Journal of Statistics and Data Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ujse21

Training Interdisciplinary Data Science Collaborators: A Comparative Case Study

Jessica L. Alzen, Ilana M. Trumble, Kimberly J. Cho & Eric A. Vance

To cite this article: Jessica L. Alzen, Ilana M. Trumble, Kimberly J. Cho & Eric A. Vance (2023): Training Interdisciplinary Data Science Collaborators: A Comparative Case Study, Journal of Statistics and Data Science Education, DOI: 10.1080/26939169.2023.2191666

To link to this article: https://doi.org/10.1080/26939169.2023.2191666

Published online: 19 Apr 2023.

JOURNAL OF STATISTICS AND DATA SCIENCE EDUCATION
2023, VOL. 00, NO. 0, 1–10
https://doi.org/10.1080/26939169.2023.2191666

Training Interdisciplinary Data Science Collaborators: A Comparative Case Study

Jessica L. Alzen [a], Ilana M. Trumble [b], Kimberly J. Cho [a], and Eric A. Vance [b]

[a]Center for Assessment, Design, Research and Evaluation, University of Colorado Boulder, Boulder, CO; [b]Laboratory for Interdisciplinary Statistical Analysis, University of Colorado Boulder, Boulder, CO

ABSTRACT

Data science is inherently collaborative as individuals across fields and sectors use quantitative data to answer relevant questions. As a result, there is a growing body of research regarding how to teach interdisciplinary collaboration skills. However, much of the work evaluating methods of teaching statistics and data science collaboration relies primarily on self-reflection data. Additionally, prior research lacks detailed methods for assessing the quality of collaboration skills. In this case study, we present a method for teaching statistics and data science collaboration, a framework for identifying elements of effective collaboration, and a comparative case study to evaluate the collaboration skills of both a team of students and an experienced collaborator on two components of effective data science collaboration: structuring a collaboration meeting and communicating with a domain expert. Results show that the students could facilitate meetings and communicate comparably well to the experienced collaborator, but that the experienced collaborator was better able to facilitate meetings and communicate to develop strong relationships, an important element for high-quality and long-term collaboration. Further work is needed to generalize these findings to a larger population, but these results begin to inform the field regarding effective ways to teach specific data science collaboration skills.

ARTICLE HISTORY
Received October 2022
Accepted March 2023

KEYWORDS
Case study; Data science collaboration; Data science education; Statistical consulting; statistical practice

1. Introduction

Statisticians and data scientists empower innovation by bridging gaps across fields and helping experts in a variety of disciplines gain critical insights from data. To appropriately equip data scientists for this complex work, it is necessary to understand effective ways of teaching collaboration. There is burgeoning research regarding how to teach interdisciplinary collaboration skills (Davidson, Dewey, and Fleming 2019; Kolaczyk, Wright, and Yajima 2020; Thabane et al. 2008; Vance 2021), but these studies most often use data such as student reflections and surveys or survey responses from the domain-specific collaborators as evidence of effectiveness (Bangdiwala et al. 2002; Davidson, Dewey, and Fleming 2019; Jersky 2002; Kolaczyk, Wright, and Yajima 2020). While this is promising data, more is needed to understand the effectiveness of these approaches to teaching statistics and data science collaboration. In this study, we present a method for teaching collaboration, a framework for identifying elements of effective collaboration, and provide a comparative case study (Yin 2017) of two collaboration meetings. The purpose of this case study is to model a way to evaluate collaboration on data science projects, provide emergent evidence of the effectiveness of one approach to teaching interdisciplinary collaboration, and contribute to the literature regarding assessing elements of effective interdisciplinary collaboration more generally.

2. Literature Review

Extant literature on teaching and learning about statistics and data science collaboration indicate that this type of learning typically occurs in one of three contexts: (a) traditional courses with collaborative project assignments, (b) on-campus consulting centers or statistics labs where students collaborate with a faculty supervisor (Vance and Pruitt 2022), and (c) direct consulting projects on which students collaborate with a faculty mentor and an outside client (Belli 2001). Much of the work around the effectiveness of these programs relies on informal documentation as opposed to empirical research with rigorous methodology (Bangdiwala et al. 2002; Jersky 2002; Mackisack and Petocz 2002; Roseth, Garfield, and Ben-Zvi 2008). Other work focuses on providing descriptions of teaching statistical collaboration in order to grow exposure to the different approaches (Davidson, Dewey, and Fleming 2019; Kolaczyk, Wright, and Yajima 2020; Sima et al. 2020; Thabane et al. 2008; Vance 2021). This work provides beginning evidence of effective training in statistics and data science collaboration, but additional research must be done to build a more robust research base.

In the balance of this article, we begin by presenting an approach to teaching statistics and data science collaboration in Section 3 and then providing a definition of effective collaboration in Section 4. Next, we describe the context and data for the current study before explaining the methodology in Sections 5

CONTACT Eric A. Vance *Eric.Vance@colorado.edu* Laboratory for Interdisciplinary Statistical Analysis, University of Colorado Boulder, 1111 Engineering Dr, Boulder, CO 80309.

and 6. Following in Section 7, we present results from comparing two different data science collaboration meetings, one facilitated by data science students and another by an experienced data science collaborator. We then provide discussion of the implications of our findings and conclude with thoughts regarding the contribution of this work to the field at large in Section 8.

3. Teaching Statistics and Data Science Collaboration

Students in this study learn statistical collaboration at University of Colorado Boulder through participation in the Laboratory for Interdisciplinary Statistical Analysis (LISA) housed in the Applied Mathematics Department. LISA has a 3-fold mission:

1. Train statisticians and data scientists to become interdisciplinary collaborators,
2. Provide research infrastructure to enable and accelerate quantitative research around the campus community, and
3. Engage with the community to improve statistical skills and literacy.

