

WPAs Reading SETs: Toward an Ethical and Effective Use of Teaching Evaluations

Courtney Adams Wooten, Brian Ray, and Jacob Babb

ABSTRACT

This article seeks to open a conversation about student evaluations of teaching (SET). Student evaluations have recently fueled much public discussion, given their role in advancement—as well as the risks they pose for contingent faculty. While SETs have historically been of interest to Rhetoric and Composition researchers, few articles in our field have engaged with their implications for writing instruction and administration. Our article uses a survey method approach to restart this conversation, reporting the results of a national survey of WPAs and identifying trends in how they perceive and use SETs at their institutions. We conclude with recommendations about the design and implementation of SETs in writing programs.

There is perhaps no topic in writing program administration more important and less studied than student evaluations of teaching (SETs). Debated fervently in professional publications such as *The Chronicle of Higher Education* and *Inside Higher Ed*, complained about after department meetings, and even lampooned in *The Onion*, student evaluations of teaching play a major role in personnel decisions made by WPAs, chairs, and deans. One of the most common methods for assessing instructors (Basow; Stark-Wroblewski et al.), SETs often carry greater weight in personnel decisions than observations or teaching portfolios. Meanwhile, writing teachers and WPAs seem to take institutionalized SET questions as unavoidable and unchangeable, or they have developed their own localized SET questions and procedures without engaging in a wider disciplinary conversation with colleagues about the design and structure of more effective evaluations. If they have, these discussions are not reflected in our discipline's publications. This article seeks to open a wider disciplinary conversation about SETs by reporting

on a national survey regarding the implementation of SETs by WPAs and concluding with recommendations about their design and implementation in writing programs.

While much research exists on SETs in higher education journals, our discipline has produced little scholarship on them since the 1960s. One major exception is Amy Dayton's collection *Assessing the Teaching of Writing*, in which WPAs address theories, histories, and best practices in a range of teaching evaluation instruments. Dayton overviews some current scholarship on SETs and concludes with recommendations that WPAs use them for formative rather than summative purposes, avoid norm referencing, and include them as one measure alongside teaching portfolios, reflective letters, and observations (41). Given the absence of research on SETs in our discipline, we designed a study drawing from survey results to determine the following: What kinds of SETs do WPAs use in their programs? How do they use them? Are they satisfied with the current systems? Do they envision the potential for change? Ultimately, our survey results have indicated WPAs do the best they can—unhappy with the design of their institutional SET forms, reluctant to propose their own, and yet attempting to use insights from them, as well as from observations and portfolios, to engage their faculty in mentorship, training, and professional development.

METHODOLOGY

We designed our survey using Qualtrics during the spring of 2014, consulting prior survey studies by Jill Gladstein; E. Shelley Reid and Heidi Estrem; and Chris Thaiss and Tara Porter. We reviewed early versions of the survey with senior WPAs at three different institutions, who helped us refine our survey before we posted the final version on the WPA-Listserv that summer. We then generated a list of several hundred institutions from sources such as the *U.S. News and World Report's* ranking of colleges and universities, state government websites, and online searches for accredited institutions in every U.S. state. From this master list, we then gathered contact information for the WPA or chair of the appropriate department at each institution. Our search for contact information resulted in a final list of 274 WPAs or department chairs to whom we emailed a detailed recruitment narrative outlining our intentions. We attempted to construct a representative sample of institutions from each state that included two-year colleges; private institutions (both secular and religious); minority-serving colleges and universities; and public universities, including regional, comprehensive, and flagship institutions. Our goal was to solicit responses from institutions

that reflect the numerous missions and varied populations of American higher education.

Our survey (see Appendix A) consisted of 28 questions in four general categories:

1. questions seeking demographic data and basic information about the design of institutional SETs (1–10),
2. focused but objective questions on how SETs are used at participants' institutions (11–16),
3. Likert scale and multiple-choice questions asking WPAs' perceptions and use of SETs in their programs (17–24), and
4. short answer questions inviting participants to elaborate on their perceptions (18–20, 24).

Our first set of questions sought data on fundamental issues regarding SETs such as how long they had been in use at a given institution, whether they were mandatory, if evaluations were completed online or on paper, and whether a program used separate forms for first-year writing courses. The second set of questions focused on key personnel who read SETs and use them to evaluate teacher performance—and specifically what role WPAs have in this process. The third and fourth sets of questions invited WPAs to assess and reflect on the effectiveness of SETs in measuring and improving instruction, as well as their impact on teachers in different ranks and positions.

