Evaluations That Respond: Prescription, Application, and Implications of Responsive Evaluation Theory for Community College Instructional Support Programs

Nathan R. Durdella
California State University, Northridge

This study examines two community college instructional support programs to explore the effectiveness of an evaluation model – responsive evaluation theory – that may ease the tensions between a concern over programs' processes and reporting requirements for program outcomes. The study uses a comparative qualitative case study design and applies responsive evaluation's prescriptive steps to assess the research questions: How effectively does responsive evaluation theory operate as an evaluation model? How does responsive evaluation theory articulate with systematic evaluation theories? Results indicate that responsive evaluation can be an effective model if evaluators consult program faculty and staff, who in turn express an interest in building a collaborative evaluation, and if the purpose of the evaluation is to examine process-oriented issues. Results further indicate that responsive and systematic evaluation models articulate well in that outcomes-oriented issues can be examined within the context of a responsive evaluation. Finally, results demonstrate that the responsive evaluation process can be highly politicized and, consequently, addresses the concerns of stakeholders to varying degrees.

Evaluation Theory for Community College Instructional Support Programs

In the health, education, and social services fields, program evaluation has been dominated by objectives-oriented, outcomes-based evaluation (Rossi & Freeman, 1993; Shaddish et al, 1991). In higher education, evaluation is an outcomes-driven, student-achievement-focused institutional tool often used to build systematic models to assess program effectiveness. In recent years, the reliance on systematic approaches has intensified with the rise of regional accreditation standards and state accountability systems, which support an institutional focus on and resources for the evaluation of academic programs and services and assessment of student learning outcomes. In most sectors of higher education, faculty and staff frequently conduct evaluations collaboratively (Green, 1981). However, competing evaluation approaches have tended to bring tension to the evaluation process. This tension could be related to the observation that faculty tend to have a more "qualitative orientation for the justification of educational programs," in contrast to administrator concerns with meeting internal and external reporting requirements (Perry, 1972). In the public sector, this tension has been exacerbated by a tendency to utilize more systematic approaches that are less resource intensive (Cohen & Brawer, 1996).

This study examines two community college instructional support programs to explore the effectiveness of an evaluation model—responsive evaluation theory—that may ease the tensions between a concern over the programs' processes and reporting requirements for program outcomes. Two primary research questions are asked in the study: How does responsive evaluation theory operate as a model to evaluate instructional support programs? How does responsive evaluation articulate with systematic evaluation theories? The two programs are evaluated using responsive evaluation theory. Data sources include interviews with program staff, faculty, administrators, and participants. Results from the evaluations and a discussion of the results are presented. The study ends by drawing conclusions about responsive evaluation theory's effectiveness and usefulness for community colleges.

Related Literature

Evaluation Research and Program Evaluation

Evaluation research is "the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs" (Rossi and Freeman, 1993, p. 5). Thus, evaluation research uses qualitative, quantitative, or mix-methods research designs and methods to examine programs and services. Program evaluations frequently use the principles of evaluation research to support administrative and managerial decisions. Indeed, Royse et. al. (2006) argue that "program evaluation systematically examines...programs for pragmatic reasons," including administrative decisions related to funding and staffing (p. 11). In general, program evaluation focuses on program needs, theory, process, outcomes, and cost effectiveness. The most common type of evaluation examines outcomes or effects of program components or interventions on participants.

The Dominance of Systematic Evaluation and Emergence of Participatory Evaluation

Historically, program evaluation in higher education developed from Ralph Tyler's work in the 1950s. Tyler's evaluation model emphasized quantifiable student outcomes—often rooted in the program under study (Alkin, 2004). Bennett (2003) claims that Tyler first defined objectives-oriented evaluation as "the process of determining to what extent educational objectives are being realized" by a program, department, or institution (p. 5). Thus, evaluation as a whole may be seen as a mechanism to assess institutions, programs, or services for their effectiveness in achieving their desired outcomes. Later, Rossi (1993) maintained that evaluation "is the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs" (p. 5). In the 1960s, alternative evaluation theories and techniques began to emerge in the literature on higher education evaluation. Frequently, faculty from women's studies departments turned to evaluation models that facilitated the participation of program staff and students in the evaluation. Initially, these researchers saw the inclusion of subjects as better suited to their disciplines, which tended to validate the importance of subjective experiences. Consequently, the evaluation models that they used were more inclusive and their value lay in "how to make [evaluations] accessible and important to users" (Cousins & Earl, 1992, p. 398).

