© Journal of Higher Education Outreach and Engagement, Volume 27, Number 1, p. 5, (2023)
 eISSN 2164-8212

# The Struggle Animates the Learning: Exploring Student Experiences with a Community-Engaged, Project-Based Course on Evaluation

*Sarah V. Suiter, Kathryn Y. Morgan, and Amie Thurber*

**Abstract**

For instructors engaged in teaching evaluation, bridging the gap between the content of formal educational experiences and what we want future evaluators to be able to do in practice remains a challenge. Studying the format and quality of university courses focused on program evaluation is one mechanism through which we might begin to narrow this gap. This article describes a community-engaged, project-based evaluation course that was taught during five semesters, and uses qualitative data to explore student experiences within the course along three dimensions: experiential education, interdisciplinary collaboration, and community partnerships. In particular, we highlight the productive yet uncomfortable role that challenge and ambiguity play in animating evaluation learning. We suggest implications for teaching evaluation based on our findings.

*Keywords: teaching evaluation, community-engaged learning, project-based learning, graduate student development*

Scriven (1991) described evaluation as a transdiscipline, noting that it is a discipline that serves and intersects with many others, as well as one that spans research and practice. This means, among other things, that evaluators are prepared in a variety of disciplinary programs and professional trajectories (LaVelle & Donaldson, 2015), leading to a diverse educational profile for professional evaluators. As a result, there exist some differences of opinion about the core knowledge and skills required for evaluators, and the best ways to deliver them to aspiring professionals (Christie et al., 2014). In general, however, scholars note the importance of hard skills (e.g., research design, instrument creating, data analysis) *and* soft skills (e.g., relationship building, communication, team management) in evaluation training (King et al., 2001; Russ-Eft et al., 2008). Despite this emphasis, researchers repeatedly find gaps between what evaluators are trained to do and what they must do in practice, especially related to soft skills (Dewey et al., 2008; Galport & Azzam, 2017). For example, in a survey of evaluation job-seekers, Dewey et al. (2008) asked respondents to rate their perceived ability in a list of evaluation competencies, and also asked which of the competencies had been taught in their degree programs. Aside from writing syntax (7%), the fewest respondents reported that they had been taught "relating to clients or stakeholders" (22%), "project and/or team management" (21%), or "project planning" (28%). At the same time, evaluation employers rated these as some of the most important competencies among people they hire, and rated relating to clients or stakeholders and project and/or team management as two of the areas in which they perceive the biggest gap between the needs of the field of evaluation and educational and/or practical experiences of potential evaluators (Dewey et al., 2008). Findings from a more recent survey of professional evaluators suggest this need persists: Respondents indicated that interpersonal competence and reflective practice

were the two domains within which evaluators are most in need of additional training (Galport & Azzam, 2017). Although the gap between the content of formal education and the expectations of practice likely exists in all professions, it is nevertheless worthwhile to consider the ways in which evaluation programs, courses, and professional development activities can better align evaluator preparation with the necessities of the field.

Scholars in the field of teaching evaluation note four primary modalities through which evaluation knowledge is typically developed: university programs, professional development workshops, webinars, and on-site training opportunities (LaVelle & Donaldson, 2015). In a survey of American Evaluation Association members, respondents who reported taking evaluation-specific courses had done so in a professional development workshop format (Christie et al., 2014). Many practicing evaluators report having taken only one evaluation-specific course (Christie et al., 2014), and most evaluation programs in the United States report having only two or three evaluation-specific courses (LaVelle & Donaldson, 2010). Nevertheless, LaVelle and Donaldson (2015) argued that "preservice education of evaluators is integral to quality evaluation practice as well as socialization into the evaluation profession" (p. 40). One way to meet this need is by integrating evaluation training into the professional education of students preparing to work in service-related fields (e.g., education, nonprofit management, public health; Bakken et al., 2014; Davis, 2006). Although a natural outgrowth of these observations is to recommend more robust academic programs devoted to evaluation (and the authors agree), an equally important strategy is to create robust opportunities for applied, project-based learning through community-engaged evaluation courses. In this approach to teaching evaluation, students apply their in-class learning to help meet the evaluation needs of a community partner, thus gaining knowledge and experience in both the hard and soft skills of evaluation (Bakken et al., 2014; Davis, 2006; Gredler & Johnson, 2001; Suiter et al., 2016).

Applied evaluation courses are an example of publicly engaged instruction (Doberneck et al., 2010); there are expected readings, assignments, and synchronous sessions for students, accompanied by community-engaged work by students, which may include developing an evaluation tool, a comprehensive evaluation plan, and/or completing analysis of evaluation data and disseminating results. The literature suggests some variation in how applied evaluation courses are structured: Some courses engage all students with a single community partner (Bakken et al., 2014); in other courses students are grouped in teams to support the needs of different community organizations (Suiter et al., 2016). Courses also differ on the extent to which the community partner participates, ranging from staff attending a single course session (Davis, 2006) to participating in the entirety of the course (Suiter et al., 2016).

