1999

How Basic Course Directors Evaluate Teaching Assistants: Social Constructionism in BasicCourseLand

Nancy L. Buerkel-Rothfuss
Central Michigan University

How Basic Course Directors (BCDs) Evaluate Teaching Assistants (TAs): Social Constructionism in BasicCourseLand

Nancy Buerkel-Rothfuss

Evaluation and feedback are crucial components of any organizational structure. Employees seek and receive feedback as a means to improve their job performance. Managers, directors, administrators and other supervisors offer feedback to subordinates in an attempt to enhance the overall quality of the organization. Knowing how others perceive us is the first step in improving those perceptions and our position within the group.

In the basic course, evaluation of teaching staff frequently falls to one individual: the director (BCD) for that course (Hugenberg, Gray & Trank, 1993). How that evaluation occurs and what criteria are used may vary widely from one program to the next. Evaluation may be as simple as reviewing student opinion survey forms or as complex as observing/videotaping class sessions and offering detailed critiques of those performances for Teaching Assistants (TAs).

By its very nature, evaluation tends to be subjective. We assess some product as "good" or "bad," "appropriate" or "inappropriate" according to some criteria we establish, but those criteria may vary from one individual to another, from one context to the next, based on how we have constructed our realities about the teaching experience (see Shotter, 1993). One's own

preferences for teaching style, comfort in the classroom, strategies for motivating students, and so on can influence what we consider "good" in others. As a result, evaluations of the same TA may vary greatly, depending on who does the assessment. Worse, there are likely to be variations in judgments even when the same person is doing the evaluation. The same BCD may see events differently from day to day, week to week, and term to term, based on differences in that person's level of interest, fatigue, comfort, stress, and so on.

Teaching is an especially difficult activity to judge objectively. Who is to say when lecture works and when it does not? Generally, it would take more than one classroom observation for a BCD to make good judgments about teaching styles selected, clarity and appropriateness of objectives, quality of activities used overall, and other pedagogical choices. BCDs can observe the quality of interaction between TAs and students, but it's often difficult to parcel out the effects of time of day, day of week, time of semester, immediately past events in the course (e.g., return of an especially difficult assignment on which most students fared poorly), and so on. Furthermore, BCDs can observe preparation, confidence, and knowledge of subject matter and may draw some conclusions about credibility but, once again, these evaluations must be couched in tentative terms if they are made only once or twice each term.

Of course, there remain the questions of validity and reliability. What do the descriptors used in those evaluations "mean" and do those meanings hold true for everyone using the same terms? What is a "competent" instructor? What makes up a "good" teacher?

According to early linguists, the terminology we have available to describe an event or observation influence how we see and think about what we experience

in the world (Whorf, 1956). Thus, the degree of differentiation inherent in our terminology determines our ability to talk and think about specific distinctions. For example, a BCD who has experience differentiating between "one-way" lectures and "interactive" lectures can talk about specific distinctions between the two without necessarily resorting to labels such as "good" or "bad." Another BCD who has never learned to differentiate among the various levels of learning (knowledge, comprehension, application, and so on) may not be able to distinguish between questions that test knowledge-level objectives and those that require synthesis of materials. Thus, the variety of terms we have for a stimulus can influence the degree to which we can identify the nuances that differentiate that stimulus from others that may be quite similar.

Additionally, people with varying experiences will have different interpretations for the same terms. For example, "competent" to one BCD may mean highly skilled; to another it may be acceptable but just barely! What constitutes a "good" lecture to one BCD may be a "dry, pedantic, one-way presentation" (with lots of good information and plenty of examples) to another. Individuals who tend to think in bipolar terms often see greater differentiation between groups of individuals (the "good" guys and the "bad" guys) than those who can see the many gradients of gray between black and white (Delia, O'Keefe, & O'Keefe, 1982). Thus, the labels BCDs routinely use to evaluate (and perhaps even to think about) their TA staff members could color their overall perceptions about those individuals.