Initiation into LISA begins when students take Statistical Collaboration, which is cross-listed as both an undergraduate and graduate course. The pre-requisite for this class is successful completion of an advanced statistical modeling or methods course to ensure students have sufficient technical skills to draw upon while they focus more on collaboration skills through the course. Students who take this course help LISA accomplish each area of its mission. During Statistical Collaboration, students work to master the following learning objectives:

1. Managing effective data science collaboration meetings with domain experts;
2. Communicating statistical concepts, analyses, and results to nonstatistical audiences;
3. Using peer feedback, self-reflection, and video analysis as a process for improving statistical collaboration skills;
4. Effectively collaborating with team members; and
5. Creating reproducible data science workflows and working ethically.

The course is comprised of three parts: theoretical preparation, working on actual data science collaboration projects, and reflections on the collaboration projects. In preparation for engaging in actual collaboration projects, students read several papers throughout the course. They complete in-class and homework exercises to apply the main concepts of the readings.

The bulk of the work to master the learning objectives occurs as students work on the three data science collaboration projects during the semester with individual researchers or decision makers across campus and in the community. On these projects, members of LISA are referred to as collaborators and the individuals who bring projects to LISA are referred to as domain experts. These labels work to recognize both the data science collaborator and the research collaborator as equal partners on the projects and to emphasize the importance of both roles.

In addition to working on data science projects, students collectively discuss and reflect on projects during class meetings. The model is for each LISA collaborator to be comfortable asking for and giving advice regarding technical issues of statistics and data science along with any nontechnical issues encountered during any given project. In addition to time where students can simply ask questions about how to approach a project, dedicated activities facilitate collaboration skill development for students. In particular, students participate in role playing activities regarding specific collaboration skills and video coaching and feedback sessions.

Role playing activities include practice with opening a meeting, talking about time allocations for meetings, asking great questions about a new project (Vance et al. 2022b), and summarizing action steps at the end of a meeting. During video feedback coaching sessions, collaborators watch three video clips (~1–5 min each) from a data science collaboration meeting, typically from the opening, middle, and end of the meeting. Following viewing of each clip, the faculty member along with other students in the class give feedback to the students whose collaboration meeting was recorded. Feedback may be given on topics such as if there are other questions students could have asked to gain better understanding, how the time could be spent better to ensure shared understanding about the project, and problem-solving when collaborators struggled to run the meeting in the way they hoped. At the end of the session, everyone in class shares at least one thing they learned to improve their own collaboration skills from the video feedback coaching session.

After completion of Statistical Collaboration, students have the opportunity to remain involved in LISA either on a volunteer basis or through taking Advanced Statistical Collaboration. In this course, students often continue collaborating on projects that were not completed in their previous semester or that need to be extended to a second semester. Students also continue their development of collaboration skills through additional readings, role-playing activities, engaging in more advanced collaboration skills, and mentoring junior students. One key skill in the advanced class is that of reflection. Students are first introduced to personal reflection on projects in Statistical Collaboration as modeled in class discussion and the video feedback coaching session activities. In Advanced Collaboration students extend this skill to work on reflecting with domain experts regarding what works best in collaboration meetings. In addition, advanced students work on five collaboration projects and specifically mentor novice students on those projects.

4. Defining Effective Interdisciplinary Collaboration

Our conceptual framework for identifying effective data science collaboration relies on the ASCCR frame for collaboration skills (Vance and Smith 2019). Under ASCCR, an effective collaboration is based on five components: **A**ttitude, **S**tructure, **C**ontent, **C**ommunication, and **R**elationship. We provide a high-level summary of ASCCR as a guide for evaluating an effective collaboration (Vance, Alzen, and Şeref 2020); details are provided in Vance and Smith (2019).

4.1. Attitude

Interdisciplinary collaboration relies on the data scientist having a positive attitude toward themself, the domain expert, and the collaboration team. Three attitudes that we believe enable

effective collaboration are that the data science collaborator has the willingness and ability to learn new data science skills as necessary to help the domain expert, the domain expert should be regarded as an expert in their field or for their project and should be treated as such, and that the data scientist and domain expert can accomplish more together than either could accomplish alone. It is key for the data science collaborator to consider both themself and the domain expert as equally important in the shared work of making a meaningful contribution to the field.

4.2. Structure

To maximize the effectiveness of collaborations, meetings should be structured to enable both the domain expert and the data scientist to accomplish their goals while fostering a positive relationship. The POWER process for organizing meetings (Zahn 2019) is a way to effectively structure meetings. The data scientist **P**repares for the meeting, **O**pens the meeting to make sure all attendees have the same expectations for the meeting, ensures that the **W**ork time is targeted and effective for accomplishing their shared goals, and devotes time at the **E**nd of the meeting to summarize what was accomplished and identify next steps. More advanced collaborators also take time to **R**eflect on the meeting with the domain expert to ensure that meeting time is used even more effectively and efficiently in the future.

4.3. Content

The content of collaboration meetings should consist of both qualitative and quantitative stages. In the ASCCR frame, effective collaboration follows the $Q_1Q_2Q_3$ method of addressing statistical content (Leman, House, and Hoegh 2015; Trumble et al. 2022). Under this approach, the data science collaborator first focuses on the qualitative (Q_1) aspects of the project before completing the quantitative (Q_2) analyses. Then the collaborator works to complete the content-specific contribution to the project by qualitatively (Q_3) conveying the analysis, results, conclusions, and recommendations to the domain expert for further action (Olubusoye, Akintande, and Vance 2021; Vance and Love 2021).

4.4. Communication

Effective communication is essential for successful collaboration. Key skills for effective communication include questioning (Vance et al. 2022b); listening, paraphrasing, summarizing (Vance et al. 2022a); and explaining. Statistical collaborators should use these skills to guarantee that both themselves and the domain expert understand one another at each step of the project. These communication skills should be used to elicit information and gain shared understanding (Vance, Alzen, and Smith 2022a). Shared understanding occurs when both the collaborator and the domain expert have common knowledge about project goals, project information, and the relevance of the information to the project goals. Clear communication also strengthens the relationship between the collaborator and the domain expert (Vance et al. 2022b).