Applied evaluation courses can serve as a critical form of professional socialization, particularly at the graduate level. O'Meara (2008) noted, "During the process of socialization, a person takes on characteristics, values, and attitudes, as well as knowledge and skills, that contribute to a new professional self" (p. 29). This socialization process is nurtured over the course of graduate education. As outlined by Weidman et al. (2001) and summarized below, graduate students entering professional programs generally progress through a four-stage developmental process. The anticipatory stage includes the application and admission process; students often enter with an idealized set of expectations for their future practice and are novices in the field. In the formal stage, students engage in coursework to receive the formal introduction to the knowledge and skills needed in the field, and move from novice to apprentice, beginning to apply their learning under the guidance of instructors. As graduate students progress into the informal stage, they deepen their knowledge acquisition and learn more about the expected behaviors and practices of their field. In this stage, learning often occurs through immersive experiences with peers and faculty, though increasingly outside formal instructional interactions. Professional students begin to shift from seeing themselves as a student to an emerging sense of self as a professional. As students successfully matriculate through graduate training, they enter the personal stage, wherein they internalize a new professional identity, synthesize learning and practice experiences, and deepen engagement in their professional community.

Pedagogically, applied evaluation courses help students bridge the formal and informal stages of their development. This

type of instruction provides some structural elements of the formal stage (such as course readings and periods of instruction) while introducing conditions of the informal stage, principally a focus on applied, project-based learning. Importantly, applied evaluation courses require students to engage with layers of complexity and difficulty that mirror the work of professional evaluators. Learning through difficulty is in fact a principal strength of community-engaged learning. As noted by Warner (2020), "As they engage in the community, students interact with unfamiliar people, settings, and tasks that generate a level of 'disequilibrium' and anxiety that stimulates deep learning" (p. 436). For students in applied evaluation settings, this disequilibrium often results from a fixed time span, limited instructor-derived requirements, and complex group dynamics (Dewey et al., 2008; Trevisan, 2004).

Furthermore, by drawing on real-world scenarios and applications, applied evaluation courses align with best practices in adult education, which emphasize a focus on collaboration, self-directed learning, and integration of past knowledge and new skills (Bakken et al., 2014; Johnson, 2017; Suiter et al., 2020). Reflecting on the value of applied evaluation experiences for graduate students, Gredler and Johnson (2001) concluded, "Perhaps most important is the nurturing of emerging professionalism. This benefit can only occur, we believe, through participation in an evaluation with real-world client concerns and time pressures, but that also provides faculty guidance and support" (p. 103). The bridging of application of skills and instructional support, such as occurs in applied evaluation courses, thus appears to be a pedagogically strong way to meet the need for more robust preservice education and socialization of evaluators.

The growing body of case descriptions of applied evaluation courses is invaluable to instructors seeking assistance in course design (Bakken et al., 2014; Davis, 2006; Gredler & Johnson, 2001; Suiter et al., 2016). However, less is known about student experiences within these courses. What aspects of these courses do students attribute to facilitating or detracting from their learning? Given that applied evaluation courses bridge formal and informal stages of graduate student development, what balance of structured and self-directed activities best serves the student's learning process? How do students experience the disequilibrium of applying their learning in real time, real life scenarios (Warner, 2020), and are there ways instructors can better scaffold their development through that process? This study begins to answer these questions through an analysis of 5 years of student evaluations from an applied evaluation course, as well as critical reflections from course instructors.

## Course Description

The course is taught as one of the core skills courses available to students in a community development master's program. The goal of the program is to prepare students for management- and leadership-level positions in local government and nonprofit organizations that work to effect change at the community level. In any given semester, roughly a third of the students in the evaluation course are enrolled in the community development program; the other students come from across the university and represent other master's and doctoral programs such as public health, international education, organizational leadership, child studies, and economic development. The diversity of perspectives, experiences, career goals, and research methods knowledge and preparation that students bring to the course is one of the greatest strengths of the course as well as being one of the aspects of the course that is most difficult to channel and manage. A shared interest among students, however, is interest in gaining skills and knowledge that are relevant to real-world problems, and are developed through real-world engagement. This course was developed to attend to those interests, as well as to provide a resource for the many local organizations that had identified program evaluation needs, but no budget or access to an evaluator to meet them. Thus, the overarching goals of the course are twofold: (1) developing evaluation knowledge, skills, and capacity in students who will (likely) eventually work in public service organizations and (2) developing evaluation knowledge and skills in local organizations that are interested in building their evaluation capacity.

Each semester, the lead author sends an email invitation to local community-based organizations through personal contacts and electronic mailing lists inviting organizations to apply to participate in the evaluation course for the upcoming semester. Applicants are asked to describe their or-

ganization's mission, describe their program evaluation need, indicate the person in the organization who will attend classes, and confirm that that person will be able to attend all courses of the semester. The lead author then selects five to seven organizations per semester to participate (the number of organizations is dictated by the number of students enrolled in the course at the time). Once organizations are selected, the students who will be taking the course are then invited to rank the organizations with which they would most like to work, and the instructor matches students with organizations to form project groups. Each project group consists of three to five students and one community partner. All students and community partners participate in weekly synchronous class sessions, where time is divided between instructor-led lecture and large-group discussion, followed by time for applied work in project-based teams. The instructor circulates and consults with all teams, and provides one in-depth consultation session to all teams toward the end of the term. The groups work together throughout the semester to develop an evaluation plan and accompanying tools (e.g., sampling plan, data collection tools and timelines, data management tools and/or strategies) that the community organization can implement after the course is over. The course has been taught using this format five times over the course of 5 years. During this time, the instructor has made some changes to assigned readings and updated lectures to provide updated examples and references; however, the core content, approach, and assignments have remained the same. The same lead instructor has taught the course each time, twice with a graduate student TA and three times without one.