Recent research has provided innovative ways for BCDs to describe and think about TAs. Some of our colleagues differentiated among TAs based on their level of professional maturity and progress toward becoming a member of the professorate. From this perspective,

faculty regard TAs as being on a continuum from freshly recruited to the academic ranks (and, as a result, very eager but unprepared) to colleagues-in-training for the day when they, too, will become tenure-track faculty. Gray & Buerkel-Rothfuss (1993) identified eight possible TA "types" in an effort to develop a scale that would allow for better selection and training of graduate students to be teaching assistants. Those types included TAs who prefer to lecture ("lecturers"), TAs who try to become close to their students and want to be liked by them ("buddies"), TAs who think they should never be wrong about anything in front of their students ("omniscient"), TAs who prefer a standardized course which requires little original thought from them ("followers"), TAs who believe that teaching is a popularity contest rather than a set of skills that can be learned and improved ("performer/personality") and TAs who would rather have a research assistantship ("researcher").

More recently, Buerkel-Rothfuss & Gray (1995) discussed various other approaches to differentiating among TAs: (1) TA attitudes toward and expectations about teaching, (2) TA attitudes toward and expectations about the overall graduate school experience, and (3) TA attitudes toward and beliefs about students. Thus, according to these researchers, it is possible to think of TAs in terms of their approach to teaching, the value they place on teaching relative to other activities in graduate school, their beliefs about what motivates students and how they should be led or managed in the classroom, and so on. While not necessarily a better coding scheme than thinking of TAs as "good/bad" or "competent/incompetent," these approaches do yield richer information about BCD perceptions and evaluations. They also offer the potential for more usable feedback for the TAs themselves.

Although no hypotheses were developed for this study, it was assumed that a BCD's experiences would shape the types of evaluations used. For example, in departments where a standardized student opinion survey form is administered, this form probably plays a role in TA evaluations. Thus, BCDs from those departments might use the terminology from the evaluation forms as a basis for discussing TA abilities (e.g., is prepared for class, respects students, etc.). Likewise, departments which focus energies on TA training and on faculty teaching improvement were expected to have lists of teaching strategies which might be evaluated in classroom observations (e.g., has set clear objectives, asks open-ended questions, uses immediacy behaviors). BCDs who have minimal responsibility for TA training and supervision probably have fewer categorization schemes for describing TAs than those who are more actively involved in TA success or failure, unless, of course, those BCDs had received prior training in communication pedagogy. BCDs who have only minimal concern for TA teaching probably have the fewest category schemes of all.

The purpose of this study was to begin to explore the ways BCDs describe and evaluate TAs. In particular, the goal was to determine what terminology/descriptors basic course directors use to describe their TA teaching staff. What do they talk about when they describe their TAs? What language do they use for assessment? Several research questions guided this investigation:

RQ1:   How systematically do BCDs evaluate TA instructors?

RQ2:   What counts as "data" for these evaluations? Course observations? Conversations with TAs? Social interactions with TAs? Specific evaluation forms?

RQ3: What terms do BCDs use to evaluate TAs as instructors? How complex are their categorization schemes? Is there any relationship between how BCDs describe TAs and research on TA "types?"

METHOD

Data were collected between Spring 1993 and Spring 1994 from a convenience sample of 46 basic course directors at both public and private institutions in four southwestern states and two large state universities in the Midwest (a total of 12 institutions). BCDs in the sample were identified by their department chairs/heads and were located using campus phone directories. They were recruited from a variety of departments, not just communication. Fifteen were from the sciences, nine were from English, nine were from communication, three were from Psychology, five were from Family Studies, one was from Communication Disorders, and four were from departments of Foreign Languages. To be in the sample, a BCD had to (1) have been a BCD for at least five years, (2) have supervised or been responsible for no fewer than five TAs each year, and (3) have had major responsibility for training/supervision of TAs on their staffs (if any was available). Initial contacts were made by telephone. Eighteen people were contacted who did not meet those criteria; after a brief conversation about their general responsibilities, those BCDs were thanked for their time and the interviews ended at that point.

