) to facilitate rigorous discussion and practice of the nontechnical skills needed for effective collaboration. Teams were intentionally formed with native and non-native English speakers grouped heterogeneously to encourage diversity in perspectives and communication styles. Students concluded the semester by working in pairs on a real project with a domain expert of their own choosing.

Weekly Staff Meetings and the Pod System

After students completed the collaboration course, they typically joined LISA for at least one semester. Weekly staff meetings provided opportunities for collaborators to discuss their and their domain experts' needs with the larger group of faculty and other collaborators to solicit feedback, advice, and mentoring on both statistical methodology and collaboration techniques. In 2013, LISA implemented a pod system to organize its Associate (novice) Collaborators and Lead Collaborators. Experienced Lead Collaborators became Pod Leaders and were responsible for supervising and mentoring four Associate Collaborators each semester. Pods met weekly to discuss projects amongst themselves or with LISA staff to engage in VCFS.

Collaboration Projects

LISA's primary service was to provide consultation to and collaboration with domain experts from across the university. When domain experts submitted project requests, they were assigned to work with a LISA collaboration team—typically of size two—including a Lead Collaborator and an Associate Collaborator. The Lead Collaborator was responsible for leading the meetings and project, communicating with the domain expert, and using the project to mentor the Associate Collaborator. Associate Collaborators were responsible for summarizing the domain expert's project goals, helping the Lead conduct effective meetings, and contributing to the analyses for the project. Figure 1 shows the distributions of total numbers of projects per student, walk-in semesters, and short courses for survey respondents. Among respondents in our study, the mean number of lifetime collaboration projects was 20.9; the median was 10; and the standard deviation was 28.4. The distributions are right-skewed due to a small proportion of collaborators having high engagement with LISA over multiple semesters. Among non-respondents, the mean was 10.4; the median was 6; and the standard deviation was 14.0. 32% of respondents served as a Lead Collaborator at least once; 12% of non-respondents did so.

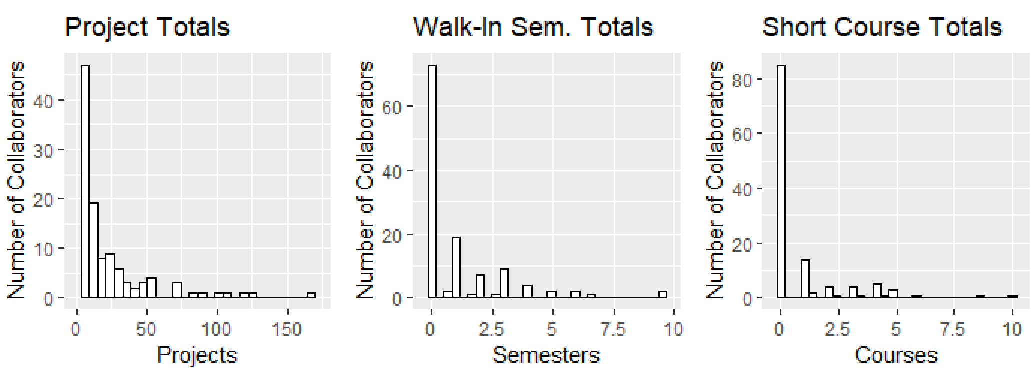

Figure 1: distribution of assigned projects, walk-in service semesters hosted, and short courses taught by each respondent.

Walk-in Consulting

LISA also provided Walk-in Consulting for relatively quick and straightforward statistical questions at pre-scheduled times and locations throughout a semester. For collaborators who were assigned to Walk-in Consulting, their hourly commitment was typically for two hours once per week. With 5-10 sessions hosted weekly at various locations on campus, this usually amounted to 10-20 person-hours per week. 41% of respondents conducted at least one semester of Walk-in Consulting. Among those, the mean number of semesters served was 2.57; the median was 2; the standard deviation was 2.10; and the 75th percentile was 3.

Short Courses

LISA offered short courses, which were typically condensed, two-hour tutorial courses intended for novice researchers and non-experts, on a variety of topics. In addition to the benefits these short courses provided the attendees, LISA collaborators also gained the experience of designing curriculum, teaching the course, acting as teaching assistants, and providing administrative support such as classroom and technology setup. Collaborators designed and taught short courses on a volunteer basis or as part of a formal paid assistantship with LISA. For data coding purposes in our study, students who co-taught a short course were credited with teaching 0.5 courses. 31% of study respondents taught at least one short course. Among those, the mean number of short courses taught was 2.82; the median was 2; the standard deviation was 2.15; and the 75th percentile was 4.

Video Coaching and Feedback Sessions (VCFS)

A Video Coaching and Feedback Session involved peer analysis of a recorded collaboration meeting involving a Lead Collaborator, an Associate Collaborator, and one or more domain experts. The collaborators typically choose 2–4 short selections (1–5 minutes) of the recording for a small group of peers and LISA staff to watch and provide feedback on what the collaborators did well and what opportunities there were for improvement in leading the meeting and addressing the domain expert's wants. All participants were expected to reflect on what they learned from the VCFS to improve their collaboration skills and state their top lesson learned. From 2010–2016 LISA conducted VCFS approximately once per week for a total of 259 sessions. While not part of a formal statistical mentoring program such as those described in Vance et al. (2017a) and Vance et al. (2017b), in our experience, the VCFS presented invaluable opportunities for practicing collaboration skills through on-the-spot role plays and just-in-time mentoring of young statisticians as described in Anderson-Cook et al. (2017).

3. METHODS

3.1. PARTICIPANTS AND RESPONSE RATE

All 173 statistical collaborators who had worked on at least one collaboration project in LISA between August 2008 and December 2014 were invited to participate in our study of the self-perceived impacts of working in a statistical collaboration laboratory. Of these 173 collaborators, 8 were current M.S. students, 34 were current Ph.D. students, 82 were graduated M.S. recipients, and 42 were graduated Ph.D. recipients. Smaller populations included one current undergraduate student, three graduated undergraduate students, two visiting scholars from the global LISA 2020 program in developing countries (Vance E. A. et al., 2022) who were current Ph.D. students at other universities, and one LISA 2020 post-doc. For the purposes of this study, undergraduate students were categorized as M.S. level collaborators and visiting scholars were categorized as Ph.D. level collaborators.

The survey collected data on the students' training and amount of experience in LISA and asked five-point Likert scale questions about the collaborators' attitudes toward statistics and application areas, self-efficacy in applying statistical thinking to help solve real-world problems, satisfaction, and the impact of LISA on their technical and nontechnical skills. The survey also included a few open-ended questions about the collaborators' qualitative perception of the impact working in LISA had on them.

This study received initial approval from the Virginia Tech Institutional Review Board (IRB), and—after all data had been collected and deidentified—was designated by the University of Colorado Boulder IRB as "not human subjects research." Informed consent was required for participants to opt in to the study and access the survey. Of the 173 people invited to participate, 123 (71%) returned the survey (120 answered every quantitative question); 37 (21%) did not respond; 2 (1%) opened the survey but did not respond; 2 (1%) opted out of the survey; and 9 (5%) of the messages sent were undeliverable due to email failure. Since collaborators were contacted through their last known email address, it is unclear how many of the 37 who did not respond were actually reached.

The sample of respondents is summarized in Table 1. Respondents included all 9 current M.S. students and 33 of 36 Ph.D. students at the time of the survey, as well as 46 graduated M.S. collaborators and 35 graduated Ph.D. collaborators for a total of 123 respondents. 54 were native English speakers and 66 were non-native English speakers. Of the respondents, 120 completed the entire survey. One respondent completed the survey up to the section on nontechnical skills.