We received 62 responses to the survey from June to September, yielding a response rate of 23 percent when calculated against our pool of 272 institutions (given the overlap of our direct email list with members of WPA-L, it is difficult to report a more exact response rate). Despite the admittedly small sample, we found value in the responses we received. Of those respondents, 89 percent identified as current WPAs, although some respondents clarified later that department chairs served as WPAs at their institutions. Our results represent an overwhelming response from four-year institutions (90%), 42 of which offer some level of graduate education, and a smaller response (10%) from faculty at two-year colleges. The skew toward four-year institutions may be due to a number of factors: some institutions lack a clear administrator in charge of first-year writing courses; some WPAs may have too little time to respond to every survey request they receive; others may not identify strongly enough with the role of WPA or organizations like CWPA to see our project as relevant to their immediate concerns. Certainly, determining the WPA or equivalent contact at institutions other than four-year colleges and universities was challenging because

websites frequently only listed a department chair. It is possible that these institutions simply did not have an official WPA or someone who identifies as a WPA.

We found a number of trends that we summarize here and then explain more fully with supporting data in the next several sections. First, WPAs must negotiate authority over personnel decisions based on SETs. Often, they must share responsibility with other faculty or defer to department heads and committees. Second, WPAs often find themselves primarily fulfilling the role of mentor or coach for instructors determined by others (i.e., deans, chairs) to have received poor teaching evaluations. Third, WPAs are keenly aware of the limitations of SETs, although they report that acontextual data focused on raw scores compared against department averages is the norm. Fourth, almost half of the WPAs surveyed do not have direct access to students' narrative comments on SETs—a source of data seen as crucial for interpreting numerical scores. Finally, WPAs express a need to counter the largely punitive and biased instrument with other measures of teaching effectiveness.

NEGOTIATING AUTHORITY BETWEEN WPAs AND OTHER FACULTY IN SET USE

Agency and authority over SETs figured as a prominent theme in survey responses, especially when WPAs described themselves as having minimal or incomplete control over personnel decisions, including mentoring and hiring or firing. In some cases, WPAs only made suggestions to department chairs or committees on hiring and renewal. Generally, however, WPAs have access to SET data, even if personnel decisions are the purview of other administrators. Survey results showed that 76% are able to read SET responses, but 78% stated that department chairs also read the SETs. The 78% figure seems to be a clear indication that many WPAs lack clear autonomy over personnel decisions that may derive from course evaluations. We became especially interested in expressions of tension in responses to one survey question that asked WPAs to “elaborate on decisions or actions the WPA at your institution takes to address what are seen as negative/concerning course evaluations for an instructor.” Of the 41 responses, 10 explicitly referenced department chairs while another 4 described WPAs having limited authority. One response stated that the writing program is housed in a first-year experience program and that the director of the FYE program—not the WPA—handles potential problems that arise in SETs. So, on the one hand, WPAs are accountable for student success in first-year writing courses. They are the ones directly charged with training and communi-

cating with the instructors who teach these courses. When programs fall short of expectations, WPAs bear the most responsibility. However, when it comes to the most important decisions about who works for them, it appears that WPAs often have a disproportionately small level of oversight.

Respondents explained varying levels of authority concerning SET results. One respondent simply states, “Lecturers and tenure-line faculty course evaluations are mostly dealt with at the department committee level” in contrast to adjuncts, whose evaluations are read by first-year composition coordinators. Another outlines the limits of their authority: “We do have TT faculty who teach first-year writing, but I (the WPA) have no control over their curriculum, no access to their evaluations, and no authority to change anything about their course assignments.” These comments are consistent with the WPA role at many institutions—limited or no authority over other tenure-track faculty.

The second theme, regarding power differentials between department chairs and WPAs, splits along two lines: mentoring abilities and hiring/firing abilities. Several WPAs identify their central mode of intervention as mentoring in contrast with hiring/firing, a power often possessed solely by the chair, regardless of rank or position. One respondent states, “At our institution the head rather than the WPA makes decisions regarding appointments and course assignments. The writing coordinator is in a coaching position.” Another responds, “The WPA role is one of mentorship and support. The assessments are formative in nature, not evaluative. Evaluation is at the discretion of the head with input and advice from the WPA.” A third responded, “The WPA may provide professional development for writing instruction, but the WPA has absolutely no authority over who teaches writing intensive courses or how they are taught.” In these responses, mentoring is coded as *coaching*, *mentorship and support*, and *professional development*.