4.5. Relationship

The final component of the ASCCR frame is that of relationship. Vance (2020) makes the case that, in addition to completing the project tasks, building a strong relationship with the domain expert should be a goal for every statistics and data science collaboration. Stronger relationships lead to better contributions to the field and sustained working relationships between domain experts and data scientists. As relationships grow stronger in statistics and data science collaborations, interdisciplinary fields are subsequently improved as well.

In the LISA model, students specifically learn about the ASCCR frame by reading papers explicating the five components (Azad 2015; Kimball 1957; Trumble et al. 2022; Vance et al. 2022a; Vance and Smith 2019) and conducting in-class and homework exercises. For example, early in the semester students discuss an inventory of attitudes they may or may not have regarding collaboration and which ones they believe promote or detract from collaboration. Near the end of the semester, students are presented four models for thinking about relationships, reflect on the strength of their relationships with the domain experts with whom they have collaborated, and discuss barriers and facilitators for creating strong relationships in collaborations. Additionally, the components of ASCCR are discussed during each video coaching feedback session. Students provide constructive feedback to one another regarding the ways that relevant elements of the ASCCR frame are addressed in the specific video clips.

We present this summary of the ASCCR frame to help the reader understand the overall ethos of effective collaborations. For the current study, we focus on two specific components, structure and communication, because they are the most easily observable in a single meeting out of several in a full project. We specifically seek to answer two questions. First, how do student collaborators trained in ASCCR compare to an experienced data scientist in the structure and communication components when conducting a collaboration meeting? Second, to what extent does this case study provide evidence of the effectiveness of ASCCR for interdisciplinary collaboration?

5. Data

Data for this study come from one data science collaboration project from Fall 2021 in which all individuals involved in the project completed informed consent forms approved by the University of Colorado Boulder's Institutional Review Board (Protocol 18-0554). Participants agreed to have their project documents, survey data, and video observation data used for research purposes. We use data from two Zoom recordings of the initial collaboration meeting between a Ph.D. candidate in geological sciences (the domain expert) and statistical collaborators. The domain expert first met with two student collaborators and then attended another "initial" meeting for the collaboration project with a LISA post-doctoral researcher (referred to as the experienced collaborator).[1]

[1]We purposefully chose the ordering of the meetings to favor the more experienced collaborator. Since the domain expert met with the students first, by the time she met with the experienced collaborator, she had a better idea of how to explain her project to someone in another field and

Table 1. Rubric Scores for Structure and Communication in Collaboration Meetings Scored from 1 (low) – 4 (high).

Element	Definition	Student Collaborator Scores	Experienced Collaborator Scores
	Structure		
Prepare	The collaborator has clearly prepared for the meeting as evidenced by some knowledge of the project and/or domain expert as well as practical preparation such as a shared notes document.	2	3
Opening	Meeting is opened with a friendly tone. Time is spent discussing everyone's wants for the meeting.	4	4
Work	Collaborator shapes the meeting in a way that specifically responds to the wants agreed upon during the opening conversation.	4	4
Ending	Time is reserved at the end of the meeting for summarizing outcomes and next steps.	3	3
Reflection	Collaborator allows space for and invites reflection regarding the meeting with the domain expert.	2	4
	Communication[a]		
Questioning	Collaborator effectively uses questioning throughout the meeting to provide clarity for both self and the domain expert Asks "great" questions that both elicit useful information and strengthen the relationship.	3	3
Listening	The collaborator listens attentively throughout the meeting. When the collaborator speaks, it is in direct response to the previous comment by the domain expert.	4	4
Paraphrasing	Collaborator paraphrases regularly throughout the meeting to clarify both their own language and the language of the domain expert.	3	4
Summarizing	Collaborator summarizes regularly throughout the meeting to solidify discussion topics and decisions made.	3	4

[a]Note that since this was an initial meeting during which we typically expect for the domain expert to do the majority of the talking to describe the project, we did not expect to observe or code for collaborators to explain statistical concepts as is defined under Communication in the ASCCR frame.

When student teams meet with a domain expert, one student leads the meeting. In this instance, the student leading the meeting was a senior undergraduate double-majoring in computer science and statistics/data science. This meeting was the first he led after observing and supporting in two other meetings. The second student was a Ph.D. student in aerospace engineering, and this was her second collaborative project. The experienced collaborator was a recent biostatistics Ph.D. graduate. At the time of the meeting, she had 3 years of experience working on approximately 15 collaborative projects at statistical centers in two research universities. The purpose of this meeting was for the domain expert to explain the project, give a general orientation to the statistics or data science need, and then for the collaborators to ensure they understood the project before conducting statistical analyses or providing advice regarding quantitative analysis. Meetings occurred at the end of September during Fall 2021, when the students were enrolled in Statistical Collaboration and learned about the ASCCR frame.

We selected this project for a case study because we had complete data for the project, and it is representative of a common project context for LISA. In the particular semester from which we selected this case study, LISA participated in 36 collaboration projects. Thirty-two of the projects came from the campus community. Of those 32 campus projects, 18 came from the College of Arts and Sciences, which includes geological sciences, the department in which this domain expert works. Further, 19 of the 32 campus domain experts during this semester were graduate students.

could anticipate questions or points of confusion. To this end, we expect the data to be biased in favor of the experienced collaborator and that evidence suggesting success for the students to minimize the extent of that effectiveness.

In addition to data from the meeting Zoom recordings, we also use data from a survey the domain expert completed about the students and the experienced collaborator separately after the close of the project. The survey consisted of 21 Likert-scale items and 4 open response items. We use these data sources to answer our second research question regarding the effectiveness of ASCCR for teaching collaboration from the viewpoint of a domain expert.

6. Methods

We present a comparative case study (Yin 2017) for collaboration meetings. The purpose of this case study is to provide a rich description of the different enactments of two elements of the ASCCR frame. These rich descriptions allow for a holistic understanding of the difference in these two meetings and provide insight into the varying skill levels among the collaborators and the implication of that variation on project outcomes (Stake 2013; Yin 2017).