A more thorough description of the course, as well as findings from a small qualitative study conducted with students and community participants the first semester the course was taught, can be found elsewhere (Suiter et al., 2016). Likewise, the outcomes of the course related to evaluation capacity-building in the participating community-based organizations are forthcoming (Suiter et al., 2020). The purpose of this article is to examine students' perceptions of their own learning and perspectives on the course.

## Methods

This course was taught in Fall 2014, Spring 2016, Spring 2018, Fall 2018, and Spring 2019. Data for this study were collected using the course evaluations that students complete at the end of each semester. Methodologically, course evaluations are acknowledged as a rich source for gaining insight regarding student experiences within a course, what they are learning, and how a course might be improved (Benton & Ryalls, 2016; Medina et al., 2019). In addition to course evaluation data, administrative data about students (e.g., number of students per semester, students' degree programs) were used to inform richer descriptions of the study participants.

### Study Participants

Students who participated in the program evaluation course came from 11 different programs across the university. The course is designed and offered as a master's level course, and so the vast majority of students, 109 (94%), were master's students, and seven (6%) students were doctoral students. The course enrollment target is 20 for each semester, and course enrollment was relatively stable over time (25, 22, 19, 27, 25). Due to the anonymity of course evaluations, there is no way to know if there were trends in responses based on a student's program, year in school, or other demographic factors. Response rates for course evaluations ranged from 52% to 68% in any given semester, which is consistent with the university's average.

### Data Collection

At the end of each semester, students are invited to complete course evaluations that contain a standard battery of quantitative and qualitative items. The evaluations are typically completed sometime within the last 2 weeks of class, are anonymous, and are released to the instructor of the course after all grades for the semester have been submitted. The university sends the link to complete course evaluations to students via their university email; however, the instructor informs students during class time that the course evaluations are coming, and encourages students to submit course evaluations. The instructor informs students that she reads and heeds course evaluations, both to improve her own teaching practice (Boysen, 2016), and to understand students' experiences with the course and how it might be improved, expanded, or replicated (Medina et al., 2019). We were

able to access course evaluation data from all five semesters of the course. During the 2017 school year, the university changed its course evaluation system and adopted slightly different question prompts, which had implications for our coding strategy. In Fall 2014 and Spring 2016, students were prompted to describe the weakest feature of the course, suggestions for improvement of the course, and suggestions for the instructor to improve their teaching style. In Spring 2018, Fall 2018, and Spring 2019, students were asked about the elements of the course that most contributed to their learning, improvements to the course they would recommend, and were given a space to include any additional comments about the course. Quantitative assessments of the course were available only after the 2017 change (i.e., for the Fall 2018, Spring 2018, and Spring 2019 semesters). The quantitative assessments of the course are summarized in the Findings section, but for the purposes of this study, we focused most of our analysis on the qualitative items, as they provide a richer source of data in terms of what students did and did not think was beneficial about the course, or what could be improved. Because course evaluations are submitted anonymously, use of these data for the purposes of this article was granted exempt status by the Vanderbilt Institutional Review Board.

**Analysis**

The first author compiled all available deidentified copies of course evaluations and shared them with the second author over a secure university server. The second author conducted a thematic analysis of the data (Braun & Clarke, 2012) by drawing on the course evaluation prompts to develop initial thematic domains. We then analyzed the data in NVivo (Version 12) to generate an iterative codebook by applying inductive codes to the data. The second author then discussed the inductive codes with the first author, who served as a critical friend (Kember et al., 1997). Specifically, the first author served as a sounding board for the second author's suggested coding scheme: listening, suggesting alternative explanations, and helping focus the analysis. For example, the analysis initially included suggestions from a few students who said they would have liked to implement an evaluation in the course rather than design one. Although this is an important consideration for course design and offerings more generally, it tells us less about the students' experiences within the course and was thus omitted as a code. After this discussion, the codebook was revised and organized into overall themes that fell within the domains constructed by the course evaluation prompts: course strengths and course suggestions. We organized codes within our thematic frame, with particular attention to the pedagogical links between many of the course strengths and weaknesses. This round of coding revealed three pedagogical and instructional practices from the curriculum that were present in both the course strengths and course suggestions domains: experiential education, interdisciplinary collaboration, and community partnerships. The second author then coded each instructional practice to capture students' self-reported affective responses to each central theme. Instances in which an instructional practice supported students' learning were coded as positive, and instances in which a practice hindered learning were coded as negative. This round of coding also revealed several nuanced responses in which students described struggling with a particular instructional practice while understanding the value of that practice to support their learning. These instances were coded separately. In reporting quotations in our findings below, we indicate the semester from which the data originated to demonstrate the reliance on multiple participants. Because the evaluations are anonymous, it is not possible to attribute the data to a particular student or student attributes, which is often done when reporting qualitative research. Instead, we report the respondent number and semester date by each quote (e.g., Respondent 5, Spring 2016) to demonstrate our efforts to include as many student voices as possible when reporting the data.