Table 1: Summary of population and respondents

	Population	Respondents (%)	Non-Respondents
Current M.S.	9	9 (100%)	0
Current Ph.D.	36	33 (92%)	3
Graduated M.S.	85	46 (54%)	39
Graduated Ph.D.	43	35 (81%)	8
Total	173	123 (71%)	50

Collaborator experience and demographic data were gathered from LISA records to augment the survey data. The data from administrative records included the subjects' current/graduated and M.S./Ph.D. designation, the number of projects they worked on, whether they had served as a Lead Collaborator, the number of short courses they had taught, the number of semesters in which they served as a Walk-in Consultant, what year they took the Communications in Statistical Collaborations course, and what year they first worked with LISA. All personally identifying information such as names, email addresses, etc. were permanently deleted from the final dataset and the key linking the survey data to LISA administrative records was maintained and then destroyed by the fourth author.

During this time, the 173 statistical collaborators served on 3,088 LISA collaboration project teams. The experiences of respondents are summarized in Table 2. Respondents accounted for 2,566 (83%) of these projects. Forty-five of the individuals had served as Lead Collaborators with 39 (87%) of these responding to the survey. Of the 144 Walk-in person-semesters and 107.5 short courses taught, respondents accounted for 128.5 (89%) and 107 (99.5%) respectively. The response rate was highest for collaborators who had taken the Communications in Statistical Collaborations course in 2014 or 2015 (100%) and lowest for those who took the course in 2009 (50%) or who had not taken the course (60%). Similarly, the response rate was highest for those who had first started working with LISA in the 2014-15 (100%) and 2013-14 (90%) cohorts and lowest for those from the 2007-08 (44%) or 2009-10 (67%) cohorts.

Table 2: Response rate compared to collaborator experience.

	Respondents	Non-Respondents	Total
Projects	2,566 (83%)	522	3,088
Lead Collaborators	39 (87%)	6	45
Walk-In Semesters	128.5 (89%)	15.5	144
Short Courses Taught	107 (99.5%)	0.5	107.5

3.2. SURVEY INSTRUMENT

Survey

We developed an online survey using Qualtrics that asked questions about the impact of involvement in LISA on the respondent's technical and nontechnical skills, experiences in the workforce, and other perceived impacts of working in LISA. The survey included ten multiple-choice, 25 five-point Likert scale, and three open-ended items. One additional Likert scale item was asked of graduated Ph.D. students about LISA's impact on their dissertation topics. Current M.S. students did not receive the five Likert scale items about the impact of LISA on their experience in the workforce. The survey was disseminated to all former and current LISA collaborators in late January 2015 and remained available through March 2015. The administered survey for graduated Ph.D. students can be found in Appendix A.

Procedures

Using the Qualtrics survey mailer feature, recruitment emails with survey links were sent to all 173 collaborators. The recruitment email explained the purpose of the study and how results would be used. After a week or so, non-responders received a follow-up reminder email that included the current response rate and median response time for each group. In total, six reminder emails were sent to encourage an increased response rate. Each of the collaborators was assigned a unique numerical code used to link the collaborator's experience in LISA and demographic information with their survey responses. The linking codes were then permanently deleted.

Validity and Reliability

To find evidence of content validity (Lawshe, 1975), the draft survey was reviewed by several directors of other stat labs and improved based on their feedback. We piloted the survey in early spring semester 2015 with the newest cohort of LISA collaborators. Feedback from this cohort was used to make minor refinements to the survey; no pilot data was included in the analyzed dataset. Evidence of strong reliability was provided by very high Cronbach's alpha values (Taber, 2018; for details see Section 4).

Qualitative Data Collection

The survey instrument also included three qualitative free-response items: How has your involvement in LISA (positively or negatively) impacted you? Please consider ALL the impacts (academic, social, professional, etc.); "Do you have any suggestions for improving LISA?"; and "Any other comments?" All 146 qualitative comments from 78 respondents were read and categorized initially into four categories: technical, nontechnical, career, and general comments. From a subsequent analysis, an additional five topics emerged: "students appreciated the practical, real-life application of statistics in their LISA experience," "LISA positively impacted life in graduate school and beyond," "I only fully recognized the importance of LISA after I graduated," "I wish I had worked on more projects," and "the LISA opportunity should be expanded to others." We provide some of the relevant comments, both positive and negative, in Section 4.2.

4. RESULTS

4.1. ANALYSIS OF QUANTITATIVE RESULTS

The broad categories of technical impacts, nontechnical impacts, and job impacts were each assessed with a binary response item and a series of five-point Likert scale items. First, we provide the binary results. 94% of respondents (116 of 123) agreed that their involvement in LISA helped them gain technical skills in statistics, including understanding statistical theory and methods, applying statistics, manipulating data sets, and computational skills. 95% of respondents (117 of 123) agreed that their involvement in LISA helped them gain nontechnical skills essential for success in statistics, including communication, collaboration, structuring meetings, and explaining statistics. 83% of respondents (100 of 120) agreed that their involvement in LISA will help or has helped them advance in their professional career, including getting an interview, getting a job, and getting promoted.

Proportions of Likert responses are provided in Figure 2. They show a heavy skew toward Agree and Strongly Agree for most items. Note that some career-related questions had fewer responses, as not all students had entered the workforce by the time they completed the survey. The mean and 95% confidence interval for each individual Likert question are shown in Figure 3. Taken individually, each response's mean demonstrates a clearly positive perceived impact. The 95% confidence intervals all exceed the baseline of 3, and indeed many exceed 4, indicating positive and highly positive mean perceived impacts. Tables B1-3 in the appendix provide the number of responses, means, and standard deviations for each individual question.