Even when WPAs are in charge of mentoring, they sometimes do not decide who to mentor or why. In response to “How does your institution distinguish between good, average, and concerning course evaluations?” one respondent answered, “scores in the low 3’s or worse will cause the Dean to ask the Chair to ask the WPA to contact you for mentoring.” The dean determines who is mentored even though it is ultimately the WPA who carries out this process. Taking such decisions out of WPAs’ hands tends to enforce the idea of SETs as inevitable measures of teaching effectiveness that cannot be contested. In such a scenario, the tacit assumption is that the WPA is not capable of interpreting SET data, only carrying out the will of the university. Sometimes, the WPA has more authority to enact sanctions on NTT faculty than on TT faculty, although lines of authority

can be muddled. One respondent stated, “With TAs, direct action; with others, chair’s responsibility.” Another distinguishes between GTAs, whom the WPA can choose to mentor and potentially dismiss from employment, and TT faculty, over whom the WPA has no control. Finally, one respondent answers,

Adjunct, FT/NTT, or TA faculty who earn low scores will prompt the chair to email the WPA to suggest further mentoring. Hiring/firing are the purview of salary committee and adjunct committee, for whom course evaluations are one of multiple factors.

In such cases, the WPA must defer to multiple stakeholders when deciding how to act on SET data and can take, at best, an indirect stance on such information—even when they are perhaps interacting the most with the target instructors.

A few responses indicated more collaborative relationships between these figures. As one respondent indicated, the “WPA collaborates with chair to adjust teaching load, assignments, and mentoring.” Another echoed the positive relationship between WPA and chair: “In conjunction with the chair of dept and chair of the dept personnel committee, the WPA identifies instructors who seem to be struggling and creates a plan for support and ‘intervention.’” A third replied, “The WPA works closely with the chair to identify concerns among our composition faculty and to establish systems for both formative/constructive feedback/mentoring and increased oversight/clear benchmarks for improvement and follow-through for contract renewal.” In summary, a recurring situation is that WPAs usually negotiate or share interpretation and decision-making processes with a department chair or committee regarding SETs for all faculty who teach first-year courses. One of the survey respondents directly addressed this issue in his or her closing comment:

I would also like to see all faculty, regardless of employment status have to administer evaluations [sic] every semester. Full-time faculty teach about 50% of our classes. Without the input of students in their courses, we have an incomplete picture of our students’ experiences in FYW classes on our campus.

Clearly, struggles over departmental governance and authority affect not just our individual work lives but our writing programs. In order to be effective, WPAs must have decision-making power. They are the most qualified to make programmatic decisions. However, our survey results show that, at least when it comes to SETs, WPAs often have less autonomy than they desire. We find this trend concerning because course evaluation forms are an important tool for assessing a program’s strengths and weak-

nesses, and the data they yield has an undeniable impact by determining who teaches what and how.

WPAs' ATTITUDES TOWARD SETs

Our survey revealed a widespread dissatisfaction among WPAs toward SETs. Despite such frustration, though, they feel constrained by their institutions in several ways that we explore in this section. For instance, our study found that 49% of WPAs “do not have plans to offer a separate [writing program] form,” and an additional 25% “do not have a separate evaluation form, but would like to in the future,” while only 13% said they have a separate form for their first-year writing courses. Finally, 13% said they use a separate form “in addition to the department and/or university’s form.” The 74% of respondents who only have access to results on generic SETs seem at a disadvantage. While the generic forms may ensure reliability across programs and departments, they do so at the expensive of validity. The types of general questions asked by generic forms address vague notions of teaching effectiveness, irrespective of what many WPAs are likely to identify as disciplinary best practices—use of peer review, introduction to research strategies and library resources, emphasis on higher-order concerns in papers. Therefore, the generic forms do not provide the most accurate information about a teacher’s ability to help students learn to write.

As respondents reported, the most common questions on university SET forms instead address “Clarity of expectations/assignments” (80%), “Preparedness” (74%), “Instructor’s level of knowledge” (69%), “Clear presentation of material” (63%), and “Quality and/or relevance of course materials” (61%). All of these questions relate to students’ perceptions of their instructors and their instructors’ handling of course materials. However, these questions are at best indirect measures of teacher effectiveness. At worst, they allow students to deflect responsibility for their learning onto their instructors. Questions directed at writing-specific instructional practices typically appear in writing program specific SETs rather than these general SETs. Most WPAs encounter information from generalized SETs that does not provide a clear view of how instructors teach writing. In some cases, student responses to such forms may obfuscate or misrepresent teacher effectiveness, given that students may not immediately appreciate the value of our best practices or desire instructional methods we train teachers to avoid—lecture-based teaching, feedback emphasizing grammar and mechanics, tests on knowledge acquisition rather than writing to learn, etc.