The authors' distinct relationships to the course reflexively sparked our interest in this inquiry and informed our investigation. As described by Etherington (2004), researcher reflexivity is "the capacity of the researcher to acknowledge how their own experiences and contexts (which might be fluid and changing) inform the process and outcomes of inquiry" (p. 19). The first author of this article was the course instructor, the second author was a student in the course, and the third author was the graduate teaching assistant for the course, who has since completed her degree and replicated this course at another university. These varied roles provided different entry points

into the focus of this study: understanding how challenge and ambiguity function in the context of evaluation learning.

## Findings

Student assessments of the course were generally positive (Table 1), indicating that course content and assignments aligned with the course goals, and that the course helped students appreciate the significance of program evaluation as an academic and professional field. Perhaps most importantly for the purposes of this study, 91% of students across all years indicated that the course helped them make connections between the context of the course and broader personal and professional goals.

Results from our qualitative analysis of course evaluation responses demonstrate the strengths and limitations of engaging graduate students and community partners in an applied, interdisciplinary program evaluation course. In what follows, we outline participants' perceptions of their learning experiences in this course. Interestingly, the three instructional and pedagogical approaches that students found most central to their learning—experiential education, interdisciplinary collaboration, and community partnerships—are the same learning contexts that students found most difficult to navigate. Table 2 outlines the frequency with which the 81 students in our sample mentioned experiential learning, interdisciplinary collaboration, and community partnerships in their course evaluations, along with their affective response to each theme. "Overall" codes indicate that the topic was mentioned at all (positively, negatively, or both) by students in the qualitative portion of the course evaluations. "Positive" codes indicate that the student identified that particular theme as supportive of their learning, and "negative" codes indicate that the student identified the theme as something that caused discomfort or frustration. "Both" codes indicate instances in which the student offered, within the same comment, more nuanced interpretations of how these

**Table 1. Response Rates and Select Quantitative Items From Course Evaluations 2014–2019**

| Course evaluation domain | Fall 2014 | Spring 2016 | Spring 2018 | Fall 2018 | Spring 2019 | All years |
|---|---|---|---|---|---|---|
| Number of students enrolled in course | 25 | 21 | 27 | 19 | 25 | 117 |
| Course evaluation response rate | 75% | 85% | 52% | 68% | 64% | 69% |
| % responding "Agree" or "Strongly Agree" to the following prompts:[1] | | | | | | |
| This course helped me appreciate the significance of the subject matter. | Not included in course evaluations | | 86% | 92% | 94% | 91% |
| The components of this course, such as activities, assessments, and assignments, were consistent with the course goals. | Not included in course evaluations | | 100% | 92% | 100% | 98% |
| This course helped me consider connections between course material and other areas of my personal, academic, or professional life. | Not included in course evaluations | | 93% | 92% | 88% | 91% |
| % responding "Very Good" or "Excellent" to the following prompt:[2] | | | | | | |
| Overall, the course was: | Not included in course evaluations | | 78% | 69% | 81% | 77% |

*Note*. [1] Response options included Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree.
[2] Response options included Poor, Fair, Good, Very Good, Excellent.

**Table 2. Frequency of Assigned Coding Across Themes**

| Theme | Assigned coding | Frequency % of all responses | Frequency % within theme |
|---|---|---|---|
| Experiential education | Overall | 48% | 100% |
| | Positive | 30% | 63% |
| | Negative | 14% | 29% |
| | Both | 4% | 11% |
| Interdisciplinary collaboration | Overall | 41% | 100% |
| | Positive | 17% | 42% |
| | Negative | 17% | 42% |
| | Both | 6% | 16% |
| Community partnerships | Overall | 68% | 100% |
| | Positive | 46% | 67% |
| | Negative | 15% | 22% |
| | Both | 7% | 11% |

practices served to complicate *and* enhance their learning in the course.

### Experiential Education

Many students noted that the experiential nature of the course had important implications for their learning. They noted the importance of "being able to practice the skills we learned about" (Respondent 1, Spring 2019), "complete work that [was] in line with program evaluation" (Respondent 3, Fall 2018), and "apply the skills we were learning" (Respondent 3, Spring 2016) in grasping the course concepts. Much of the applied nature of the course was rooted in generating a multistep program evaluation plan that groups constructed with their community partner over the course of the semester. This project served an instructional purpose through offering students scaffolded practice in developing technical evaluation skills and resulted in a comprehensive evaluation plan that could be implemented by the community partner. One student reflected on how this project shaped her group's learning by noting that it "helped [us] review course materials and understand what concepts are most important and applicable" (Respondent 3, Spring 2019). This was particularly salient for students who expressed an interest in continuing to work with nonprofits to carry out the evaluations that they had designed, as the course served as a catalyst for meaningful partnership development.

Despite a consensus among students that this experiential approach to evaluator training was unique and generative, students often admitted missing the "organization and structure" (Respondent 2, Fall 2018) of traditional graduate classes. As they were in an applied, community-based course, students were tasked with responsibly addressing a range of difficult practical problems while gaining basic technical competencies. As students learned that evaluators are often tasked with navigating complex institutional structures, defining and measuring change, and balancing methodological rigor with organizational capacity, they were met with many of the same real-word challenges and tasked with "critically applying new skills as they were being taught" (Respondent 2, Fall 2014). The approach to evaluation training leveraged in this course included several factors that were challenging to students, including a fixed time span, limited instructor-derived requirements, and complex group dynamics.