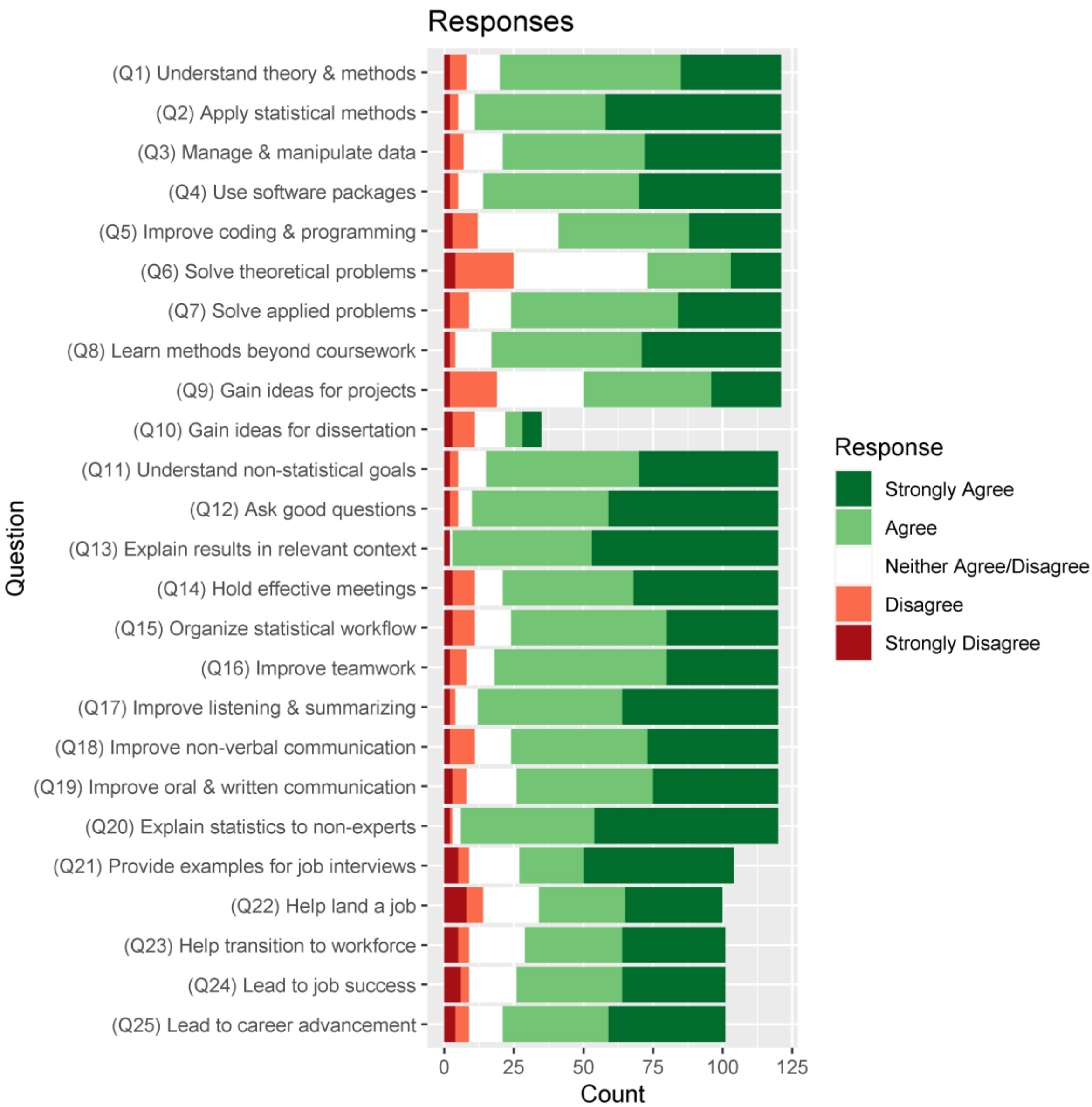

Figure 2: Proportions of Likert responses to items on Technical, Non-technical, and Job-related skills and impacts. The wording of the survey questions can be found in Appendix A.

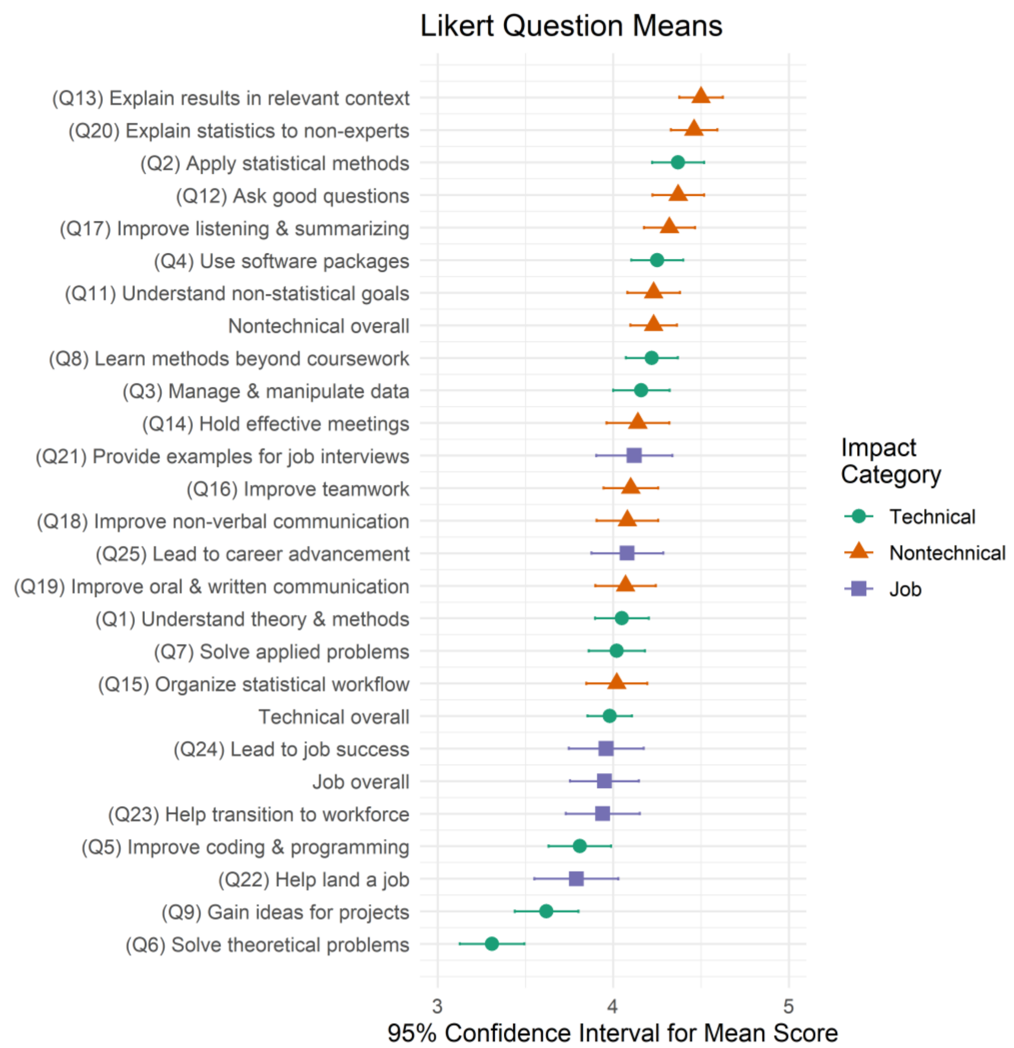

Figure 3: Means and 95% confidence intervals for Likert item responses and impact category scores. The wording of the survey questions can be found in Appendix A.

Technical, Nontechnical, and Job-Related Impacts

We calculated a perceived impact score for the categories of technical, nontechnical, and job-related impacts by averaging each collaborator's corresponding Likert scale items. The questions included in each category are color-coded in Figure 3 and provided in Tables B1-3. Job-related questions were only asked to students who joined the workforce after their involvement in LISA. The average technical impact score (reflecting 9 items) was 3.98 (SD = 0.71); the average nontechnical impact score (reflecting 10 items) was 4.23 (SD = 0.74); the average job-related impact score (reflecting 5 items) was 3.95 (SD = 1.04). The distributions of these impact scores are shown in Figure 4.

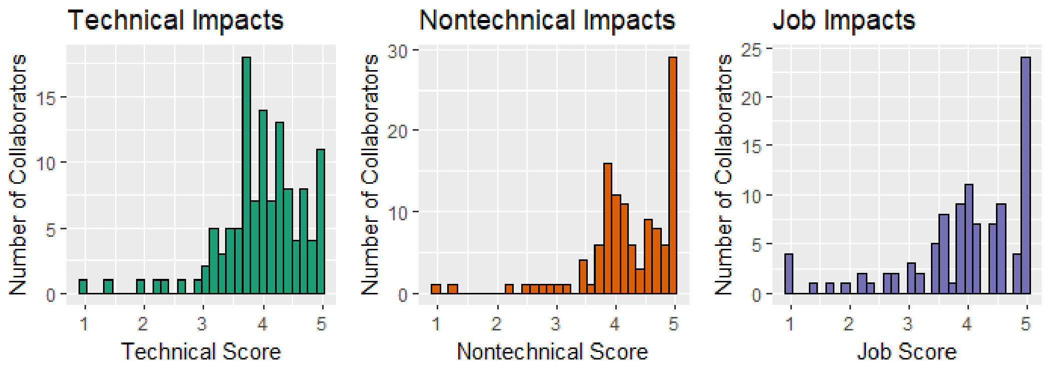

Figure 4: distributions of technical, nontechnical, and job-related impact scores by collaborator.

Reliability of Perceived Impact Scores

To test how related individual items were within each perceived impact category, we calculated Cronbach's alpha for each set of items and simulated 95% bootstrap confidence intervals for these values. For the category of "technical skills", questions 1-9 (excluding Question 10 due to its limited scope) had a Cronbach's alpha value of 0.919, with a 95% bootstrap confidence interval of (0.872, 0.945). This high value of alpha, along with the unique impacts measured by each item, indicate that the technical skills questions measure related quantities (Taber, 2018).

For the category of "nontechnical skills", questions 11-20 had a Cronbach's alpha value of 0.955, with a 95% bootstrap confidence interval of (0.925, 0.972), indicating that the nontechnical questions measured related quantities.

For the category of "job-related impacts", questions 21-25 had a Cronbach's alpha value of 0.933, with a 95% bootstrap confidence interval of (0.884, 0.96), indicating the job-related questions measured related quantities.