After establishing that they met the three criteria for inclusion in the sample, each BCD was asked a series of questions from a scheduled, open-ended questionnaire. In particular, BCDs were asked (1) how fre-

quently and in what way(s) they observe TA teaching, (2) what other methods they use to assess TA ability and competence, and (3) on what types of interactions and in which contexts TAs are evaluated. These questions were not probed to any significant degree. BCDs were also asked to describe what training, if any, TAs in their department receive prior to or during their teaching experience and the degree to which the BCDs participate in that process.

The directors were then asked to describe the "types" of teaching assistants they have had working for them over the years. The question was open-ended and the only clarification offered was that the BCD could offer whatever descriptions seemed most appropriate for the nature of his/her staff and the context in which they work. The interviewer recorded any use of descriptors (adjectives, labels, etc.) that could be equated to a categorization or evaluation scheme. After those descriptors were recorded, the interviewer further prompted subjects to describe "types" by asking again how the BCDs might differentiate among a given staff at any given time. This second question generally stimulated thinking on the subject of how to differentiate other than through simple evaluation.

Phone conversations lasted from 10 to 45 minutes in length. No one who was contacted by phone refused to participate in the research, although several asked for time to think about the topic and then returned the call to the researcher when they were ready to be interviewed. Five BCDs were contacted initially by phone but later were interviewed in person. These interviews took place in the BCDs' offices, at their request.

RESULTS

The first research question asked how systematically BCDs evaluate TA instructors. Only half (23) of the BCDs in the sample based their evaluations of TAs on personal in-class observations, and only five of those BCDs scheduled observations for every term of teaching. Most indicated that they only observe during the first term of teaching and then sporadically after that. Three of the BCDs said they observe TAs only at the TA's request. Only one, a communication faculty member, indicated that she observes TAs without advance warning; the others all set appointments for observations well in advance.

Of the remaining 23 BCDs in the sample, most (19) indicated that they rely on two sources of information about TA teaching for their evaluations: (1) student opinion survey forms and (2) complaints (or compliments?) from students enrolled in the course. These BCDs tended to schedule feedback appointments only when there were difficulties in a section of the course. The remaining four BCDs in this group tended to view themselves as resource people, not supervisors. TAs could come to them for advice but were likely to go to other faculty advisors instead. These BCDs had no formal control over TA performance evaluation, nor were there expectations in their departments that they would offer such services. All four indicated that their departments focus on graduate student research performance, not teaching. None of these four was a communication faculty member.

The second research question further explored the nature of the evaluations: "What counts as "data" for these evaluations? Course observations? Conversations

with TAs? Social interactions with TAs? Specific evaluation forms?"

As just discussed, student feedback was considered by BCDs in this sample to be a valid and reliable source of information about TA teaching. All 46 indicated that they examine and compare means on standardized teaching evaluation forms completed by students at the end of the term. All BCDs had a mental "cut-off point" below which performance is considered to be questionable. For most, this cut-off point was a mean score on the scale corresponding to "poor" or "inferior" ratings by students. Three of the BCDs indicated that they consider performance below the department and/or college/university mean to be cause for concern. Forty of the 46 said they read selected written comments, either before the TA receives them or as a courtesy to the TA after he or she has puzzled over the feedback alone. Twelve said they read all student written comments for all TAs in their charge. Coincidentally, these 12 BCDs were from departments that offered the smallest number of TA-taught courses or used TAs as discussion leaders in fairly small-size recitation sections. One BCD who supervises 35-40 TAs, each teaching two or three sections of their various basic courses, literally hee-hawed when asked if he read student comments: ". . . I'd go blind if I had to do that!" Only ten of the 46 indicated that they discuss student opinion forms with TAs directly.