In their course evaluations, some students offered suggestions to mitigate the ambiguity inherent in this learning experience. Suggestions referenced the pace of the

course, noting that they "move very quickly through the semester," which made it difficult to "dig deep" into the course concepts (Respondent 1, Spring 2018). Some asked for a more hands-on pedagogical approach from the instructor, as they missed the uniform "course design and guidance on the deliverables" (Respondent 11, Spring 2019) that are offered in traditional coursework. One student suggested scaling back the experiential nature of the class in order to facilitate time for concept mastery before groups applied their learning to their evaluation plan through having future participants "learn the necessary skills/tools during the first few weeks and get practice applying with scenarios, [and] after receiving the basic skills then allowing clients to come in and apply with more feedback and oversight from [the] instructor" (Respondent 6, Spring 2019). Similarly, one student shared a belief that the pacing negatively impacted their progress in the course:

> Towards the end of the course, it felt a bit rushed to develop a data management and analysis plan. It was also difficult to develop qualitative and quantitative evaluation instruments in one course period. Having more time to work on these sections would help in more thoroughly understanding the material. (Respondent 10, Spring 2018)

Other students shared the sentiment that the course felt "rushed" but attributed this feeling to their own practitioner identity development and the learning curve they were faced with in the beginning of the course. They offered process-oriented critiques of their own evaluation practice, noting that they would have benefited from "starting to think earlier about how all of the pieces for the final project will fit together, having our one-on-one group meetings with [the instructor] much closer to the final project date so we can ask for help" (Respondent 4, Fall 2018). Overall, within these critiques of the complexities of experiential learning, students demonstrated a nuanced understanding of what is and is not possible within a single program evaluation training course, as "the nature of program evaluation makes it difficult to cover all the bases in one semester" (Respondent 5, Spring 2016).

**Interdisciplinary Collaboration**

Disciplinary and programmatic diversity in the course was an often-cited contributor to students' perceived growth. Many students noted that they benefited from working with others from across the university to create an evaluation plan that would benefit their community partner. Students appreciated having a portion of each 3-hour class period devoted exclusively to carrying out group work with their community partner present, noting that "allowance of class time made the work load more manageable" (Respondent 5, Spring 2019) and gave groups space to "actively consider the topic in relation to the [partnering] programs" (Respondent 10, Spring 2016), as content covered in each class session was immediately applied to the community-based organization being evaluated. For example, when the topic of logic modeling was introduced in class, students spent their group time in "collaboration with our organization's representatives" (Respondent 13, Fall 2018) to translate insights about their organization into a visual model of how that group functions to achieve their goals. Many students commented on the value of bringing multiple perspectives to bear to support their community partner. One student described the richness of their collective experience in this way:

> The group setting of the class is very helpful to learning. Though we are situated within the larger class, it is extremely helpful to work on assignments and work through things that may have been misunderstood or perspectives that may not have been considered. (Respondent 3, Spring 2019)

This opportunity to bring a range of perspectives into an evaluation plan speaks to a push in the field to promote interpersonal skills and competencies within evaluation training, as employers increasingly seek out candidates with a commitment to and confidence in collaboration.

Despite the demonstrated benefits of interdisciplinary learning, course evaluations revealed a range of tensions that arose within groups. Much of the intragroup conflict was rooted in role clarity and varying levels of experience among group members. Interestingly, students with more and less evaluation experience both spoke to this tension as a barrier to their development.

For example, one novice evaluator shared that they felt "dragged through certain things by more knowledgeable students in [their] group" (Respondent 12, Spring 2018). In their course evaluation, they described the perceived implications of being a relative novice in the group on their development as an evaluation practitioner:

> I feel like I could participate in a program evaluation moving forward, but am not sure if I could successfully lead one in an organization because there were tasks that were given to people who were already highly skilled in certain areas, rather than allowing for growth of other members in our group due to time constraints. I still feel like what I learned was a huge asset to my skills, but wonder if this could be addressed on the front-end by the instructor when groups begin developing group norms the first week. (Respondent 12, Spring 2018)

This student struggled to apply skills they were developing in the course, allowing more practiced group members to take the lead at the expense of their own experiential education. Given that this was an introductory course, some professional students with previous evaluation and research experience struggled to remain engaged in technical training instruction, noting frustration with members of their group "not understanding basic concepts (how to write, basic logic, research methods)" (Respondent 2, Spring 2018). One student described their perception of the impact that working in a group with differing levels of experience had on quality of their final evaluation plan:

> It was always a struggle to wait for everyone to figure out the basic principles behind an assignment. We never got work done in class, and the outside work individuals did was sloppy. I am really embarrassed by our work. That being said, this has been a huge learning experience for me in terms of how I work with groups. This is my first time having a negative working experience, and I think having it will likely make me a better practitioner (and evaluator). (Respondent 2, Spring 2018)

Although students who shared their frustrations with their group members candidly in their course evaluations represent a small minority of course participants, their contributions illuminate an important instructional challenge inherent in group work in general, and interdisciplinary coursework in particular.