Administering SETs electronically, a trend that corresponds with an increase in online education, also determines to some extent how WPAs regard their current SETs. Among respondents, 45% stated they use online forms, 31% print, and 27% both online and print. Comments regarding whether WPAs are satisfied with “the current course evaluation process” and what they would change indicate mixed feelings about this trend. Several respondents indicated that they were moving their SETs online by choice. One was in a period “to test the online system and begin collecting data on the evaluation questions themselves.” Another had “just put in an online [sic] procedure” to start in Fall 2015, while another hoped “to go online for our mid-semester evals” since their end-of-semester SETs were already online. Part of the motivation to go online may, for some WPAs, be the ease with which data can be captured and analyzed. However, other respondents indicated that low response rates diminished their ability to rely on this data. The clearest statement of this problem was from one respondent who answered, “Moving it [the SET form] online has produced a decrease in participation. We need strategies to increase student participation.” Several other respondents also spoke to this issue, stating “response rates are too low to get any kind of reliable data,” “the return rate is usually so low that the results are invalid,” and “moving to an online system has greatly decreased participation in students.” A couple of WPAs expressed distress over low response rates and emphasized the need for strategies to encourage response so SETs could provide meaningful data.

Most surprisingly, according to the survey, WPAs do not have as much access to narrative comments on SETs as one might expect. Although 76% of survey respondents indicated that they can (and do) read the statistical summaries, only 58% reported that they have access to comments. Those WPAs who do not have access to comments find themselves significantly hamstrung in their ability to read and interpret numerical data. In fact, numerical scores without such narrative commentary are acontextual at best. Two respondents expressed a clear sense of frustration at this situation: “I do not have access to non-numeric data (including comments) for all faculty, which I find richer” and “Access to written comments [is a desired change]. The only information released to WPA are the statistical figures for each instructor.” WPAs recognize the disconnect between available information and desired reading practices. The survey asked WPAs to rate the “relative importance of comments to numerical scores or statistical summaries” on a Likert scale of 1–10 (least to most important). The average value was a 7.16, with a majority of responses falling into the upper half of the scale.

WPAs also expressed ambivalence regarding established practices for determining what counts as good SET results, aside from numerically high scores. A question about the accuracy of evaluations produced an average score of 5.60 on a Likert scale of 1–10. On a subsequent question, “How does your institution distinguish between good, average, and concerning course evaluations?”, 8 of the 43 responses (19%) indicated uncertainty about this process or failure on the institution’s part to make such distinctions. Comments on what made the difference between good and bad course evaluations ranged from “I’m not sure” to “Uncertain.” Meanwhile, comments on how institutions helped included “Not very well,” “Not very systematically,” and “Depends. There’s no specific formula.” Such responses illustrate a need for more discussion on how to use SETs effectively and ethically. If the only clear value of SETs lies in their quantitative data and WPAs are unable to articulate what this data means in terms of teaching effectiveness, then it is not reasonable to rely on them very heavily for personnel decisions.

One of the only commonalities in SET use is the calculation of averages, with 23 out of the 43 respondents (53%) relying on such raw quantitative data to distinguish between good, average, and concerning SETs. Some WPAs indicated that numerical data compared to averages was the only way institutions distinguished between SETs: “against the department average,” “Statistically,” and “Comparison to average” were typical comments. Others described in more detail how these numbers were interpreted: “Statistical averages per course: higher than average is excellent, within a couple decimal tenths is fine, lower is concerning.” Another stated that “It varies from department to department, school to school. I believe our department sees a 3.0 (5-pt scale) as acceptable, 4.0 as good,” and another replied that

Comparison to university means based on final two questions: rating of teaching and rating of this course. Means are high (4-point-something out of 5); scores of very high 4’s may earn you a commendation letter for excellence in general education teaching from the vice provost; scores in the low 3’s or worse will cause the Dean to ask the Chair to ask the WPA to contact you for mentoring.