According to the BCDs in this sample, student complaints about individual TAs tend to be taken seriously only when they occur in significant numbers. In fact, student opinions in general seemed to be of lesser concern than BCD or other faculty perceptions. A common sentiment was summed up this way: "If students knew what they needed from the course, they wouldn't be the students. They'd be the teachers." Many

of the BCDs in the sample indicated that they receive complaints from students but they tend to discount such problems as typical of any new instructor and only report results of such discussions to the TA when they focus on a common theme or complaint over time. Conversely, two BCDs viewed student feedback as the only valid perceptions. "If a student isn't happy, we have a dissatisfied customer. In this environment, that is close to unforgivable!"

When asked whether or not they give feedback based on social or casual interactions, virtually all of the BCDs in the sample emphatically claimed to discuss only teaching-related behaviors. Problems noted in informal settings tended not to enter into their discussions of TA ability or competence. One BCD went on to emphasize that it is his job to supervise teaching, not personal skills. He described some of his TAs as "very socially inept" but indicated he would never even consider addressing those concerns in discussions with or about them. The lone hold-out, a communication BCD, argued that it is his responsibility to turn out well-rounded graduates from the program. A communication student with a Ph.D. who cannot communicate would be "a blight on the reputation of the department."

The final research question focused on the specificity and complexity of the mental coding schemes used by BCDs to evaluate their staff: "What terms do BCDs use to evaluate TAs as instructors?" Although no hypothesis was posed, the expectation was that most BCDs would describe their staff members in fairly simplistic, bipolar terms.

All of the BCDs interviewed used evaluative words to differentiate among their TAs. In particular, over 90% began by dividing their staffs into two groups: "good" teachers and "not-so- good" or "bad" teachers. A

similar percentage referred to TA attitudes toward their jobs as a way to divide them into two groups: those who enjoy teaching and those who do not. All of the BCDs in the sample used bipolar terms to describe their TAs, suggesting that they evaluate them using a variety of judgments that put TAs into "good" or "bad" groupings. Adjectives used were the following:

- competent
- hard-working
- intelligent
- curious
- prepared
- goal-directed
- creative
- assertive
- dependable
- confident
- personable
- motivated
- bright
- mature
- professional
- dedicated
- task-oriented
- innovative
- respectful
- responsible
- likable
- successful

The implication was that some TAs fit into those descriptions while the others did not. Only two BCDs in this sample talked about using those terms as a continuum under which some TAs fit strongly and others to varying degrees (very dependable, generally dependable, somewhat dependable, etc.). One BCD explained that he rank-orders his new staff members based on how "competent" he perceives them to be after two or three weeks of teaching. With over 20 TAs on his staff, this procedure creates a finely differentiated scale. However, this BCD did not elaborate in any detail on how he made those assessments, even when asked follow-up probing questions. He can "just tell" how they should be ranked.

When probed further to differentiate among staff members, most BCDs moved to a categorization scheme

based on demographic information: Ph.D. students only/M.A. students only/a mix of both, from our institution/from other institutions, older/younger than the typical graduate student, majoring in X or Y, from a specific mix of ethnic or geographic backgrounds, etc. Two-thirds (31 of the 46) of the BCDs in the sample stopped at that point, unable to come up with other ways to describe their TAs, or returned to the earlier discussion of bipolar adjectives.

The 15 BCDs (five from communication) who offered other classification schemes described their TAs from a variety of perspectives, many of which were relevant to the TA expectation and attitude scales developed by Buerkel-Rothfuss and Gray (1995). These categorizations seemed not to come easily or naturally for the BCDs in the sample, however.

Five BCDs talked about general expectations for how TAs should interact with their students as ways to differentiate among their staff members. All five mentioned that TAs can get "into trouble" by trying to be "too similar to their students" and "trying to relate to them as equals." These BCDs described TAs who were "too close" to students (buddies) and those who tried for more of a professional distance. Problems with the TAs who tried to interact on the same level as their students included the following: difficulty with grading credibility later in the semester, conflicts with the BCDs over course policies, student complaints that the instructor was unprepared, and a tendency for the TA not to follow course policies and guidelines (especially dress codes). Behaviors observed (or learned about from third-party sources) included socializing with students at bars or parties; dating students; offering what might be considered "too much help" on assignments, especially those the TA did not like; holding office hours at

inappropriate times or in inappropriate places; and missing staff meetings.