Professional Curiosity

Finally, 89% of respondents (110 of 123) agreed that their participation in LISA fostered their desire to apply statistics, solve problems, and answer questions in other fields of study. 90% (111 of 123) believe their involvement in LISA will increase the chance that their work will have a positive impact on their organization.

The eleven-point scale item, "How likely is it that you would recommend involvement in LISA to a prospective statistics graduate student at VT [Virginia Tech]?" was used to create a net promoter score for each respondent (Reichheld, 2003). As shown in Figure 5, 81% of respondents are categorized as Net Promoters (9 or 10), 15% of respondents are categorized as Net Passives (7 or 8), and 4% of respondents are categorized as Net Detractors (from 0 to 6), for a net promoter score of +77, which could be considered excellent (Qualtrics, 2022).

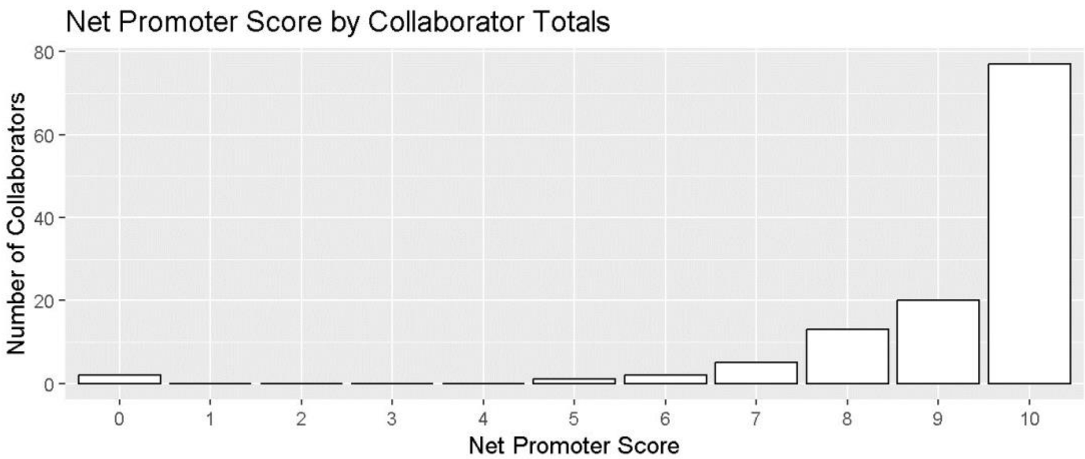

Figure 5: summary of net promoter scores.

Subgroup Comparisons

To understand how the average perceived impact scores varied between student subgroups, we compared the perceived impact scores by subgroups of collaborators. For technical, nontechnical, and job perceived impacts, we compared the following pairs of subgroups: lead vs. non-lead collaborators, master's vs. PhD students, walk-in hosts vs. non-walk-in hosts, and native English speakers vs. non-native English speakers. Using two-sample t-tests, we created 95% confidence intervals for the difference in mean scores for each comparison. All but two such intervals overlapped 0, indicating no difference for most of our comparisons. The mean perceived technical impact score for non-native English speakers was 4.12, and for native English speakers, 3.80. The 95% confidence interval for the respective difference was (0.0699, 0.575). The mean perceived technical impact score for walk-in hosts was 4.16, and for non-walk-in hosts, 3.85. The 95% confidence interval for the respective differences was (0.0538, 0.552).

We also compared the perceived technical, nontechnical, and job impacts by LISA cohort (first year as a LISA collaborator) and by the year in which the collaborator took the collaboration course (see Figure 6). With one-factor ANOVA of each impact score category, we found no statistical evidence that average impact scores differed by LISA cohort, or by course year.

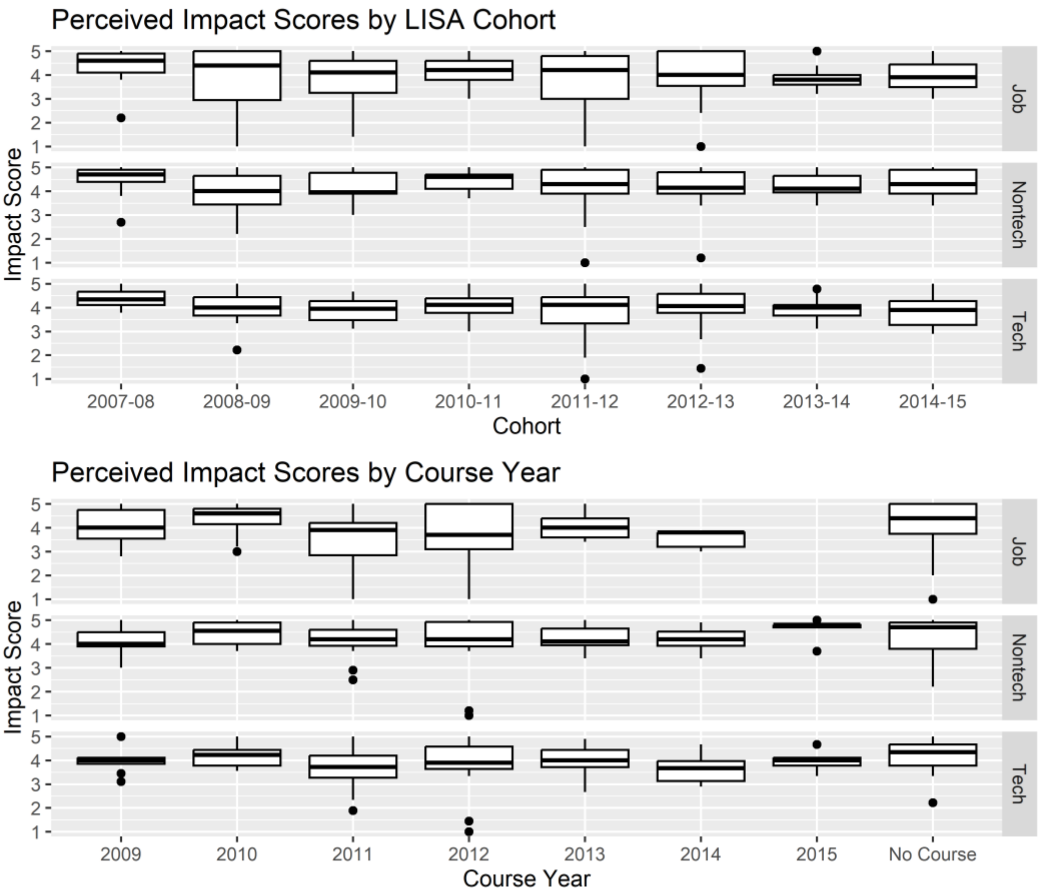

Figure 6: Perceived impact scores by category, for LISA cohort (academic year of first LISA service) and course year.

Finally, we also quantified the perceived impact of a doubling of project totals (see Figure 7), as we did not expect the perceived impact to change linearly as students worked on more collaboration projects. We fit linear regression models for the perceived impact scores by the $\log_2$ of the project totals, and computed 95% confidence intervals for the corresponding slope effects. For perceived technical impacts, the effect estimate was 0.15 with 95% confidence interval (0.0243, 0.249); for perceived nontechnical impacts, the effect estimate was 0.10 with 95% confidence interval (0.0271, 0.180); and for perceived job impacts, the effect estimate was 0.14 with 95% confidence interval (0.0751, 0.218). This means that every doubling of a student's number of collaboration projects is associated with an increased in perceived impact of 0.15, 0.10, and 0.14 for the technical, nontechnical, and job-related impacts, respectively.

.

Figure 7: Average impact scores by log (base 2) project totals.

4.2 QUALITATIVE RESULTS

To gain more insight into the impacts of LISA on collaborators, we analyzed the 146 qualitative comments from the 78 collaborators (63.4% of responders) who provided them in the survey. We discuss the quantitative results presented above by including representative qualitative comments, which we label GP, GM, CP, and CM to indicate if the collaborator was a graduated Ph.D., graduated M.S., current Ph.D., or current M.S. student, respectively.