Responsive Evaluation Theory

Like participatory evaluation models, responsive evaluation emerged in response to the dominant model that often relegated program staff to the role of passive observers in the evaluation process. Guba (1978) reports that Stake distinguished responsive evaluation from systematic evaluation. The term "responsive" in Stake's evaluation model derived from his desire to, literally, respond to the needs of his clients. Guba (1978) finds that the primary purpose of evaluation should be "to respond to audience requirements for information, particularly in terms of the value perspectives held by each audience" (Guba, 1978, p. 34). Responsive evaluation (a) orients to program concerns and activities, (b) addresses the requirements for information from program staff, and (c) includes the different perspectives of program staff and participants (Shaddish, et al., 1991, p. 275).

Stake (1995a) broadly conceptualizes responsive evaluation as "a general perspective in the search for quality and the representation of quality in a program" (p. 87). If an evaluation uses a responsive model, Stake sees important implications. First, the evaluation design changes as the program changes, making original research design decisions more or less relevant depending on how the evaluation unfolds (Stake, 1995a). Second, a responsive perspective requires that the evaluator examine the program more closely. Finally, the responsive evaluator pays careful attention to all stakeholders so that he or she does not look for how representative or typical the program or program's participants are but rather the uniqueness of an event or perspective (Stake, 1995a).

Although appropriate in educational contexts, Stake's responsive model has drawbacks. House (2001) argues that a weakness of responsive evaluation "is that it can be conservative and relativistic" (p. 28). Shaddish, et al. (1991) called Stake's model conservative in the sense that program improvement is an incremental, piecemeal process. Indeed, Stake's emphasis on getting to know the program, understanding the program's issues, and surveying the perspectives of all participants requires that an evaluator stays with a program for a long period. Consequently, recommendations for program improvement occur slowly. In contrast, systematic models would be capable of measuring program outcomes through statistical analyses and/or survey instruments much more efficiently than Stake's model could, allowing for more efficient processing of evaluation results.

In Stake's responsive model, evaluators proceed through prescriptive steps to conduct the evaluation. Guba and Lincoln (1989) conceptualize Stake's model as four phases. In the first phase, key program staff and participants "are identified and solicited for those claims, concerns, and issues that they may wish to introduce" (p. 42).

The second phase consists of processing the claims of program staff and participants and bringing them to staff members, participants, and other audience members (e.g., administrators) for individual comment. In the third phase, the claims, issues, and concerns that have not been resolved by consultation and negotiation are used as the "organizers for information collection [i.e., data collection] by the evaluator" (Guba & Lincoln, 1981, p. 42). Finally, in the fourth stage the evaluator approaches each audience member with the evaluation results to resolve all issues, concerns, and claims in an effort to improve the program.

Procedures

The purpose of this study is to examine responsive evaluation theory as a model for program evaluation in the community colleges. Further, the study seeks to understand how responsive evaluation theory articulates with systematic evaluation theories used in the two programs. Accordingly, two primary questions are asked in the study: How effectively does responsive evaluation theory operate as an evaluation model? How does responsive evaluation theory articulate with systematic evaluation theories? The goal of the study is to assess the effectiveness of responsive evaluation theory as both an alternative and complement to systematic evaluation theory. To this end, the study uses a comparative case study design and applies responsive evaluation theory's prescriptive steps to assess the research questions. The study proceeds by identifying program participants to inquire about issues or concerns of program directors, developing an evaluation framework to address these issues or concerns with key program and campus constituents, collecting data about the issues or concerns, and finally presenting and discussing the results with program directors to bring some resolution to them. The model's effectiveness will be demonstrated by the extent to which responsive evaluation theory's prescriptive steps (a) identify, (b) investigate, and (c) resolve keys concerns, issues, and claims of program participants.