Six BCDs mentioned amount and type of prior training and expertise as a way to differentiate among staff members. All six were responsible for staffs which included both Ph.D. and M.A. candidates, thus creating significant differences in experience levels among staff members. All six discussed the value of students beginning their Ph.D. programs having already had teaching experience and/or training elsewhere and the problems that arise when a TA has little or no prior experience: reticence in the classroom, loss of control, lack of credibility, perceptions of non-professionalism and a lack of preplanning for class. TA training was provided in all of the departments represented by these six BCDs. The two communication BCDs in this group referred directly to research by Nyquist and colleagues which differentiates among TAs based on their relative maturity as teachers: from newcomers to faculty-in-training.

Five BCDs talked about TAs' attitudes toward and expectations about students as ways to differentiate among them. In particular, some TAs tend to exaggerate the difference between them and their students, resulting in a tendency for those TAs to "talk down" to undergraduates (the omniscient TA type?). Others become excessively frustrated with their classes because they assume that all students are like they were as undergraduates: striving to get As, in class every day to learn the material, eager to read and complete assignments in advance, etc. Although no one directly addressed these expectations as being ways of viewing students (externally motivated vs. internally motivated), some of the comments suggested a recognition that TAs as instructors can influence how their experience will go as teachers based on the assumptions they

make about the nature of their audiences. Those TAs who believe students need rewards and punishments tend to over-structure their courses, rely heavily on "pop quizzes" to assure reading, call on students in class and put them on the spot as a way to make sure they will be prepared, cover the material from the book with little elaboration, and feel threatened by student questions in class. TAs who believe students are more like them often fail to cover material in sufficient depth or set objectives that are "over the heads" of their students, sometimes use language that is too sophisticated, and are frustrated with their teaching experiences earlier in the term than others.

Three BCDs, all from science departments, talked about the tendency for some TAs to accept an assistantship merely for the money (TA as researcher?), which all three found to be frustrating. According to those BCDs, TAs in this category frequently neglect their teaching responsibilities in favor of their own graduate work. Those who take the assistantship seriously view it as a "job" and resent intrusions into their lives that would not be expected to be part of a job, such as phone calls from students at home, surprise visits from students at times other than office hours, etc. In one subject's department, teaching is something the first and second-year TAs must do; after that, about half of the best and brightest among them can shift to research assistantships, which carry a 20% higher stipend. The message in that department is that teaching is something you must do but research is something the privileged are allowed to do.

No BCD in this sample directly referred to TAs as assuming different types of teaching styles, such "lecturer" or "follower" (Gray & Buerkel-Rothfuss, 1993). An occasional mention was made of TAs who expect too much from the BCD ("He expected me to provide him

with a syllabus, lesson plans, exam questions--everything!), which might suggest a "follower" type of TA. Several BCDs noted that some TAs are more reticent than others and that the reticent ones are better at leading small group discussions or working in lab or study sections than as lecturers. At least one BCD noted that TAs can get into trouble when they pretend to know everything (the "omniscient") or when they answer every question with "I don't know." Three BCDs referred to themselves as "actors" or "performers" when teaching. Of those, one speech communication BCD trains her TAs to be as engaging as possible and provides them with as many visual aids or other attention-getting devices as possible. She maintains a list of films appropriate for the course, has a set of PowerPoint presentations to be used with a portable projection computer set-up, has a file of fairly elaborate simulations and activities in her office, and uses much of her staff time to generate creative ways of presenting information to students. In an effort to adapt to the MTV generation, some lessons are loosely based around popular media personalities such as Seinfeld, the characters on <u>Friends</u> and even "Spooky Fox" Mulder!