Negative comments

Of the 146 comments, four were negative. A GM who worked on only two projects stated that their LISA experience provided, "Little to no impact." After listing two positive impacts, a CP who worked on five projects commented, "I do feel, however, that a fair amount of time was wasted in [staff] meetings." A GM who worked on four projects commented, "I learned nothing technical from LISA, which I was disappointed in." Finally, a GP who worked on 87 projects stated, "I'll start with the one negative impact LISA had on me, which was the amount of time it required, which was overwhelming at first. This was partly my own fault in terms of managing associate collaborators but another reason was I wanted to provide the best possible answer." This GP then described seven positive impacts of LISA.

Technical skills

Collaborators reported generally positive impacts of their experience in LISA on their technical skills. Many commented on how LISA helped them better understand how statistical theory and methods learned in the classroom could be applied in real world situations, and some commented on how LISA made them better statisticians. For example, a GM stated, "[LISA experience] helped me put the puzzle pieces from my course work together in a practical way." GP commented, "It provided me with real-world applications of the statistical theory that we learned in our classes, which gave me a stronger understanding of the practical applications of the theory." Another GP commented, "It challenged me to explain complicated statistical concepts to non-statisticians, which dramatically improved my understanding of the concepts. Working on LISA projects also helped me become more creative with my analysis and encouraged me to research multiple statistical methods."

Nontechnical skills

Comments on impacts on nontechnical skills were even more positive. In the words of a GP, "LISA not only provided valuable opportunities for statistics students to apply textbook learning to solving real world problems, but also cultivated a professional training environment for students to assess and develop many critical non-technical … skills such as communication, project management, and meeting facilitation." A GM commented, "It is obviously very important for a statistician to be able to analyze data properly, as well as understand and speak the language of statistics, but sometimes I think it is more important for a statistician to be able to explain statistics in plain language because what good are your results if only a small fraction of the population can understand you? LISA … truly helped me value that significance and sharpen my skills." Another GM wrote, "LISA was invaluable, as it provided a chance to practice communication, listening, presenting, and the application of statistics to a diverse set of problems and questions." Another GM: "A major lesson that I continue to keep in mind is to first fully understand what the researcher or client is truly interested in before digging deeper into specifics. Everything else does not matter if you are not on the same page with the researcher in terms of what he or she is trying to learn." Current (at the time of the survey) students also valued the nontechnical skills they learned. CP: "[Experience in LISA] improved my ability to communicate with people and cooperate as a team." CM: "It has improved my ability to effectively meet with others and help them solve problems." Another CP: "It also helped me explain statistics to non-statisticians, which is very helpful."

Job-related impacts

A CP reflected that LISA experience gets "me well prepared for job interviews and real work environment. In addition to technical and non-technical skills of collaboration, LISA gives me a chance to lead a team which definitely fostered my leadership skills." Many collaborators considered the importance of LISA in helping them get their current job. For example, a GM wrote:

> "I participated in LISA about 2 years ago, and it was a great experience. It has helped me by 1. Communicating statistics and software to scientists of different fields. I now work with engineers of many backgrounds and it's important for us to leave our jargon behind (or translate it to layman's terms) to communicate clearly. 2. Applying statistics to real problems, and dealing with the natural imperfections of real experiments and data. 3. Learning material outside of classes (models, methods, software) 4. **My time at LISA was extremely helpful in getting a job. Nearly all the employers I interviewed** [with] **were looking for applicants with the experiences listed above."**

Perhaps the most important impact of LISA was on the success of its collaborators after landing a job. A GM expressed, "LISA was exactly like my job after grad school, being an internal consultant/collaborator on methodology questions around the organization. I can't imagine anything that could have prepared me better and made me successful from day one upon graduation." A GP wrote, "My first employer counted my time with LISA as experience toward a higher pay level. When special projects came up which didn't fit into the regular process, I was often asked to work on them because I had consulting experience."

LISA experience also translated to success in academia. A GP wrote:

> "I work as a professor, and my classes are organized such that motivations of new topics are based on real life examples. This technique I learned while working for LISA, because that was for me an easier way to approach an unknown problem. I noticed that approaching this way to the problem at hand made my clients more open and more collaborative during the meetings, and now my students react similarly. Thanks to LISA I became a more proactive statistician while working on projects, and a more easy-to-follow instructor."

Impacts of LISA on life in graduate school and beyond

Almost one-sixth of the collaborators responding (19 out of 120) indicated that LISA was, "The most important positive aspect of my graduate career." GM: "LISA was one of my favorite parts of being at VT. It fired my passion for what I do. I found it rewarding, enriching, and enjoyable. I enjoyed working with everyone there." Collaborators also indicated that lessons learned in LISA influenced other aspects of their life. Another GM reflected, "LISA has made me much more aware of how communicating statistics well is just as integral as doing the actual work for a client. In my personal life, I started to notice that I make the effort to genuinely listen and understand more so than in the past." A GP summarized their experience: "**LISA improved the quality of person I am, professionally and personally.**" Some collaborators had the opportunity to travel internationally to work on LISA projects. For example, another GP stated, "My involvement in LISA had a strongly positive impact on my professional career and personal life. Socially, my time in LISA provided me the opportunity to travel to areas of the world that are not accessible to most people and to have experiences that most will not have the opportunity to have."

5. DISCUSSION

5.1. DISCUSSION OF THE 0^{TH} ORDER IMPACT OF STATISTICAL COLLABORATION LABORATORIES

Vance (2015) described a virtuous cycle of impacts achieved by stat labs that begins with the "0^{th} order impacts" on the students who gain experience moving from theory to practice when working on collaborative projects. When students are well trained and mentored, they will excel as collaborative statisticians to help domain experts solve problems and implement solutions to real-world research, business, and policy questions. These students are the immediate beneficiaries of working in a stat lab, and their experience will have long-lasting impacts. A CM wrote, "The experience at LISA let me realize I could be a good statistician, and give me advantage in job seeking. I thank LISA for leading me to the career path I'm choosing after graduation." A CP commented, "LISA has given me the confidence to speak knowledgeably about applied statistics problems. It has given me several opportunities to make an impact on real research, and to become a significantly more desirable professional candidate. I am very fortunate to have participated in LISA, and as an alumnus I will do what I can to ensure its future success."

By focusing on educating and training its collaborators and providing them with mentored experiences working on real projects with real domain experts, a stat lab can achieve the 0^{th} order impacts described in this paper. To our knowledge, this is the first study to quantify and describe these 0^{th} order impacts directly on graduate students' educational and outcomes. These impacts on the students working in a stat lab may be reason enough for statistics and data science departments and their universities to invest in similar programs. In our experience, when students are well trained, they provide excellent service to domain experts, leading to the 1^{st} order impacts of having many satisfied domain experts benefitting from the stat lab's expertise and then extolling the stat lab's virtues to their colleagues and administrators (Vance, 2015; Vance & Pruitt, 2016). Perhaps the impacts of a stat lab on its students and associated domain experts would be enough to justify such a program, yet the virtuous cycle continues. The 2^{nd} order impact of a stat lab is what it enables domain experts to do with the stat lab's statistics and data science expertise, e.g., innovate, make discoveries, advance science, improve decision-making, publish in high-impact journals, and receive grant funding. An example of this impact on a collaborator's organization was provided by a GM:

> "When interviewing for jobs, and even now in the workplace, LISA is consistently the item most interesting to employers/clients on my resume. I am in a field where data analysis is a huge part of the work, but where most analysts do not have adequate statistical skills. Experience in LISA has allowed me to relate our problems to other fields and demonstrate with real examples that problems we encounter are not unique and are in fact seen elsewhere. My management team routinely uses the experience to aggressively bid me and to justify my inclusion on projects. In summary, it has helped me associate with multiple projects internally **allowing me to have a larger impact on my organization, and has helped my organization position me onto projects that they normally would not be able to."**