Case Selection and Descriptions

The two programs under examination were selected purposefully as cases. Program characteristics and the uniqueness of program activities were considered during case selection. The two programs were unique to the instructional support function of community colleges and also shared key similarities. First, the two programs shared the same target student population: underrepresented and/or educationally disadvantaged students. In both cases, Latino students comprised the majority of program participants. Second, both programs focused on math and sciences; however, Program B at one community college also included engineering, while Program A at the second community college was geared towards the health science fields.[1] Third, they both were funded categorically. In addition to the general and demographic student characteristics of the programs, the community colleges with which the two programs were affiliated influenced their selection as cases. Although the two colleges diverge on many student demographic characteristics, they share a few key characteristics. As Table 1 demonstrates, Latino students comprised approximately half of the total enrollment in Fall 2004 at both community colleges—both large, ethnically diverse community colleges in Southern California. Further, the percentage of students enrolled in 12 or more units at both institutions was similar (23.8% at Program A's college and 25% at Program B's college).

Table 1. Case Comparison of Selected Student Demographics, Program A's College (N=21,428) and Program B's College (N=22,223), Fall 2004

	Percentage of Students Who Identified with a Demographic Category at Each College	
Category	Program A's College	Program B's College
Latino/a	50.1	48.9
Female	58.5	43.6
Age 20-29*	47.8	25.0
Enrolled in 12 or more units	23.8	25.4
Enrollment Status: Continuing Student	61.3	76.2

Source: Program A's College Office of Research and Planning, Program B's College Research Department
*Program B's College figure represents students age 22-29.

Data Sources and Sampling

Data sources consisted of interview and journal data. Table 2 lists the 4 subject groups—program participants, program faculty, program staff, and college administrators—selected through purposeful and random sampling techniques. The use of both sampling strategies accomplished the study's need to seek out cases that were most relevant to the research questions and to maintain rigor in the design. In general, the use of purposeful sampling techniques supports comprehension of the research problem (Creswell, 2003; Stake, 1995b; Patton, 1980). Further, sample sizes in qualitative research studies generally do not need to meet the same requirement for representativeness as studies intended to generalize to larger populations from which samples are drawn (Patton, 2002).

Program participants were selected using both purposeful and random sampling strategies in a multi-stage design. For student interviews, program directors were first asked to respectively identify students (n=20) who were actively engaged in program activities. Second, among the active participants identified in the first stage, 10 per program were randomly sampled and invited to interview. Finally, 3 active participants per program responded to invitations to interview. As Table 2 reveals, program faculty (N=4) and staff (N=2) were invited purposefully to participate in interviews. The final sample consisted of 2 program faculty and 2 program staff. Three administrators from Program A's college and two from Program B's college were purposefully selected for their supervisorial roles over the programs. In addition to interview data, 5 Program A students and 4 Program B students were invited and accepted invitations to journal. Students who program directors perceived as being likely to commit to maintaining journals were invited to participate.

Data Collection and Analysis

Data collection commenced in spring 2005 with program staff and faculty. Program directors were interviewed first for the purpose of identifying salient issues, concerns, or claims. Over the course of the study, program directors were interviewed three times using a semi-structured, in-depth interview protocol. Interview questions were guided by the issues identified by program directors and varied by program. Interview questions for Program A related to perceptions about concerns over the use of cultural pedagogy and the perceived value of the program. Interview questions for Program B related to perceptions of program impact. The purpose of the series of interviews was to ensure that program faculty and staff confirmed the issues identified by program directors prior to data collection and to perform member checks during data collection and analysis. Semi-structured, in-depth formats were used in interviews. Interviews were conducted over a 3-month period in spring 2005. Over a period of 4 weeks in late spring 2005, program participants selected for journaling were asked to maintain journals of their experiences related to issues identified by program faculty and staff.

Once all interview and journal data were collected, data transcription, organization, reduction, and analysis commenced. Patton (1980) recommends "organizing the data by specific cases which permits in-depth study of these cases" (p. 303). Consequently, data from the 2 programs were analyzed separately using qualitative data analysis software and qualitative content analysis, a technique used to uncover core consistencies and meanings (Patton, 2002). Once transcribed, data were organized by case and source, indexed by issue or theme, and reduced to that which was most relevant to evaluating the research questions.