First, two authors used the element codes in Table 1 to independently code Zoom recording transcripts[2] for evidence of each element of effective structure and communication. Following the qualitative coding process of Miles, Huberman, and Saldaña (2014), the two coders then met to reach consensus and validate codes. Next, each coder independently scored elements for overall quality using a four-point rubric[3] (see Appendix)

[2]Due to the ways the collaborators explained the multiple collaboration meetings, we were unable to blind the transcripts regarding student or experienced collaborator status and still code for all elements in Table 1.

[3]Authors from a research and evaluation center on campus, who are not affiliated with LISA, developed the rubric based on the ASCCR frame and piloted it with five coders across four disciplines outside of LISA and four

and met a second time to reach consensus on scores (Miles, Huberman, and Saldaña 2014). The rubric scores provide evidence of success in interdisciplinary collaboration as defined by adherence to the structure and communication components of ASCCR. These data[4] allow us to answer our research question about if the student team could perform comparably well to an experienced data science collaborator.

After identifying elements of effective meeting structure and communication, we used the data from the domain expert surveys as a second data source for successful collaboration. These data provide some initial insight into the effectiveness of the collaboration meetings overall.

7. Results

Table 1 shows that for each element of the rubric, the experienced collaborator scored as well or higher than the students. Of the four elements where scores differed, the students received scores one level lower than the experienced collaborator on three elements and two levels lower on the relationship element.

To better understand the evidence for these scores and to answer our first research question regarding how well students compare to an experienced collaborator, we present contrasting descriptions of the ways each collaborator engaged with structure and communication within the meetings. We examine key excerpts to identify patterns in the differences between the facilitation of each collaboration meeting. We then provide emerging descriptive evidence of the effectiveness of this approach for teaching interdisciplinary collaboration from the lens of a domain expert.

7.1. Structure in Collaboration Meetings

By organizing meetings with Zahn's (2019) POWER structure, the statistical collaborator can lessen the cognitive load of the meeting for themselves and the domain expert and help themselves focus on providing domain-specific statistics and data science advice (Vance and Smith 2019). Both the students and the experienced collaborator attend well to each element of the POWER structure in their meetings, as evidenced by the scores in Table 1. Although differences in scores do illustrate some variability in the quality of meeting structure between students and an experienced collaborator, our evidence also shows that the students learned how to follow the POWER structure well by this point in the semester.

7.1.1. Prepare

There is one clear way the experienced collaborator is more advanced in preparing for the meeting than the students. Near the beginning of the meeting, the experienced collaborator tells the domain expert that she read the domain expert's initial request for a collaboration meeting and knew some basic information about the domain expert herself.

> *Experienced Collaborator: Why don't we just start with you giving me kind of the big picture? I have read your collaboration request, so I know you're a student. You're doing geological sciences. There were a lot of words that I didn't know. {Laughs}*

The experienced collaborator shows she thought about the content of this collaboration prior to the meeting. She knows something about the domain expert and recognizes from the beginning how she will learn from the domain expert. While the students are clearly prepared to facilitate the meeting, they do not show evidence of having thought about the content of the study prior to the meeting. The lack of demonstrated preparation, however, does not seem to detract from the meeting.

7.1.2. Opening

Both the students and the experienced collaborator received the highest score possible for the meeting openings. The leading student starts the meeting with smiles and cordial introductions and then states:

> Okay. Thank you so much for meeting with us. I'm looking forward to collaborating on this project. First of all, I'd like to have a time conversation. Right now we're scheduled from 10:00 to 11:00 a.m. Does that still work for everybody? […] Okay. Great. And then the next thing that I'd like to discuss is what are your wants from this collaboration both today in this meeting and also longer term? What are you looking to get out of this collaboration? […] It sounds like you would like help with analyzing your data using statistical tools. […] Okay. Great. So if you want to tell us a little bit about your overall research goals, that would be helpful.

This student demonstrates all elements of the opening of the conversation discussed during class. The students are friendly and make sure everyone is on the same page regarding meeting timing. The student leading the meeting also provides the domain expert an opportunity to define objectives for the meeting and asks about the overall research project. This is very similar to the way the experienced collaborator begins her meeting:

> Usually how we start is to have a time discussion. I have an hour blocked out for this meeting until 10:00 a.m. Does that work for you? […] If we're being productive, can you go a little over 10:00 a.m. like ten or 15 minutes? […] Cool. And then the next thing we usually like to go through is what we each want to get out of the meeting. I've listed some generic things that I usually try to get out of my initial meeting: an overview of your research, the big picture of what you're doing and the real-world implications that it has, what exactly your hypothesis question is, if you know that, what does your data look like, and why you've come to us for collaboration. And then the last thing is what your deadlines are. Is there anything else you want to put on the agenda for today?

The experienced collaborator includes the same elements of the opening to a meeting as the student collaborators with just a bit more detail. Both openings ensure everyone is on

collaboration videos. Raters independently scored videos in teams of two and then met to discuss necessary changes to the rubric to appropriately identify differences in quality across score levels as well as to clarify language in the rubric.

[4]Full transcripts are available on the Open Science Framework (OSF; *https://osf.io/c95jd/*) for readers to use with the included rubrics and paper text to better understand the coding, reproduce the results, and repeat the evaluation process in other contexts.

the same page regarding the time available for the meeting and identifying and paraphrasing the domain expert's objectives for the meeting. The experienced collaborator provides a bit more in her opening by naming specific objectives she typically has for an initial collaboration meeting. Although a bit less direct, the students solicit similar information about the project by inviting the domain expert to share her overall research goals.

7.1.3. Work

Effective collaborators shape meetings in a way that specifically responds to the wants or objectives agreed upon during the opening conversation. The overall goal for this meeting was for the domain expert to give the collaborators a general sense of her project and to identify the statistical problem about which she came to LISA for help. The bulk of this section of the meeting includes the domain expert sharing details about her project and onboarding the collaborators to her content. Throughout this process, both the students and the experienced collaborator show evidence of attending to the stated objectives of understanding the research project and helping the domain expert think about the research analysis. The students regularly make statements like the following:

> To check my understanding…
> Really quickly before we go [on], I want to understand […] better.
> What does that comparison look like?