DISCUSSION

The results of this study illustrate the diversity of approaches with which we attack the problem of evaluating basic course staff and give some credence to the claim that BCDs would benefit from exposure to alternative evaluation strategies. Only 15 of the 46 BCDs interviewed for this study could go beyond simple evaluations and demographic descriptors to talk about the TAs in their teaching staffs. However, many of those 15 provided multiple approaches for categorization.

While there is nothing inherently wrong in labeling a TA instructor as "competent" or "incompetent," "motivated" or "unmotivated," and so on, these labels do little to provide information to the TA about how to improve. Furthermore, beginning with such a label may start the appraisal interview on a defensive note, leading to little acceptance of the feedback. Use of such labels could color future observations and conversations by structuring the BCDs' expectations about that TA (Shotter, 1993).

Instead, there would be value in feedback that further describes behaviors and attaches a more behavior-based "label" to the observations. For example, "You are trying too hard to be liked by your students. I have concluded this because I see you grading much more easily than other staff members, using examples that would tend to appeal to less-than-dedicated students (going to the bar, getting "wasted") but could be offensive to the more serious students, allowing students to get you off track during class, and socializing with students during your office hours" might be a better way to offer this feedback than to say "You need to take this job more seriously. You seem more concerned with being liked than being a good teacher."

Perhaps this claim does no more than reinforce interpersonal communication research that argues that descriptive, specific feedback is preferable to general comments and likely to lead to better relationships and more productive results. We can improve behaviors that are specified and described. We can acknowledge attitudes that are identified. Whatever the theoretical basis for the assertion, we can assume that complex, detailed, specific, descriptive feedback will produce better results than thinking of a TA as a "good" teacher or a "not good" teacher. Presumably most of us believe we already know how to give specific, descriptive feed-

back. Nevertheless, it would appear from these conversational data that BCDs could use some assistance in making those evaluations.

In many of our basic communication courses, we discuss social constructionism as a way to explain differences in perceptions (Whorf, 1956). What we do not always remember to add, however, is what advantages having a variety of labels can provide. Being able to differentiate among TAs on more than a gross "good/bad" level could help BCDs offer job performance feedback and ongoing supervision better tailored to the specific needs, values and expectations of staff members. "Buddy" TAs can be taught the disadvantages of getting too close to students. Knowing that they tend to be "buddies" can alert their supervisors to keep a closer watch on their behaviors, too. These are the TAs that could attract the favoritism and/or sexual harassment claims. "Follower" TAs can be motivated to take more responsibility for their students and development of their classes. Because "follower" types tend to be speech anxious (Gray & Buerkel- Rothfuss, 1993), attention to building their presentational confidence could provide the motivation they need to become more self-directed instructors. "Omniscient" TAs can be assured that perfection is not necessary, which may reduce much strain for them and create a more flexible classroom environment for their students. All of the TAs in our charge could benefit, if we make the effort to determine what makes them unique.

REFERENCES

Buerkel-Rothfuss, N. L., & Gray, P. L. (1995). *Predispositions and expectations about teaching style scales: Developing a TA profile.* Paper presented at

the annual meeting of the Speech Communication Association, San Antonio, TX.

Delia, J. G., O'Keefe, B. J., O'Keefe, D. J. (1982). The constructivist approach to communication. In E. X. Dance (Ed.) *Human communication theory: Comparative essays.* New York: Harper & Row.

Gray, P. L., Buerkel-Rothfuss, N. L. (1993). Adapting to attitude differences: A factor analytic profile of graduate teaching assistants. In K. G. Lewis (Ed.), *The TA experience: Preparing for multiple roles* (pp. 41-50). Stillwater, OK: New Forums Press.

Hugenberg, L. W., Gray, P. L., & Trank, D. M. (Eds.). (1993). *Teaching and directing the basic communication course.* Dubuque, IA: Kendall/Hunt.

Shotter, J. (1993). *Conversational realities: Constructing life through language.* London: Sage.

Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf.* J. B. Carroll (Ed.). Cambridge, MA: MIT Press.