These comments illustrate a major weakness with institutional readings of SETs and WPA involvement in this process: when SETs are read against an average, the number itself begins to assume meaning without appropriate context and attention to external factors such as the instructor’s age, gender, level of course, and/or if course is required or elective. Focusing on SET averages alone is difficult to justify, unless those numbers are somehow cor-

related to other measures of teaching effectiveness and interpreted in light of narrative comments.

Another unsettling trend in survey responses was WPAs' understanding of how external factors can affect SETs. In response to one question which asked WPAs to rate the influence of such factors on a scale of 1 to 10, WPAs readily admitted that expected grade (average response of 7) and course difficulty (average response of 7.06) might influence a teacher's SET scores. However, WPAs are less likely to admit that SETs are influenced by intrinsic factors: gender (average response of 5.58), ethnicity (average response of 6), and appearance (average response of 5.69). In fact, research has shown these do affect SET responses, often significantly (Basow; Basow, Phelan, and Capotosto; Basow, Codos, and Martin; Bavishi, Hebl, and Madera; Reid; Sprague and Massoni). Reluctance to acknowledge the importance of these instructor-related variables could affect how WPAs evaluate SET data. Therefore, we strongly encourage WPAs to consider how gender, age, ethnicity, and other such factors may have influenced an instructor's course evaluations when making personnel decisions. A WPA may need additional information about the class demographics, as well as the instructor's reflections on the semester, in order to determine the most appropriate course of action.

A more positive commonality lies in the practice of reading narrative comments for patterns. Among WPAs who report access to comments, 13 of the 43 respondents (30%) relied on pattern recognition, either alone or in conjunction with numerical scores. For those WPAs who use both comments and numbers, the quantitative and qualitative information provides different windows into an instructor's teaching. As one respondent stated, "High averages plus strong comments are good; moderate averages plus moderate comments are average; a pattern of student concern and low averages demonstrate cause for concern." Another described a process of "comparing comments and numerical scores." A third stated, "Instructor experience relative to numerical scores; patterns in student comments." In these cases, WPAs were able to read SETs more contextually. Other respondents wrote that "[t]he WPA and administrative team read all student evaluations; the questionnaire includes only qualitative questions - no numeric scales," while another described a process in which the "[d]epartment chair (same as WPA here) looks at trends, primarily in written comments." These WPAs see the combined use of numerical scores and narrative comments as essential to the appropriate, contextualized use of SETs. Several reasons may explain an instructor's high or low numerical scores, and it is incumbent on WPAs to discover those reasons rather than risk false assumptions about someone's effectiveness based on numbers alone.

NTT faculty are perhaps the most vulnerable to the merely summative use of SETs, which focuses only on statistical summaries and department averages. Survey respondents ranked the importance of SETs for different types of faculty, saying SETs were most important for adjunct faculty, GTAs, tenure-track faculty, and lecturers. SETs are least important for tenured faculty. Such institutional hierarchies are reflected in survey responses about how likely instructors are to “modify materials or pedagogy based on course evaluations.” WPAs tended to rate adjunct, non-tenure track, and untenured faculty as the most likely to modify their materials in response to SETs. By contrast, they ranked tenured faculty as the least likely to change their materials. Although these are WPAs’ perceptions and do not directly reflect reality, they speak to the differences WPAs notice in the ways SETs are used or not by different faculty ranks in their departments. The job security offered by tenure lends the perception, if nothing else, of senior faculty as immune to SETs. By contrast, NTT faculty have no such security; the perception of their positions as replaceable or even disposable means they must always demonstrate their worth through the means authorized by the university—in this case, SETs.

Although SETs appear to serve a punitive function at many universities, they play a surprisingly minor role in the promotion of teaching effectiveness. Here, we wish to distinguish between teaching effectiveness and teaching excellence. Teaching effectiveness refers to a basic level of competency required for continued employment. By contrast, excellence connotes a higher standard held up for other teachers as an ever elusive goal. Poor SETs often signal ineffectiveness at teaching to administrators and therefore can result in disciplinary action. Meanwhile, our survey found that excellent SETs contribute little to any notion of a teacher’s excellence. Other than securing continued employment, they do little for a teacher’s career advancement compared to other measures, such as classroom observations and teaching portfolios. We asked survey respondents, “How do course evaluations figure into non-promotional decisions such as teaching awards and other recognition?” In this context, 19 out of 45 respondents (42%) indicated that SETs figured very little or not at all into such awards. Only 7 respondents indicated that SETs figured centrally into these awards, and the remaining 19 respondents indicated that they were used in conjunction with other factors such as course observations or that they were uncertain about how SETs were used in such decisions. Clearly, SETs, which are supposed to indicate teaching effectiveness broadly, are often not used to point to how well a teacher instructs students but, instead, to how badly a teacher instructs students. If SETs are not valued as part of the criteria for giving

teaching awards, WPAs may want to openly question why they would be used to sanction some instructors and not used to commend others.