**Community Partnerships**

Across semesters, the aspect of the course that students most often cited as the primary contributor to their learning was the opportunity for close collaboration with community partners, as it "made the work [they] did feel more significant and relevant than work in other classes" (Respondent 8, Spring 2016). Creating a program evaluation effectively "turned classwork into real world work" (Respondent 6, Spring 2016) and offered students a space to enact an evaluator identity as they worked to deliver an actionable product to their community partner. Students adjusted their level of engagement in the project accordingly, noting that "doing an actual program evaluation for an actual community partner gave [them] a level of commitment and buy-in [they] might not otherwise have had for just a grade" (Respondent 10, Spring 2016).

Through this university–community partnership, students found that "having the community partners in the room was exceptionally helpful for building a collaborative relationship" (Respondent 7, Spring 2016). Having community partners present provided for constant member-checking of students' evaluation plans, as community partners could speak to their organizations' goals, values, and capacity. Students appreciated that by "working with local organizations to develop tools that will be useful to them" (Respondent 7, Spring 2018) both parties benefited, noting that the "course work [was] extremely practical and beneficial, not only for students but (from what I gathered) for the community partners as well" (Respondent 5, Spring 2016). Students viewed these partnerships as a way to build skills beyond those covered in the syllabus, including a deeper understanding of "nonprofit organizations and development thereof" (Respondent 8, Spring 2018). They also saw these partnerships as spaces for networking and career development through "build[ing] a strong relationship with our community partner, which has lent itself to future opportunities" (Respondent 5, Spring 2019).

Despite largely positive experiences between

groups and their community partners, a small subset of students reported barriers to success that stemmed from limitations of the community-based organization or the community partner. For example, a few students noted programmatic limitations that made it difficult to make a complete evaluation plan. Critiques of this nature were rooted in the perceived absence of an "existing, robust program to evaluate" (Respondent 4, Spring 2018). This variability extended to the readiness of the participating community partner, who was sent from the community-based organization to represent its evaluation needs and capacity. For example, a student observed that the relative level of readiness among partners contributed to "a lot of variation in the workload between groups—some community partners had clear goals and brought in materials and some were quite vague with no materials" (Respondent 10, Spring 2019). Additionally, one student shared that "relying on a community stakeholder for information and guidance was often difficult" (Respondent 10, Spring 2016), and another felt that their partner "was not very organized or prepared" (Respondent 7, Spring 2016). We noted one critique that addressed an important limitation in developing generative community partnerships: organizational power. Although organizations were asked to nominate staff in leadership positions to participate in the class, this was not always a possibility. In those instances, students sometimes questioned the ability of the community partner to guide the development of an evaluation plan that was aligned with organizational priorities, and thus likely to be implemented by organizational leadership. As one student noted, without "enough organizational authority to answer questions on what they were looking for" (Respondent 10, Spring 2019), it was difficult for community partners to communicate the organization's needs and goals to their team of evaluators.

Finally, as groups spend most of their time addressing the distinct needs of their community partner, several students described "feeling siloed in [their] group" (Respondent 3, Spring 2016) and desired more opportunities to "share across group projects during the semester" (Respondent 10, Fall 2018). One student shared the way that "limited interaction with students outside of [their] group" (Respondent 4, Spring 2016) impacted their ability to provide a high-quality evaluation plan to their community partner:

> Not having an opportunity to catch up with the other groups about their own program evaluation led to relatively minor weaknesses . . . it would be nice to hear about how [other groups are] navigating hurdles or coming up with great ideas. (Respondent 7, Spring 2016)

Although organizational siloing can provide direction and keep professionals focused on addressing one task well, students' course evaluations pointed to the ways that siloing and infrequent interaction among the different project teams may have inhibited creative problem-solving and innovation.

## Discussion

We began this article discussing the importance of evaluation training opportunities that combine the hard and soft skills of evaluation (Dewey et al., 2008; Galport & Azzam, 2017; King et al., 2001; Russ-Eft et al., 2008) and engage students across a range of disciplines (Bakken et al., 2014; Davis, 2006). This course provides one such opportunity, and we used students' course evaluation data to investigate their experiences, as well as the short-term effects of the course on their learning and professional development. The course encouraged students to straddle the formal and informal stages of professional development (Weidman et al., 2001), and—as noted in scholarship on experiential learning in graduate education—emerging professionalism produced discomfort along with feelings of accomplishment and success (O'Meara, 2008). Part of what makes real-world scenarios so generative for teaching and learning is their complexity, which results in a tension between giving students what they want (e.g., order, sense of control, linear progression) and what may ultimately help better prepare them for practice (e.g., ambiguity, trial and error, immersive practice; Warner, 2020). This course supports evaluation knowledge and skill building among graduate students with diverse professional interests while supporting evaluation capacity in local community-based organizations. The course functions as both a traditional course and a practice space. Although many markers of the traditional classroom remain (e.g., course texts, summative assessments), the inclusion of community partners required students to enact an evaluator identity and

create a work product with real-world implications. In this hybrid space, there was an expectation that learning would involve struggle and would at times be uncomfortable and messy. Ultimately, students found that learning is often animated by this difficulty.

Although the ambiguity that students noted in their course evaluations could be mitigated by additional scaffolding associated with the traditional classroom, we believe that scaling back the experiential and emergent nature of the course would disrupt students' ability to enact a learner identity and an evaluator identity iteratively as the course progresses. Additionally, we maintain that the tension students describe regarding their own developing expertise and finite resources is salient in real-world evaluation practice, and is therefore generative for students of evaluation to experience within instructional settings.