Similarly, the experienced collaborator made comments such as these:

> Hold on. I'm trying to understand this…
> We're focusing on this first bullet point?
> What is this called on the x-axis?

In addition to pausing the domain expert to ask questions and make sure they understand details, the collaborators also ask the domain expert to check their understanding. In both quotations below, the collaborators summarize the details of what they identify as the main task of the collaboration. The domain expert is interested in using statistical analysis to show that there are significant differences in facies, or rock types, about which she had collected data.

> *Student 1: You're looking for the significant differences in your measurements between facie at the same stratigraphic height or between stratigraphic heights or both? […] What you're looking for from us right now. It's more of just the analysis between stratigraphic heights. […] You want to show that the facies are in fact facies.*
> *Experienced Collaborator: You want to look at carbon and oxygen composition between the facies and then you also want to look at temperature between the facies. […] And in order for you to classify something as different, you would be looking at if the delta 180 thing that you have is different. You want to tell between the different [facie]?*

In response, the domain expert affirms that the collaborators all understand her goals for the project. The use of domain-specific vocabulary by all collaborators helps ensure that they have a deep understanding of the domain expert's problem and research questions and strengthens the relationship by demonstrating an effort to use the domain expert's language, thus, making for a more effective collaboration.

7.1.4. Ending and Reflection

During the end of the meeting, the students and the experienced collaborator make similar moves. Both reserve time at the end of the meeting for a closing conversation, summarize what was discussed, and make plans for next steps.

> Student1: Okay. We only have about ten minutes left in our initial planned meeting time. […] I think we can summarize.
> Experienced Collaborator: Okay. I'm just looking at the wants, goals again. So I think I've gotten through everything.

Both also have some element of reflection in the meeting, but the experienced collaborator is more explicit about it and provides more space for the domain expert to give direct feedback.

The students exchange multiple generally positive and reflective comments with the domain expert toward the end of the meeting such as the following:

> Student1: This was super helpful.
> Student2: I think we made a lot of progress today. And it helps us to understand your greater goals first before we jump into the analysis just so we understand what's going on.
> Domain Expert: Super
> Student2: Thanks for being patient with us [as we understand your project]
> Domain Expert: No. This is fun for me.
> Student1: I think I have a few ideas for potential analyses, but I have some things I want to look in to first.
> Domain Expert: Totally
> Student2: I have some ideas, but I would want to look into them more before I would give a recommendation as well.
> Domain Expert: That sounds great.

The domain expert responds positively to the collaborators throughout the meeting and has a general tenor of excitement moving forward at the end of the meeting. This is all evidence to suggest that the students developed a positive connection with the domain expert, and they are all eager to continue working together.

The experienced collaborator and the domain expert exchange similarly positive comments throughout their meeting.

Experienced Collaborator: I didn't quite catch the last thing you said.

Domain Expert: Oh, I was just being stoked about my own science.

> Experienced Collaborator: It's really cool.
> —
> Domain Expert: I appreciate you letting me take a lot – so much time to be like, "This is a rock."
> Experienced Collaborator: {Laughs} I'm so glad you did. It's so interesting.
> Domain Expert: Yes, yes. It's fun.

These two exchanges are not unlike the example showed above from the student collaborators. However, the experienced

collaborator makes one additional move the students did not that elevates her reflection during the meeting and resulted in the higher rubric scores in Table 1.

> Experienced Collaborator: Another thing we like to do at the end of the meetings is to reflect with the domain expert and say, "How did this meeting go for you? In the future, is there anything that we could do differently to make the meeting easier or anything that for you?"

In response to this question the domain expert first comments on how the process of onboarding multiple collaborators to her project was a good experience for her and helped her think about how she explains her research to others. She also gives some feedback regarding how there were some vocabulary words (e.g., independent) that seemed to operate differently in her field of geology compared to statistics. The domain expert points out that making sure everyone understood vocabulary in the same way would help to avoid misconceptions.

By and large, the students and the experienced collaborator facilitate their initial meetings with this domain expert in very similar ways. The meetings are well-organized, cover all aspects of the POWER structure (e.g., preparation, open, work, ending, and reflection), and achieve the goals of the collaborators understanding the overall project research goals. There are some small ways in which the experienced collaborator added additional levels of expertise to the structure of the meeting that not only led to greater shared understanding about the project but also worked to build a stronger relationship between herself and the domain expert.

7.2. Communication in Collaboration Meetings

Both the student and experienced collaborators demonstrate evidence of every element of communication in their respective meetings. All collaborators clearly listen intently to the domain expert as she explains her project. Each collaborator also responds to the domain expert input with thoughtful comments and intelligent questions such as the following:

> Student1: Would this whole stratigraphic section be considered one type, then, in your analysis or in your data the way that you've been recording it? Do you designate it based on what it is right now? How do you designate type?
> Experienced Collaborator: In the geological sciences, is there a certain number of categories of facies, or is it kind of like a continuous scale? Or what is it? How does it work?

Both the students and the experienced collaborator also make sure to paraphrase regularly to ensure they understand the new domain-specific information. Finally, both meetings have regular moments where the collaborators summarize the thoughts just shared or next steps in the project. There are few differences in the skills shown in these elements of communication throughout both meetings.

The student collaborators received lower scores on paraphrasing and summarizing because the experienced collaborator more frequently made connections to the overall project rather than focusing on specific details, likely resulting in better comprehension of the domain expert's research. The experienced collaborator also more frequently uses the domain expert's domain-specific vocabulary.

> Student1: I think what you were saying is you think that it might be different at these different stratigraphic heights, but you're not sure yet.
> Experienced Collaborator: The Jacob's staff allows you to be able to tell the height or how deep these things are, what kind of rock it is, and then, also, you can look at sort of – is it the CaCO3 nodules that tell you what kind of environment it was?