5.2. DISCUSSION OF POTENTIALLY SURPRISING RESULTS

Based on our experience, we expected to see positive self-reported impacts on technical skills because working in LISA required understanding and applying statistical methods. We expected to see even higher impacts on nontechnical skills and job-related items because this was a major focus of LISA's training, particularly for non-native English speakers. The average score for nontechnical skills (4.23) was indeed higher than the score for technical skills (3.98), which was similar to the average job-related impact score (3.95). We also expected to see higher impacts on later cohorts of collaborators as LISA's training program refined and improved, for example with the introduction of Zahn's POWER structure (2019) and VCFS to LISA in 2010. Yet, as shown in Figure 6, average scores on the technical, nontechnical, and job-related impacts sections did not increase over time. On the other hand, we also recognize that many of the impacts of a stat lab on students need time to develop and become apparent to the student, and therefore we might expect greater impacts from older cohorts as they had more time to apply and appreciate their skills post-graduation. For example, a GM reflected, "I don't think I appreciated the [collaboration course] for what it was at the time. It really prepared me for collaborating with others in business settings." A second GM reflected, "I am grateful for having the opportunity to be part of LISA. I did not fully appreciate how helpful LISA was until I graduated." A third GM reflected, "LISA is a terrific experience for anybody going on into consulting. I am not sure I fully appreciated the value of the experience until after I graduated."

We also expected to see higher impacts based on how much of a "dose" of LISA experience the student received, based on their number of collaborative projects (see Figure 7), being a Lead Collaborator, or engaging in advanced LISA activities such as Walk-in Consulting and teaching short courses. We found that each additional doubling of projects was correlated with an increase in technical impact score of 0.15, an increase in nontechnical impact score of 0.10, and an increase in job impact score of 0.14 (see Section 4.1, Figure 7). Based on our own personal experience, we believe that collaborators continue to improve their skills the more projects they work on, and this trend is evident within our data. A CM wrote, "I wish I could have started my experience in LISA earlier." A CP reflected, "If I had the chance to do the graduate study again, I would have got involved in LISA much more."

5.3. LIMITATIONS AND FUTURE WORK

One of the limitations of this study is that it is based entirely on self-reflection from LISA collaborators from 2008–2015. Ideally, we would like to assess the students' collaboration outcomes and impacts objectively and correlate these measures of the effectiveness or success of their collaboration projects with objective measures of their technical and nontechnical skills. Unfortunately, tools and measures for assessing the effectiveness of collaborations and for assessing students' collaboration skills remain elusive. Solely relying on feedback from clients/domain experts is problematic because they may not fully recognize what appropriate statistical work or advice may be. For example, they would likely not be able to detect the commission of a Type III error (i.e., the right answer to the wrong question; Kimball, 1957).

Our study did not take into consideration students' initial attitudes or intrinsic starting motivation to work on collaborative projects. In our experience working with graduate students in LISA, we know that individual motivations, prior work experience, and attitudes can vary greatly and that these can impact students' engagement with, willingness to volunteer for, and learnings taken from the laboratory. Because of this, we believe it is important to explicitly teach and foster attitudes that enable collaboration (Vance & Smith, 2019; Vance 2020). While the students in our study may have started with different perspectives and attitudes, their ending attitudes were overwhelmingly positive toward statistical collaboration. Future work could explicitly measure the attitudes or students and determine how or if attitudes affect the quality of collaborations and/or the impacts on the students of working in a stat lab.

We also call for others to conduct additional surveys of their collaborators to justify the reliability and validity of such surveys, as we acknowledge the limitations of simultaneously testing our own instrument's reliability and using it to draw conclusions on the same population of graduate student statistical collaborators.

We are not aware of any published methods for evaluating the success of statistics and data science collaborations beyond the framework for assessing students' attitudes, skills, performance, and improvement described in Vance et al. (2020). These methods were not available in 2015 when this study was conducted and are still being developed and refined to better understand what specific aspects of the LISA education and training program are most effective in helping students learn essential technical and nontechnical skills and prepare them for fulfilling, meaningful careers in statistics and data science. Further research is needed on students' experience in stat labs and their impacts on student outcomes.

Finally, we understand that the nature of academic and statistical collaboration, and indeed work itself, has drastically changed after the COVID-19 pandemic of the early 2020s. None of the results described here account for now common concepts such as remote collaboration. We also call for more investigation into how statistical collaboration has changed since the pandemic, as well as attitudes toward remote collaboration.

5.4. RECOMMENDATIONS

As we have shown, when a stat lab starts with a mission to educate and train students and develops a robust training and mentoring program, its impacts on its students can be profound and wide-ranging. These initial 0^{th} order impacts on well-trained collaborators provide the foundation for providing excellent collaboration services leading to satisfied domain experts who achieve their research, business, and policy goals because of the stat lab. Therefore, we call on the statistics and data science community and higher education administrators to create more stat labs and/or strengthen their education and training programs to generate more opportunities for more students to gain impactful collaboration experiences.

Vance and Pruitt (2022) describe a seven-step process for creating new stat labs. To improve the education and training of collaborators, we recommend teaching the ASCCR Framework for collaboration (Vance & Smith, 2019; Alzen et al., 2023), Zahn's POWER process (2019), how to navigate the content of a project with the $Q_1Q_2Q_3$ framework (Trumble et al, 2022), how to create shared understanding (Vance, Alzen, & Smith, 2022), and how to ask great questions (Vance, et al., 2022). Additional resources for teaching collaboration skills are available at www.osf.io/xmtce.

6. CONCLUSION

This work is the first to our knowledge that quantifies the perceived improvement of students' technical and nontechnical skills and their preparation to enter the workforce, as opposed to outcomes and long-term impacts on researchers and scientific growth. In our study of 123 students who engaged in LISA's education and training program from 2008 to 2015, we found that involvement in LISA, as self-reported by the students:

- Enhanced technical skills, such as statistical theory, methods, and applications;
- Improved computational skills;
- Enhanced non-technical skills such as communication, collaboration, structuring meetings, and explaining statistics to non-statisticians;
- Had a positive impact on job acquisition and performance;
- Resulted in students enthusiastically recommending the experience for other students;
- Fostered a desire to apply statistics, solve problems, and answer questions in other fields;
- Increased the positive impact on one's workplace.

Based on these positive perceived impacts of experience in a stat lab and our own experience, we call for the creation and integration of more stat labs into statistics and data science programs worldwide as part of a comprehensive and modern statistics education.

ACKNOWLEDGEMENTS

We thank our past and present students and colleagues in LISA, as well as the editor and reviewers for their valuable comments. This work was partially supported by the National Science Foundation under Grant No. 1955109 for the project, "IGE: Transforming the Education and Training of Interdisciplinary Data Scientists (TETRIDS)" and by the United States Agency for International Development under Cooperative Agreement Number 7200AA18CA00022 for the project, "LISA 2020: Creating Institutional Statistical Analysis and Data Science Capacity to Transform Evidence to Action."

REFERENCES

Aldieri, L., Kotsemir, M., & Vinci, C. P. (2018). The impact of research collaboration on academic performance: An empirical analysis for some European countries. *Socio-Economic Planning Sciences, 62*, 13-30.

Alzen J. L., Trumble I. M., Cho K., Vance E. A. (2023). Training Interdisciplinary Data Science Collaborators: A Comparative Case Study. *Journal of Statistics and Data Science Education*. https://doi.org/10.1080/26939169.2023.2191666

American Statistical Association Undergraduate Guidelines Workgroup. (2014). *2014 curriculum guidelines for undergraduate programs in statistical science.* American Statistical Association.

Anderson-Cook, C. M., Hamada, M. S., Moore, L. M., & Wendelberger, J. R. (2017). Statistical Mentoring at Early Training and Career Stages. *The American Statistician, 71*(1), 6–14. Doi:https://doi.org/10.1080/00031305.2016.1200491

Arnold, J. C. (2000). Virginia Tech Department of Statistics: The First Fifty Years. *Journal of Statistical Computation and Simulation, 66*(1), 1-17. Doi:https://doi.org/10.1080/00949650008812008

Arnold, J. C., Hinkelmann, K., Vining, G. G., & Smith, E. P. (2013). Virginia Tech Department of Statistics. In A. Agresti, & X. L. Meng, *Strength in Numbers: The Rising of Academic Statistics Departments in the U. S.* (pp. 537-546). Springer. Doi:https://doi.org/10.1007/978-1-4614-3649-2_39

Blei, D. M., & Smyth, P. (2017). Science and data science. *Proceedings of the National Academy of Sciences, 114*(33), 8689-8692.