Our survey results emphasize how much SETs are interwoven into the larger discussions in our field about unethical labor practices and the evaluation of instructors. These responses illustrate the punitive aspects of SETs when they are used inappropriately to pinpoint, sanction, and sometimes dismiss faculty. Summative use of SETs alone can have severe, lasting consequences for individuals teaching in our programs, especially those non-tenure track faculty who are often the most vulnerable to administrative decisions. These survey results, therefore, serve as a collective call to action to re-evaluate how we gather and interpret SET information and how we choose to use this information with different groups of faculty. In the conclusion, we offer a set of best practices based, in large part, on our survey results, to help the field recognize changes that need to be made so that SETs are used appropriately and fairly.

REFRAMING EXPECTATIONS: TOWARD BEST PRACTICES IN DESIGNING AND READING SETs

No WPA responding to our survey supported the strictly summative use of SETs. Rather, WPAs wanted greater agency in implementing formative use meant to professionalize teachers and inform their practices. This position also reflects prevailing scholarship on SETs in higher education research. As Ramsden puts it, “No one is going to be frightened into becoming a better teacher by the threat of student-ratings” (232). A culture of fear and anxiety about SETs is more likely to produce lenient grading and lowered standards, primarily because of the belief that they may raise an instructor’s evaluation scores. Furthermore, research also affirms WPAs’ reluctance to use acontextual statistics on SETs based merely on department averages, which are not statistically reliable or valid (Stark and Freishtat). Instead, WPAs should look at distributions of responses where clusters of scores occur and at students’ response rates.

Based on the survey results and secondary research, then, we wish to offer a set of best practices regarding the use of SETs to assess teacher effectiveness:

1. Writing programs should implement SETs that reflect their own pedagogical values and local conditions. These evaluations may be administered in addition to, if not in place of, their institution’s mandated forms.

2. Those who make personnel decisions using SETs should collect as much context as possible and resist the temptation and institutional pressure to only use statistical summaries. WPAs should always be given access to narrative comments, and they should consider external factors such as gender, ethnicity, and appearance as possible influences on students' evaluations of their instructors.
3. WPAs should ensure high response rates to course evaluations by having them completed in class (in a computer lab or on smartphones, tablets, or laptops) or by offering incentives (extra credit to everyone if a certain percentage complete them, a replaced quiz grade, etc.). Otherwise, response rates are too low for results to be usable.
4. WPAs should clarify with other faculty (department chairs, committees) who has the right to read composition instructors' SETs and why. If other faculty have access to these SETs, then they should inform, not countermand, WPAs' decisions about personnel.
5. Programs and departments should develop consistent criteria across faculty ranks regarding the impact of SETs on personnel decisions as well as the promotion of excellence in the form of raises and awards. WPAs should be given the authority to mentor tenured faculty who receive negative feedback on composition courses. At the least, tenured faculty who consistently have problems teaching composition courses should be assigned elsewhere.
6. WPAs and other supervisors should collaborate to ensure that faculty at all ranks are incorporating feedback from SETs into their course materials—such as syllabi, lesson plans, teaching strategies and assignments. The annual review process for all faculty can be adapted to provide accountability for making such pedagogical changes.
7. Programs and departments should strive for transparency in how they use SETs. They can achieve this by composing specific policy language on what qualifies as excellent, good, fair, and poor SET results. Written policies would also describe standard procedures for addressing lower SETs. This language could be used in a separate document or integrated into existing documentation on ten-

ure, promotion, and expectations of graduate and other non-tenure track faculty.

We hope these best practices will help to foster more disciplinary discussion among WPAs about the appropriate design and use of SETs. They may be especially helpful for newer, less experienced WPAs as well as those who have less control over the evaluation process.