The ASCCR frame suggests that while statistical collaborators work to understand and to be understood in a project, they also use communication to build a relationship with the domain expert. As seen in the structure results above, this is where we see nuanced differences between the students and experienced collaborator. Although the students certainly have moments where we observe a positive relationship being built with the domain expert, we see deeper moments with the experienced collaborator.

> Experienced Collaborator: In the geological sciences, is there a certain number of categories of facies, or is it kind of like a continuous scale? Or what is it? How does it work?
> Domain Expert: […It's like] "Okay. I have this rock in my hand, so it had to have formed somehow." So now I have to figure out what setting and sequence of conditions provided the right moment for this rock to form.
> Experienced Collaborator: Kind of like playing detective.
> Domain Expert: It's totally like playing a detective with rocks {Laughter}
> Experienced Collaborator: You're a professional rock detective.
> Domain Expert: Yes, exactly. It's very fun.
> —
> Domain Expert: I ran a couple of ANOVAs on my data and was like, "I don't know how to do this prudently." And I know that there is a lot of ways that one could just get the right – or get an answer you want out of stats, and I just really felt like a kid using a power tool that I wasn't supposed to be doing without supervision. I was like, "I probably should go get some handrails to make sure I use those correctly." {laughs}
> Experienced Collaborator: Well, I appreciate that. Yay, I'm so excited that you're here. So as part of what you maybe want in addition to us helping with this particular analysis is maybe kind of – it's a learning experience for you too, and you want to be able to understand what we're doing. And then maybe you came across a similar problem in the future, like you would have enough understanding to kind of do it on your own. Is that what your goal is?

In both exchanges the experienced collaborator and the domain expert share moments of connection around the work. The experienced collaborator not only speaks with the goal of understanding the project, but she includes an element of interpersonal relationship. These excerpts show how the experienced collaborator used communication to accomplish the tasks of the meeting and develop a relationship with the domain expert, the two end goals of a collaboration under ASCCR.

7.3. Domain Expert Evaluation

Our qualitative data show that both the student and experienced collaborators include all the expected elements of structure and communication from the ASCCR frame in their collaboration meetings with this domain expert. More than that, our data show that the quality of the structure and communication are similar between the students and the experienced collaborator. We see some evidence of the experienced collaborator adding a subtle level of improved collaboration skills, but the data suggest that both collaboration meetings were successful.

In addition to coding transcripts, we also asked the domain expert to complete a survey[5] at the end of the project in which she evaluated her experiences with both collaboration teams. A limitation to this data source is that the domain expert completed these surveys after six meetings with the student collaborators over the course of the semester and after only the initial meeting with the experienced collaborator. The repeated initial meetings allowed us to study how student collaborators compared to an experienced collaborator during the initial meeting, but the experienced collaborator only served as a consultant for the student team following the double initial meeting. While this survey data is problematic for direct comparisons between the experienced collaborator and the student collaborators in projects overall, it does still serve as a useful tool for understanding the domain expert's experience with the student collaborators as well as some nascent information regarding effectiveness of collaboration in general.

When asked if her collaboration with LISA was helpful and if she was satisfied with her experience, the domain expert gave the students and the experienced collaborator scores of 6 (strongly agree) and 4 (somewhat agree; on a 6-point scale), respectively. This makes sense as the students were the ones who helped the domain expert through the full project. In another item about feeling welcomed during collaboration meetings that used the same scale, the domain expert rated both the students and the experienced collaborator a 5 (agree). These scores represent the general pattern in scoring across all items. For each, the domain expert either gave both the students and the experienced collaborator equally high scores, or she gave the experienced collaborator a slightly lower score than the students. However, all ratings were on the positive end of the scales. The domain expert reported that all collaborators made her feel welcomed both during and outside of meetings and that she would recommend the statistical lab to colleagues after her experiences with this collaboration. Finally, the domain expert indicated an equally high level of relationship between herself and all collaborators.

When given the chance to provide some qualitative feedback regarding the collaboration, the domain expert indicated that the student collaborators "were very respectful and clearly were taking time to understand aspects of my research most relevant to the statistical goals. [Student1] was really excellent. [He was] very professional and thoughtful in the questions he asked regarding the research." In the qualitative items regarding the experienced collaborator, the domain expert noted:

> Meeting with [the experienced collaborator] was pleasant and a positive experience, but I have to admit that I was confused to as why I met with her separately to the other student collaborators and then had no further contact with her. I thought we were going to all meet together (me, student collaborators, and her) after we had had separate onboarding meetings, but I guess I misunderstood that aspect of the process. With that in mind, giving detailed feedback on how effective the collaboration was with [the experienced collaborator] is difficult because I didn't work with her besides our one meeting during which I explained my research.

In this instance, the experienced collaborator did not communicate to the domain expert how the project would progress, resulting in a lack of shared understanding. The domain expert had the impression that the experienced collaborator would be involved in all subsequent meetings, but she only provided support to the students when needed. This resulted in the experienced collaborator receiving lower scores than the students, which is evidence for the need to create shared understanding for a successful collaboration.

8. Discussion

This comparative case study provides illustrative examples of the structure and communication elements of ASCCR in two collaboration meetings. Further, this study provides some emerging tools for evaluating collaboration meetings grounded in collaboration theory (Vance and Smith 2019). This study differs from previous research regarding the effectiveness of teaching interdisciplinary statistical collaboration in that it explicitly defines specific elements of statistics and data science collaboration, uses pre-defined criteria to evaluate particular elements of collaboration meetings, and provides data from student collaborators alongside an experienced collaborator (Bangdiwala et al. 2002; Jersky 2002; Mackisack and Petocz 2002; Roseth, Garfield, and Ben-Zvi 2008). The case study method allows for the comparative inspection of detailed elements of the collaboration meetings. It also provides real world examples of structure and communication skills in action. By aligning evidence from collaboration meeting transcripts with the theoretical framing of effective collaboration and feedback from the domain expert, this study expands the body of evidence regarding the effectiveness of teaching interdisciplinary collaboration.