Many of the ambiguities students detailed in their course evaluations are consistent with authentic learning and the development of an evaluation practitioner identity (Brown, 1985). As students gained creative confidence, they experienced a shift in their perception of learning in the course. For example, some students suggested a standardized, instructor-driven team-building process early in the course to ensure that all students are able to "lead [an evaluation] in an organization" (Respondent 12, Spring 2018). However, these same students' evaluations reflect that they ultimately understood the course model to provide an important cross-training opportunity that allowed students to engage with experienced others in structured practice as a modality for socialization (LaVelle & Donaldson, 2010). In this way, the course offered a unique evaluator socialization opportunity and had real implications for practice, which ultimately shifted students' orientation to the work.

### Limitations

An important limitation of this study derives from the fact that course evaluation data is collected at the end of the semester, immediately after students have participated in the evaluation class. Consequently, we were not able to test for things such as the durability of the hard and soft skills students learned in the course, or if students made sense of the experience differently after having entered the workplace. Although tracking students after graduation is difficult, subsequent studies could contact students once they are 1 or 2 years out of their respective programs to investigate such questions as whether they use evaluation in their current job, what contributions they believe the course made to their evaluation practice, and whether their perspectives on the nature of the course—especially aspects of the course that students found challenging—had changed over time.

An additional limitation comes from the response rates to the course evaluations. Although they were on par with university averages and represent acceptable response rates in general for survey research, it is nevertheless true that 31% of students who took the course did not provide feedback. We therefore lack the perspectives of those students, who might or might not have had similar experiences and opinions regarding the course. Finally, there is a question of whether students are prepared, immediately following a course, to assess its value (Deslauriers et al., 2019). Certainly, the advantages and disadvantages of various content and forms of learning may emerge or become visible to students only long after the course is complete (Boud & Falchikov, 2006; Deslauriers et al., 2019). Despite this potential limitation, we trust students' ability to assess their learning and experiences, and regard them as an essential source of information in evaluating the pedagogical value of the course (Benton & Ryalls, 2016; Medina et al., 2019).

## Conclusion and Implications for Teaching Evaluation

In this study, we used student course evaluation data to assess student experiences with an applied program evaluation course and further the scholarship of teaching evaluation, especially teaching evaluation through community-engaged and project-based learning formats. In our own teaching practice, we also plan to use student course evaluation data to inform improvements in the course. We share these implications for our own practice as an entry point into suggesting implications for teaching evaluation more generally.

First, we believe that forecasting the tension that students might experience, as well as providing more transparency about why and how the course is designed, could help students understand the structure and activities of the course. For example, telling students

from the beginning of the course that the necessary ambiguity of real-world practice was likely to produce some anxiety and/or discomfort, and that educational research as well as our data show that such discomfort can be productive, might help to lessen the anxiety students feel. Second, the feedback on group process has made us realize that guiding students in establishing good group process and norms at the beginning of the course—and continuing to check in with students about it throughout the course—is an essential aspect of facilitating group learning. We are exploring literature on the design of interdisciplinary teams in order to ensure we are providing optimal conditions for group functioning (Choi & Pak, 2007). Third, student feedback as well as our own reflections on the fitness of various community partners has given us a better sense of the types of community partners who are best positioned to contribute to and benefit from this course. Specifically, the best community partners are those who can send organizational leaders with enough power to guide and implement an evaluation, those with a clearly articulated program evaluation need, and those who are able to explain the work of their organization in a straightforward manner.

This study allows us to better understand student experiences and perspectives on participation in a community-engaged, project-based evaluation course. The qualitative research strategy allowed us to observe and relay the complicated and sometimes contradictory nature of student experiences—specifically that the best and worst parts of the course were often the same things. Importantly, these "contradictions" were not only experienced between students (meaning, some students liked aspects of the course that others did not), but also experienced within individual students throughout the course of the class. Ultimately, given just enough support and scaffolding, students' struggles animate their learning.

## Author Note

Article correspondence should be directed to Sarah V. Suiter, Vanderbilt University, sarah.v.suiter@vanderbilt.edu

## Conflict of Interest

We have no conflict of interest to disclose.

## Funding

The project described was supported by CTSA award No. UL1TR000445 from the National Center for Advancing Translational Sciences. Its contents are solely the responsibility of the authors and do not necessarily represent official views of the National Center for Advancing Translational Sciences or the National Institutes of Health.

## About the Authors

***Sarah V. Suiter*** *is an associate professor of the practice in the Department of Human & Organizational Development at Vanderbilt University.*

***Kathryn Y. Morgan*** *is an assistant professor in the Department of Psychology at The University of the South–Sewanee.*

***Amie Thurber*** *is an assistant professor in the School of Social Work at Portland State University.*

## References


Bakken, L., Núñez, J., & Couture, C. (2014). A course model for building evaluation capacity through a university–community partnership. *American Journal of Evaluation, 35*(4), 579–593. https://doi.org/10.1177/1098214014523671

Benton, S. L., & Ryalls, K. R. (2016). *Challenging misconceptions about student ratings of instruction* (IDEA Paper No. 58). https://www.ideaedu.org/Portals/0/Uploads/Documents/Challenging_Misconceptions_About_Student_Ratings_of_Instruction.pdf