Our aim in this article is not to answer all of our questions about SETs but to orient WPA discourse toward them in hopes of articulating a set of coherent positions on how to design and implement SETs that can serve various stakeholders. By learning how WPAs approach SETs, we can begin to formulate a cohesive set of principles rooted in our own tested values and theories about writing instruction. Ideally, we envision a movement toward the creation of locally-appropriate evaluation forms and procedures that balance the needs of all stakeholders—students, teachers, and administrators. This move would better represent the teaching that occurs in writing programs and practices that take a broader view of teaching effectiveness than simply SET results. Such revision, however, must take place within the context of discussions in the field and at individual institutions about best practices regarding the formative and equitable use of SETs. Finally, we conclude by noting that a vast majority of research on SETs gathered for this article appears in publications from other fields such as psychology, business, accounting, and foreign language instruction. Such scholarship, though not specific to composition, can help orient our own investigations into SETs, but we want to assert the importance of studies about SETs specific to writing instruction. Teachers and WPAs can benefit from richer study and discussion of SETs in the context of our own pedagogies and values.

APPENDIX A: SURVEY QUESTIONS

1. What is your administrative role at your institution?
2. Please identify your type of institution
3. Please indicate the region of your university:
4. Please describe the size of your institution;
5. Please indicate how many of the following you have teaching first-year writing courses (adjunct faculty, lecturers, graduate teaching assistants (GTAs), tenure-track faculty, tenured faculty);
6. How long can you say with certainty that your institution and/or program has conducted course evaluations?
7. Are course evaluations mandatory at your institution?

8. Does your institution and/or program use print or online processes for course evaluations?
9. Does your institution have separate course evaluation forms for first-year writing courses?
10. Please indicate the nature and type of questions your course evaluation form asks. Check all that apply:
11. Does your institution produce statistical summaries of individual instructors' course evaluations on a semesterly or yearly basis?
12. Does the WPA at your institution read these statistical summaries?
13. Who else reads these summaries? (Please check all that apply.)
14. Does the WPA and/or institution generate statistical averages for evaluations by course or department?
15. Does the WPA see or read students' evaluative comments that have been excerpted by faculty in cover letters and/or teaching philosophies?
16. Does the WPA see or read students' comments on the evaluation forms themselves?
17. What do you see as the relative importance of comments to numerical scores or statistical summaries? (1 = much less important, 5=equal, 10 = much more important)
18. How does your institution distinguish between good, average, and concerning course evaluations?
19. What assessment of course evaluations is the WPA most likely to perform for the following: (adjunct faculty, lecturers, GTAs, tenure-track faculty, tenured faculty)
20. What decisions is the WPA most likely to make based on course evaluations for the following: (adjunct faculty, lecturers, graduate teaching assistants (GTAs), tenure-track faculty, tenured faculty)
21. Please elaborate on decisions or actions the WPA at your institution takes to address what are seen as negative/concerning course evaluations for an instructor:
22. How do course evaluations figure into non-promotional decisions such as teaching awards or other recognition?
23. How important would you say course evaluations are for the following: (adjunct faculty, lecturers, graduate teaching assistants (GTAs), tenure-track faculty, tenured faculty)
24. How likely do you think an instructor would be to modify materials or pedagogy based on course evaluations?
25. How accurately do you think course evaluations reflect teacher effectiveness? (10 = very accurately, 1 = not accurately at all)

26. How likely do you think course evaluations are influenced by external factors (grades, personality, appearance, gender, sexuality, ethnicity)?
27. Are you satisfied with the current course evaluation process for first-year writing courses at your institution? What would you change?
28. Please enter any final comments you have about the survey here:

WORKS CITED

- Basow, Susan. "Student Evaluations of College Professors: When Gender Matters." *Journal of Educational Psychology*, vol. 87, no. 4, 1995, pp. 656–65.
- , Julie E. Phelan, and Laura Capotosto. "Gender Patterns in College Students' Choices of Their Best and Worst Professors." *Psychology of Women Quarterly*, vol. 30, no. 1, 2006, pp. 25–35.
- , Stephanie Codos, and Julie L. Martin. "The Effects of Professors' Race and Gender on Student Evaluations and Performance." *College Student Journal*, vol. 47, no. 2, 2013, pp. 352–63.
- Bavishi, Anish, Juan M. Madera, and Michelle R. Hebl. "The Effect of Professor Ethnicity and Gender on Student Evaluations: Judged Before Met." *Journal of Diversity in Higher Education*, vol. 3, no. 4, 2010, pp. 245–56.
- Berrett, Dan. "Scholars Take Aim at Student Evaluations' 'Air of Objectivity.'" *The Chronicle of Higher Education*, 18 September 2014, chronicle.com/article/Scholars-Take-Aim-at-Student/148859/.
- Centra, John A. "Will Teachers Receive Higher Student Evaluations by Giving Higher Grades and Less Course Work?" *Research in Higher Education*, vol. 44, no. 5, 2003, pp. 495–518.
- Creswell, John. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 4th ed., Thousand Oaks Sage, 2014.
- Dayton, Amy E., editor. *Assessing the Teaching of Writing*. Utah State UP, 2015.
- Dresel, Markus, and Heiner Rindermann. "Counseling University Instructors Based on Student Evaluations of Their Teaching Effectiveness: A Multilevel Test of Its Effectiveness Under Consideration of Bias and Unfairness Variables." *Research in Higher Education*, vol. 52, no. 7, 2011, pp. 717–37.
- Gladstein, Jill, and Dara Rossman Regaignon. *Writing Program Administration at Small Liberal Arts Colleges*. Parlor P, 2012.
- Gump, Steven E. "Student Evaluations of Teaching Effectiveness and the Leniency Hypothesis: A Literature Review." *Educational Research Quarterly*, vol. 30, no. 3, 2007, pp. 56–69.
- Hansen, W. Lee. "Rethinking the Student Course Evaluation." *Liberal Education*, vol. 100, no. 3, 2014. *Association of American Colleges and Universities*, 2014, www.aacu.org/publications-research/periodicals/rethinking-student-course-evaluation.
- Marsh, Herbert W. "Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research." *International Journal of Education Research*, vol. 11, 1987, pp. 253–388.

- Nasser, Fadia, and Knut A. Hagtvet. "Multilevel Analysis of the Effects of Student and Instructor/course Characteristics on Student Ratings." *Research in Higher Education*, vol. 47, no. 5, 2006, pp. 559–90.
- Newton, Camille, and Ellen Schendel. "More Than Meets the Eye: Teaching Portfolios As Sites of Institutional and Disciplinary Inquiry." *WPA: Writing Program Administration*, vol. 24, no. 1–2, 2000, pp. 55–68.
- Ramsden, Paul. *Learning to Teach in Higher Education*. Taylor and Francis, 2003.
- Reid, E. Shelley, and Heidi Estrem. "The Effects of Writing Pedagogy Education on Graduate Teaching Assistants' Approaches to Teaching Composition." *WPA: Writing Program Administration*, vol. 36, no. 1, 2012, pp. 32–73.
- Reid, Landon D. "The Role of Perceived Race and Gender in the Evaluation of College Teaching on Ratemyprofessor.com." *Journal of Diversity in Higher Education*, vol. 3, no. 3, 2010, pp. 137–52.
- Sprague, Joey, and Kelley Massoni. "Student Evaluations and Gendered Expectations: What We Can't Count Can Hurt Us." *Sex Roles*, vol. 53, no. 11/12, 2005, pp. 779–93.
- Stark, Philip B., and Richard Freishtat. "An Evaluation of Course Evaluations." *ScienceOpen*, 26 September 2014, doi: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1.
- Stark-Wroblewski, Kimberly, Robert F. Ahlering, and Flannery M. Brill. "Toward a More Comprehensive Approach to Evaluating Teaching Effectiveness: Supplementing Student Evaluations of Teaching with Pre-post Learning Measures." *Assessment and Evaluation in Higher Education*, vol. 32, no. 4, 2007, pp. 403–15.
- Thaiss, Chris, and Tara Porter. "The State of WAC/WID in 2010: Methods and Results of the U.S. Survey of the International WAC/WID Mapping Project." *College Composition and Communication*, vol. 61, no. 3, 2010, pp. 534–70.

Courtney Adams Wooten is Assistant Professor of Rhetoric and Composition and the First-Year Writing Program Administrator at Stephen F. Austin State University in Nacogdoches, TX. She studies feminist rhetorics, writing program administration, and first-year composition. Her work has previously appeared in Composition Studies and Harlot. She is currently co-editing the collection WPAs in Transition and is working on a book manuscript about the rhetorics of childless-by-choice women.

Brian Ray is Assistant Professor of Rhetoric and Writing at the University of Arkansas at Little Rock, where he also serves as Director of Composition. He is author of Style: An Introduction to History, Theory, Research, and Pedagogy. He has also published in Written Communication, Composition Studies, Rhetoric Review, Computers and Composition, and College Composition and Communication.

Jacob Babb is Assistant Professor of English at Indiana University Southeast. He is also the Associate Book Review Editor for WPA: Writing Program Administration. He publishes on composition theory and pedagogy, writing program administration, and rhetoric. He has published articles in Harlot and Composition Forum, and he has published chapters in several edited collections.