A limitation of this approach is the small sample size. While an in-depth look at two meetings allows for careful scrutiny of individual collaboration skills within meetings, there is limited ability to generalize the results to comparisons between student and experienced collaborators at large. Future work will involve expanding the datasets by comparing more student and experienced data science collaboration meetings in order to bolster these claims. Such research will provide more information regarding the most effective ways to teach data science students to become successful interdisciplinary collaborators.

Practice is essential for developing collaboration skills. For example, using the domain expert's vocabulary and making clear connections between specific points in the project and the overall project goals are ways in which novice collaborative

[5]Full survey available at (*https://osf.io/c95jd/*)

data scientists may build their craft through practice. Additionally, building relationships throughout collaboration projects is something that may come more naturally through practice as inexperienced collaborators become more comfortable with asking questions and learning about new domains in ways that not only build their understanding of new fields but also develop strong relationships with new colleagues.

We believe that statistics and data science educators should consider the ASCCR frame for teaching collaboration skills to their students. It provides a straightforward and clear way to implement successful interdisciplinary collaboration and names discrete skills to model and practice. Teaching collaboration with ASCCR can advance the field in training data scientists and push our thinking regarding successful interdisciplinary collaboration. Professional data scientists can also benefit from implementing ASCCR in their future collaborations.

Most statisticians and data scientists will inevitably find themselves working on interdisciplinary teams at some point in their careers. To equip future data scientists to contribute most effectively to their own field as well as others, we must understand the best ways to support them in developing collaboration skills. The emerging evidence provided here suggests that training data scientists in the ASCCR frame can help improve collaboration skills. We encourage statistics and data science educators and practitioners to consider implementing aspects of the ASCCR frame and the related rubric to expand the data from this study to continue to build the body of evidence of effectively teaching interdisciplinary collaboration. We believe that the ASCCR frame can improve the quality of collaborations and ultimately improve collaborative data scientists' ability to transform evidence into action.

Acknowledgments

The authors report there are no acknowledgments to declare.

Data Availability Statement

The data that support the findings of this study are openly available in the Open Science Framework at *https://osf.io/c95jd/*.

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

This work was supported by the National Science Foundation under Grant No. 1955109 and Grant No. 2022138 for the projects, "IGE: Transforming the Education and Training of Interdisciplinary Data Scientists (TETRDIS)" and "NRT-HDR: Integrated Data Science (Int dS): Teams for Advancing Bioscience Discovery." This work was also supported 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."

ORCID

Jessica L. Alzen https://orcid.org/0000-0002-1706-2975
Ilana M. Trumble https://orcid.org/0000-0003-0531-6584
Eric A. Vance https://orcid.org/0000-0001-5545-1878

References

Azad, K. (2015), *Math, Better Explained: Learn to Unlock Your Math Intuition*, Scotts Valley, CA: CreateSpace Independent Publishing Platform.

Bangdiwala, S. I., Amarillo, M. L., Ughade, S., Rodriguez, M. N., Singer, J. M., Munoz-Navarro, S. R., Komoltri, C., and Cumsille-Garib, J. F. (2002), "Teaching Consultancy Through Direct Experience in Research: The Approach of the Department of Biostatistics at the University of North Carolina at Chapel Hill," in *Proceedings of the Sixth International Conference on Teaching Statistics*, pp. 1–5.

Belli, G. (2001), "The Teaching/Learning Process in University Statistical Consulting Labs in the United States ," in *Training Researchers in the Use of Statistics*, eds. C. Batanero, pp. 325–338, Granada, Spain: International Association for Statistical Education and International Statistical Institute.

Davidson, M. A., Dewey, C. M., and Fleming, A. E. (2019), "Teaching Communication in a Statistical Collaboration Course: A Feasible, Project-Based, Multimodal Curriculum," *The American Statistician*, 73, 61–69. DOI:10.1080/00031305.2018.1448890

Jersky, B. (2002), "Statistical Consulting with Undergraduates – A Community Outreach Approach," in *Proceedings of the Sixth International Conference on Teaching Statistics*, 1–5.

Kimball, A. W. (1957), "Errors of the Third Kind in Statistical Consulting," *Journal of the American Statistical Association*, 52, 133–142.

Kolaczyk, E., Wright, H., and Yajima, M. (2020), "Statistics Practicum: Placing 'Practice' at the Center of Data Science Education," *Harvard Data Science Review*, 3, 1–28. DOI:10.1162/99608f92.2d65fc70.

Leman, S., House, L., and Hoegh, A. (2015), "Developing a New Interdisciplinary Computational Analytics Undergraduate Program: A Qualitative-Quantitative-Qualitative Approach," *The American Statistician*, 69, 397–408. DOI:10.1080/00031305.2015.1090337.

Mackisack, M., and Petocz, P. (2002), "Projects for Advanced Undergraduates," in *Proceedings of the Sixth International Conference on Teaching Statistics*, pp. 1–6.

Miles, M. B., Huberman, A. M., and Saldaña, J. (2014), *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.), Los Angeles: Sage Publications.

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

Roseth, C. J., Garfield, J. B., and Ben-Zvi, D. (2008), "Collaboration in Learning and Teaching Statistics," *Journal of Statistics Education*, 16, 1–15. DOI:10.1080/10691898.2008.11889557.

Sima, A. P., Rodriguez, V. A., Bradbrook, K. E., DiPace, B. S., and Okhomina, V. (2020), "Incorporating Professional Recommendations Into a Graduate-Level Statistical Consulting Laboratory: A Case Study," *Journal of Clinical and Translational Science*, 5, 1–9. DOI:10.1017/cts.2020.527.

Stake, R. E. (2013), *Multiple Case Study Analysis*, New York: Guilford Press.