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399–413. https://doi.org/10.1080/02602930600679050

Boysen, G. (2016). Using student evaluations to improve teaching: Evidence-based recommendations. *Scholarship of Teaching and Learning in Psychology, 2*(4), 273–284. https://doi.org/10.1037/stl0000069

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology: Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Brown, R. D. (1985). Supervising evaluation practicum and intern students: A developmental model. *Educational Evaluation and Policy Analysis, 7*(2), 161–167. https://doi.org/10.2307/1163763

Choi, B. C., & Pak, A. W. (2007). Multidisciplinarity, interdisciplinarity, and transdisciplinarity in health research, services, education and policy: Promotors, barriers, and strategies of enhancement. *Clinical and Investigative Medicine, 30*(6), 224–232. https://doi.org/10.25011/cim.v30i6.2950

Christie, C. A., Quiñones, P., & Fierro, L. (2014). Informing the discussion on evaluator training: A look at evaluators' course taking and professional practice. *American Journal of Evaluation, 35*(2), 274–290. https://doi.org/10.1177/1098214013503697

Davis, M. (2006). Teaching practical public health evaluation methods. *American Journal of Evaluation, 27*(2), 247–256. https://doi.org/10.1177/0198214006286422

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences 116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Dewey, J. D., Montrosse, B. E., Schröter, D. C., Sullins, C. D., & Mattox, J. R., II (2008). Evaluator competencies: What's taught and what's sought. *American Journal of Evaluation, 29*(3), 268–287. https://doi.org/10.1177/1098214008321152

Doberneck, D. M., Glass, C. R., & Schweitzer, J. H. (2010). From rhetoric to reality: A typology of publicly engaged scholarship. *Journal of Higher Education Outreach and Engagement, 14*(4), 5–35. https://openjournals.libs.uga.edu/jheoe/article/view/794

Etherington, K. (2004). *Becoming reflexive researchers: Using ourselves in research*. Jessica Kingsley Publishers.

Galport, N., & Azzam, T. (2017). Evaluator training needs and competencies: A gap analysis. *American Journal of Evaluation, 38*(1), 80–100. https://doi.org/10.1177/1098214016643183

Gredler, M., & Johnson, R. (2001). Lessons learned from the directed evaluation experience. *American Journal of Evaluation, 22*(1), 97–104. https://doi.org/10.1177/109821400102200110

Johnson, S. M. (2017). *Teaching adult undergraduate students*. https://cft.vanderbilt.edu/teaching-adult-undergraduate-students/

Kember, D., Ha, T. S., Lam, B. H., Lee, A., NG, S., Yan, L., & Yum, J. C. (1997). The diverse role of the critical friend in supporting educational action research projects. *Educational Action Research, 5*(3), 463–481. https://doi.org/10.1080/09650799700200036

King, J. A., Stevahn, L., Ghere, G., & Minnema, J. (2001). Toward a taxonomy of essential evaluator competencies. *American Journal of Evaluation, 22*(2), 229–247. https://doi.org/10.1177/109821400102200206

LaVelle, J. M., & Donaldson, S. (2010). University-based evaluation training programs in the United States 1980–2008: An empirical examination. *American Journal of Evaluation, 31*(1), 9–23. https://doi.org/10.1177/1098214009356022

LaVelle, J. M., & Donaldson, S. (2015). The state of preparing evaluators. *New Directions for Evaluation, 2015*(145), 39–52. https://doi.org/10.1002/ev.20110

Medina, M. S., Smith, W. T., Kolluru, S., Sheaffer, E. A., & DiVall, M. (2019). A review of strategies for designing, administering, and using student ratings of instruction. *American Journal of Pharmaceutical Education, 83*(5), Article 7177. https://doi.org/10.5688/ajpe7177

O'Meara, K. (2008). Graduate education and community engagement. *New Directions for Teaching and Learning, 113*, 27–42. https://doi.org/10.1002/tl.306

Russ-Eft, D., Boeber, M. J., de la Teja, I., Foxon, M. J., & Koszalka, T. A. (2008). *Evaluator competencies: Standards for the practice of evaluation in organizations*. Jossey-Bass.

Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage.

Suiter, S., Morgan, K, & Eccleston, S. (2020). *Evaluation capacity building through community–university partnership* [Unpublished manuscript].

Suiter, S., Thurber, A., & Sullivan, C. (2016). A co-learning model for community-engaged program evaluation. *Progress in Community Health Partnerships: Research, Education, and Action, 10*(4), 551–558. https://doi.org/10.1353/cpr.2016.0063

Trevisan, M. S. (2004). Practical training in evaluation: A review of the literature. *American Journal of Evaluation, 25*(2), 255–272. https://doi.org/10.1177/109821400402500212

Warner, L. H. (2020). Developing interpersonal skills of evaluators: A service-learning approach. *American Journal of Evaluation, 41*(3), 432–451. https://doi.org/10.1177/1098214019886064

Weidman, J. C., Twale, D. J., & Stein, E. L. (2001). *Socialization of graduate and professional students in higher education: A perilous passage?* (ASHE-ERIC Higher Education Report Vol. 28, No. 3; Jossey-Bass Higher and Adult Education Series; ED457710). Jossey-Bass.