Börner, K., Scrivner, O., Gallant, M., Ma, S., Liu, X., Chewning, K., . . . Evans, J. A. (2018). Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. *Proceedings of the National Academy of Sciences, 115*(50), 12630-12637.

Hotelling, H. (1988). Golden Oldies: Classic Articles from the World of Statistics and Probability: The Place of Statistics in the University. *Statistical Science, 3*(1), 72-83.

Jeske, D. R., Lesch, S. M., & Deng, H. (2007). The merging of statistics education, consulting and research: A case study. *Journal of Statistics Education, 15*(3). ttps://doi.org/10.1080/10691898.2007.11889544

Kettenring, J. R., Koehler, K. J., & McKenzie Jr., J. D. (2015). Challenges and Opportunities for Statistics in the Next 25 Years. *The American Statistician, 69*(2), 86-90. Doi:10.1080/00031305.2015.1033987

Kimball, A. W. (1957). Errors of the third kind in statistical consulting. *Journal of the American Statistical Association, 52*(278), 133-142.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-575.

LeBlanc, M., Rueegg, C. S., Bekiroğlu, N., Esterhuizen, T. M., Fagerland, M. W., Falk, R. S., . . . Zucknick, M. (2022). Statistical advising: Professional development opportunities for the biostatistician. *Statistics in Medicine, 41*(5), 847-859. Doi:10.1002/sim.9290

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science, 35*(5), 673-702.

Love, K., Vance, E. A., Harrell, F. E., Johnson, D. E., Kutner, M. H., Snee, R. D., & Zahn, D. (2017). Developing a Career in the Practice of Statistics: The Mentor's Perspective. *The American Statistician, 71*(1), 38–46. Doi:https://doi.org/10.1080/00031305

Morton, S. C. (2005). Statistical collaboration to impact policy decisions. *Statistics in Medicine, 24*(4), 493-501.

National Association of Colleges and Employers. (2014). *2015 Job Outlook Survey.* Bethlehem, PA.

Olubusoye, O. E., Akintande, O. J., & Vance, E. A. (2021). "Transforming Evidence to Action: The Case of Election Participation in Nigeria," *CHANCE*, *34*(3), 13–23. https://doi.org/10.1080/09332480.2021.1979807

Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation, 23*(1), 21-32.

Qualtrics. (2022, April 13). *What is a good Net Promoter Score?* Retrieved from Qualtrics Experience Management: https://www.qualtrics.com/experience-management/customer/good-net-promoter-score/

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review, 81*(12), 46–124.

Sharp, J. L., Griffith, E. H., & Higgs, M. D. (2021). Setting the Stage: Statistical Collaboration Videos for Training the Next Generation of Applied Statisticians. *Journal of Statistics and Data Science Education, 29*(2), 165–170. Doi:https://doi.org/10.1080/26939169.2021.1934202

Taber, K. (2018). The Use of Crobach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education, 48*(6), 1273-1296.

Tukey, J. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, *33*(1), 1-67. Doi:https://doi.org/10.1214/aoms/1177704711

Vance, E. A. (2015). Recent developments and their implications for the future of academic statistical consulting centers. *The American Statistician, 69*(2), 127-137.

Vance, E. A. (2020). Goals for Statistics and Data Science Collaborations. In *JSM Proceedings*, Statistical Consulting Section (pp. 2198–2209). American Statistical Association.

Vance, E. A. (2021). Using Team-Based Learning to Teach Data Science, *Journal of Statistics and Data Science Education*, *29*(3), 277–296. https://doi.org/10.1080/26939169.2021.1971587

Vance, E. A., & Love, K. (2021). Building Statistics and Data Science Capacity for Development. *CHANCE, 34*(3), 38-46. Doi:10.1080/09332480.2021.1979810

Vance, E. A., & Pruitt, T. R. (2022). "Statistics and Data Science Collaboration Laboratories: Engines for Development," in O. O. Awe, K. Love, & E. A. Vance (Eds.), *Promoting Statistical Practice and Collaboration in Developing Countries* (pp. 3–26). Chapman and Hall/CRC. https://doi.org/10.1201/9781003261148

Vance, E. A., & Smith, H. S. (2019). The ASCCR frame for learning essential collaboration skills. *Journal of Statistics Education, 27*(3), 265-274.

Vance, E. A., Alzen, J. L., & Seref, M. M. (2020). Assessing Statistical Consultations and Collaborations. *JSM Proceedings* (pp. 161–169). American Statistical Association: Statistical Consulting Section.

Vance, E. A., Alzen, J. L., & Smith, H. S. (2022). "Creating Shared Understanding in Statistics and Data Science Collaborations," *Journal of Statistics and Data Science Education, 30*(1), 54–64. https://doi.org/10.1080/26939169.2022.2035286

Vance, E. A., Trumble, I. M., Alzen, J. L., & Smith, H. S. (2022). "Asking Great Questions," *Stat, 11*(1), e471. https://doi.org/https://doi.org/10.1002/sta4.471

Vance, E. A., LaLonde, D. E., & Zhang, L. (2017a). The Big Tent for Statistics: Mentoring Required. *The American Statistician, 71*(1), 15–22. Doi:https://doi.org/10.1080/00031305.2016.1247016

Vance, E. A., Love, K., Awe, O. O., & Pruitt, T. R. (2022). LISA 2020: Promoting Statistical Practice and Collaboration in Developing Countries1. In *Promoting Statistical Practice and Collaboration in Developing Countries* (pp. 27-46). Chapman and Hall/CRC.

Vance, E. A., Tanenbaum, E., Kaur, A., Otto, M. C., & Morris, R. (2017b). An Eight-Step Guide to Creating and Sustaining a Mentoring Program. *The American Statistician, 71*(1), 23-29. Doi:https://doi.org/10.1080/00031305.2016.1251493

Vance, E., & Pruitt, T. (2016). *Virginia Tech's Laboratory for Interdisciplinary Statistical Analysis Annual Report 2015-16.* Virginia Tech Laboratory for Interdisciplinary Statistical Analysis.

Zahn, D. (2019). *Stumbling Blocks to Stepping Stones.* iUniverse.

ERIC A. VANCE
Department of Applied Mathematics
University of Colorado Boulder
1111 Engineering Drive
Boulder, Colorado 80309-0526, USA

APPENDIX A: SURVEY INSTRUMENT

SURVEY QUESTIONS FOR A STUDY ON THE 0^{TH}-ORDER IMPACT OF LISA

Consent Preamble and Introduction

Dear LISA statistical collaborator,

We are conducting a study of LISA statistical collaborators to determine what impact (if any) involvement in LISA has had on your technical and non-technical statistical skills, on-the-job performance, and attitude toward statistical collaboration.

Your responses to the survey below will be kept anonymous and will be very valuable to help us document LISA's impacts. We are hoping that everyone responds to this survey, so please do share your thoughts with us on this survey.

By completing this survey you are consenting that your anonymous responses can be used in a published research study on what we call LISA's "0^{th}-order" impacts on its statistical collaborators.

Thank you!
Eric Vance, Director of LISA

[Four different versions of the survey were created: Current students received a (slightly) different survey than graduated students. The main difference was on questions about jobs. MS students received a (slightly) different version of the survey than PhD students. Question 10 in the Technical section is the main difference.]

[Experience and demographics: We used administrative records of the number of LISA domain experts and number of semesters in LISA, whether the student was currently enrolled at Virginia Tech, whether the student was an MS or PhD student, their gender, and a coding system to maintain anonymity of responses while linking these data to the survey responses. No experience questions were asked in the survey.]