Thabane, L., Walter, S. D., Hanna, S., Goldsmith, C. H., and Pullenayegum, E. (2008), "Developing a Biostatistical Collaboration Course in a Health Science Research Methodology Program," *Journal of Statistics Education*, 16, 1–16. DOI:10.1080/10691898.2008.11889564.

Trumble, I. M., Alzen, J. L., House, L. L., and Vance, E. A. (2022), "The Content of Statistics and Data Science Collaborations: The QQQ Framework," Preprint Available on ArXiv. *https://arxiv.org/abs/2209.04887*

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

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

Vance, E. A., Alzen, J. L., and Şeref, M. M. H. (2020), "Assessing Statistical Consultations and Collaborations," in *JSM Proceedings, Statistical Consulting Section*, pp. 161–169.

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

Vance, E. A., and Love, K. (2021), "Building Statistics and Data Science Capacity for Development," *CHANCE*, 34, 38–46. DOI:10.1080/09332480.2021.1979810.

Vance, E. A., and Pruitt, T. R. (2022), "Statistics and Data Science Collaboration Laboratories: Engines for Development," in *Promoting Statistical Practice and Collaboration in Developing Countries*, eds. O. O. Awe, K. Love, and E. A. Vance, pp. 3–26, Boca Raton, FL: CRC Press. DOI:10.1201/9781003261148

Vance, E. A., and Smith, H. S. (2019), "The ASCCR Frame for Learning Essential Collaboration Skills," *Journal of Statistics Education*, 27, 265–274. DOI:10.1080/10691898.2019.1687370.

Vance, E. A., Trumble, I. M., Alzen, J. L., and Smith, H. S. (2022b), "Asking Great Questions," *Stat*, 11. DOI:10.1002/sta4.471.

Yin, R. K. (2017), *Case Study Research and Applications: Design and Methods* (6th ed.), Thousand Oaks, CA: Sage Publications.

Zahn, D. (2019), *Stumbling Blocks to Stepping Stones: A Guide to Successful Meetings and Working Relationships*, Bloomington, IN: iUniverse.

Appendix

Table A1. Rubric for structure and communication.

	Structure			
	4 Mastery	3 Competent	2 Developing	1 Minimal
Preparation	The collaborator has clearly prepared for the meeting as evidenced by general knowledge about the project as well as some thoughts about the project beyond what the domain expert shared in the project request.	It is clear that the collaborator has reviewed the collaboration request and has done some preparation for the meeting. They appear generally familiar with the project but have perhaps not done any other preparation for the meeting.	Collaborator does not do anything that specifically reveals they have spent time thinking about the project ahead of time. However, this does not appear to detract from the meeting's success.	No evidence of preparing for the meeting evident. Collaborator appears to have no prior knowledge of the project. Collaborator is ill-prepared to engage in the project.
Open	Meeting is opened with a friendly tone. Time is spent discussing everyone's wants for the meeting.	Meeting is opened with a friendly tone. Time is spent discussing everyone's wants for the meeting, but may be hurried, rushed, or incomplete.	Collaborator may or may not take time for friendly interaction at the beginning of the meeting. If goals/an agenda are shared, the collaborator tells the domain expert rather than including them in the process.	Collaborator doesn't take time for friendly interaction at the beginning of the meeting. Goals and/or an agenda are not shared at the beginning of the meeting.
Work	Collaborator shapes the bulk of the meeting in a way that specifically responds to the wants agreed upon during the opening conversation.	Collaborator attempts to address the wants agreed upon during the opening conversation but is sometimes unsuccessful.	Collaborator does little to address domain expert wants and focuses more on own ideas for addressing the project goals.	Collaborator ignores domain expert wants and jumps into project statistics with no consideration for project context of domain expert wants.
End	10%–20% of meeting time is reserved for the closing conversation. All applicable elements of "end" are present.	Time is reserved for the closing conversation, but it may be insufficient. Some elements of "end" are missed.	Collaborator attempts to cover some elements of ending a meeting via the POWER structure, but time was not reserved. The conversation is rushed and incomplete.	No time is reserved for meeting closure activities. Collaborator does not attempt to cover elements of ending a meeting via POWER structure.
Reflection	Collaborator allows space for and invites reflection with the domain expert.	Collaborators reflect on the meeting after the domain expert leaves.	Collaborator asks for or provides space for reflection but does not reserve sufficient time for the conversation.	Collaborator does not provide space for meeting reflection.
	Communication			
Questioning	Collaborator effectively uses questioning throughout to provide clarity for both self and the domain expert. Asks "great" questions that elicit useful information and strengthen the relationship.	Questioning is attempted frequently during the meeting but is sometimes insufficient for gaining clarity for collaborator or domain expert.	Collaborator asks questions to help clarify for self or the domain expert but not both regularly throughout the meeting.	Questioning is not evident in the meeting.
Listening	The collaborator listens attentively throughout the meeting. When the collaborator speaks, it is in direct response to the previous comment by the domain expert.	The collaborator listens attentively throughout the meeting. When the collaborator speaks, it is in direct response to the domain expert. However, sometimes the collaborator does not fully respond to domain expert comments.	The collaborator occasionally jumps ahead of the domain expert to talk about their own desired topics. Responses sometimes do not make sense after domain expert comments.	The collaborator regularly jumps ahead of the domain expert to talk about their own desired topics. Responses do not make sense after domain expert comments.
Paraphrasing	Collaborator paraphrases regularly throughout the meeting to clarify both their own language and the language of the domain expert.	Collaborator paraphrases regularly throughout the meeting for clarity but may only do this for the domain expert or themselves.	Collaborator occasionally paraphrases during the meeting to add clarity for either themselves or the domain expert.	Collaborator rarely, if ever, paraphrases during the meeting.
Summarizing	Collaborator summarizes regularly throughout the meeting to solidify discussion topics and decisions made.	Collaborator summarizes regularly throughout the meeting but may miss some key moments that would have benefited from summarizing.	Collaborator occasionally summarizes, but individuals will leave with lack of clear understanding of the discussions and decisions.	Collaborator rarely, if ever, summarizes during the meeting.