Five initial Yes/No questions

1. Do you think your involvement in LISA helped you gain technical skills in statistics (understanding statistical theory and methods, applying statistics, manipulating data sets, computation, etc.)?
Yes
No
Other

2. Do you think your involvement in LISA helped you gain non-technical skills essential for success in statistics (communication, collaboration, structuring meetings, explaining statistics, etc.)?
Yes
No
Other

[For current students] {For graduated students}
3. Do you think your involvement in LISA [will help] {has helped} you advance in your professional career ([getting an interview,] getting a job, getting promoted, {acceptance into a new grad program,} etc.)?
Yes
No
Other

4. Do you believe that your involvement in LISA increased your desire to apply statistics to solve problems and answer questions in other fields of study?
Yes
No
Other

[For current students] {For graduated students}
5. Do you believe that your involvement in LISA [will increase] {increased} the chance that your work [will have] {has} a positive impact on your organization (office, company, university, etc)?
Yes
No
Other

Unless otherwise indicated, all following questions are on a 5-point Likert scale of agreement.

Impacts of LISA:
Technical
"My involvement in LISA…" (Strongly Disagree[1], Disagree[2], Neither Agree nor Disagree[3], Agree[4], Strongly Agree[5])
Q1) Led to a deeper understanding of the theory and methods of statistics
Q2) Improved my ability to apply statistical methods
Q3) Improved my ability to manage and manipulate real data sets
Q4) Improved my ability to use or run various statistical software packages (JMP, SPSS, SAS, R, etc.)
Q5) Improved my computer programming skills (coding and algorithmic thinking) [for example, in R, SAS, Matlab]
Q6) Improved my ability to solve *theoretical* statistical problems in the classroom
Q7) Improved my ability to solve *applied* statistical problems in the classroom
Q8) Helped me learn statistical methodology not covered in my classes
Q9) Provided ideas or data for my class projects
Q10) Provided topics, problems, or knowledge I used in my PhD dissertation [graduated PhD students]
Provided or will provide topics, problems, or knowledge I will use in my PhD dissertation [current PhD students]
Provided topics, problems, or knowledge I used in my MS oral exam [graduated MS students]
Provided or will provide topics, problems, or knowledge I will use in my MS oral exam [current MS students]

Non-technical
"My involvement in LISA…" (Strongly Disagree[1], Disagree[2], Neither Agree nor Disagree[3], Agree[4], Strongly Agree[5])
Q11) Improved my ability to understand non-statistical aspects of problems I work on (e.g., business or research questions)
Q12) Improved my ability to ask good questions
Q13) Improved my ability to explain statistical results in the context of the client's business or research questions
Q14) Improved my ability to structure and organize meetings.
Q15) Improved my ability to structure and organize my statistical work on projects (project management or statistical workflow)
Q16) Improved my ability to work on a team
Q17) Improved my listening, paraphrasing, and summarizing skills
Q18) Increased my awareness of the importance of non-verbal communication (body language)
Q19) Improved my oral and written communication skills
Q20) Improved my ability to successfully explain statistics to non-statisticians

On the job

{For graduated MS and PhD students} and [For current PhD students]. Current MS students are not given this section.

If you have joined the workforce after your involvement in LISA, please answer the following questions. Otherwise, skip to the next page.

"My involvement in LISA…" (Strongly Disagree[1], Disagree[2], Neither Agree nor Disagree[3], Agree[4], Strongly Agree[5])

Q21) Provided real world examples for job interviews.

Q22) Was an important factor in helping me acquire the job I wanted.

Q23) Made the transition to the workforce easier.

Q24) Made me more successful in my current job (or position if still in school).

Q25) Will help me advance in my career.

Additional multiple-choice questions:

If I could redo my time as a graduate student in the Virginia Tech Department of Statistics, I would...

1. Not participate in LISA at all.
2. Be less active in LISA.
3. Participate in LISA to the same extent.
4. Be more active in LISA
5. Be much more active in LISA

How likely is it that you would recommend involvement in LISA to a prospective statistics graduate student at VT? (net promoter score)

0 (not likely at all) – 5 (neutral) – 10 (extremely likely)

Are you a native English speaker?

a. Yes, English is my native tongue
b. No, I learned English later in life

For non-native English speakers: Involvement in LISA improved my spoken English (5 point agreement scale)

Looking back and considering all my professional experiences in grad school, my involvement in LISA was:

1. A thoroughly negative experience and a detriment to my graduate career
2. An unimportant or negative aspect of my graduate career
3. Neutral
4. An important and positive aspect of my graduate career
5. The most important positive aspect of my graduate career

Do you currently work as a "statistician" (according to your own definition)? Yes/No/Other

Open-ended comment questions:

How has your involvement in LISA (positively or negatively) impacted you? Please consider ALL the impacts (academic, social, professional, etc.)

Do you have any suggestions for improving LISA?

Any other comments?

Goodbye page

Thank you for completing this survey! Your input is very much appreciated. If you are interested in receiving occasional updates about LISA, please email Eric Vance or Tonya Pruitt.

APPENDIX B: INDIVIDUAL QUESTION SUMMARY STATISTICS

Table B1: Technical impact questions and summary statistics.

My involvement in LISA…	**n**	**Mean**	**SD**
(Q2) Improved my ability to apply statistical methods	121	4.37	0.83
(Q4) Improved my ability to use or run various statistical software packages (JMP, SPSS, SAS, R, etc.)	121	4.25	0.83
(Q8) Helped me learn statistical methodology not covered in my classes	121	4.22	0.83
(Q3) Improved my ability to manage and manipulate real data sets	121	4.16	0.90
(Q1) Led to a deeper understanding of the theory and methods of statistics	121	4.05	0.86
(Q7) Improved my ability to solve applied statistical problems in the classroom	121	4.02	0.90
(Q5) Improved my computer programming, coding, and algorithmic thinking skills (for example, in R, SAS, Matlab)	121	3.81	1.00
(Q9) Provided ideas or data for my class projects	121	3.62	1.02
(Q6) Improved my ability to solve theoretical statistical problems in the classroom	121	3.31	1.03
(Q10) Provided topics, problems, or knowledge I used in my PhD dissertation	35	3.17	1.25
Average Technical Impact Score (Q10 excluded)	**121**	**3.98**	**0.71**

Table B2: Nontechnical impact questions and summary statistics.

My involvement in LISA…	**n**	**Mean**	**SD**
(Q13) Improved my ability to explain statistical results in the context of the client's business or research questions	120	4.50	0.69
(Q20) Improved my ability to successfully explain statistics to non-statisticians	120	4.46	0.74
(Q12) Improved my ability to ask good questions	120	4.37	0.82
(Q17) Improved my listening, paraphrasing, and summarizing skills	120	4.32	0.81
(Q11) Improved my ability to understand non-statistical aspects of problems I work on (e.g., business or research questions)	120	4.23	0.84
(Q14) Improved my ability to structure and organize meetings.	120	4.14	1.00
(Q16) Improved my ability to work on a team	120	4.10	0.87
(Q18) Increased my awareness of the importance of non-verbal	120	4.08	0.98

communication (body language)			
(Q19) Improved my oral and written communication skills	120	4.07	0.96
(Q15) Improved my ability to structure and organize my statistical work on projects (project management or statistical workflow)	120	4.02	0.97
Average Nontechnical Impact Score	**120**	**4.23**	**0.74**

Table B3: Job-related impact questions and summary statistics.

My involvement in LISA…	**n**	**Mean**	**SD**
(Q21) Provided real world examples for job interviews	104	4.12	1.13
(Q25) Will help me advance in my career	101	4.08	1.05
(Q24) Made me more successful in my current job (or position if still in school)	101	3.96	1.09
(Q23) Made the transition to the workforce easier	101	3.94	1.08
(Q22) Was an important factor in helping me acquire the job I wanted	100	3.79	1.22
Average Job-related Impact Score	**104**	**3.95**	**1.02**