Table 2. Interview Data Sources, Program A and Program B, Sping 2005

Interview Subject Group	Sample Population N	Purposefully Sampled n*	Randomly Sampled n*	Sample Yield n*	Interview Type
Program participants**	361-671	20	10	3	Semi-structured, in-depth
Program staff	2	2	-	2	Semi-structured, in-depth
Program faculty	4	2	-	2	Semi-structured, in-depth
College administrators***	40-55	3	-	2-3	Semi-structured

Source: Program A's College Office of Research and Planning, Program B's College Research Department
Note: Program directors were interviewed 3 times.
*Per program
**Program A (N=671), Program B (N=361)
***Program A's College (N=55), Program B's College (N=40)

Validity Threats

In qualitative research, the concept of validity has multiple meanings. Creswell (2003) argues that validity is more likely to be defined in terms of the authenticity, credibility, and trustworthiness of the data. Like validity, the quality of a study is an important consideration in qualitative research. In fact, Patton (1990) argues that "the qualitative researcher has an obligation to be methodical in reporting sufficient details of data collection *and* the processes of analysis to permit others to judge the quality of the resulting product" (p. 462). Consequently, this study details the data collection and analysis process. To proceed through data collection and analysis in a way that produces an accurate representation of the cases under investigation and supports the responsive evaluation approach, triangulation and design checks were used. Stake (1995b) recommends data source triangulation and investigator triangulation. Data source triangulation involves considering what the case would look like if the case did not change over time or space (Stake, 1995b). In terms of data source triangulation, member checking—defined subjects examining rough drafts of writing—was used (Stake, 1995b). Thus, subjects were consulted to verify responses. In addition, investigator triangulation was used to limit conclusions to "persons, contexts, and purposes for which data are applicable" (Patton, 1990, p. 471).

Limitations

This is a study of how an evaluation model works in assessing instructional support programs. The units of analysis, including the institutional characteristics, program components, and program participants, all affect the ability to generalize the findings of the study to other contexts. The exclusion of community colleges that are not classified as large, urban Hispanic-Serving Institutions (HSIs) limits the ability of the study's findings to be applied to other contexts. The fact that this study did not include quantitative data—except descriptive statistical data—limits the ability to draw conclusions that would otherwise be available to a researcher.

Results

Program A's Issues of Politics and Student Perspectives on Program Effect

Program A was launched in October 2002 as the result of a five-year, federally funded Title V grant awarded to the college. The program had three primary objectives: (a) to increase the matriculation and transfer of underrepresented students into health science programs and careers; (b) to develop culture-specific pedagogy; and (c) to develop support strategies. Program A identified six program components to meet program objectives: (a) academic, (b) student services, (c) faculty development, (d) family, (e) mentorship, and (f) community. Program A included higher percentages of Latino, female, and younger students as compared to the total student population at the college. Table 3 demonstrates that 66% of Program A students identified as Latino and 78% as female, while 49% and 46% of all students at the college identified as the same, respectively.

Program A staff identified two central issues. First, faculty's and administrators' confusion and hesitation to embrace cultural pedagogy occupied the attention of program staff. Second, program staff acknowledged the program's perceived lack of value by faculty and administrators not affiliated with the program. Additionally, program staff expressed a desire to learn from the students in the program how the program evaluation revealed a program that appeared to be in turmoil. In the middle of a five-year, federally funded grant, the program was confronting major challenges internally and externally. However, in the face of these challenges, students reported being positively affected by their experiences in the program, and program faculty and staff expressed passion for the

Table 3. Selected Student Demographics, Program A's College (N=22,223) and Prgram A (N=071), Fall 2004

	Percentage of Students Who Identified with a Demographic Category in Each Group	
Category	Program A	College
Latino/a	66.0	49.0
Female	78.0	46.0
Age 20-29*	76.0	42.0

Source: Program A's College Office of Research and Planning
*Program A figure represents program participants age 18 to 35.

program. From results of the evaluation, a number of observations can be made about the program.

The process of identifying and processing claims during the evaluation clarified and confirmed program faculty and staff concerns of resistance to their advocacy of cultural pedagogy. Program staff and faculty perceived that the development and promotion of an alternative pedagogy to what they believed was the "traditional" lecture format was an integral part of what they had accomplished. They posited that perhaps some faculty opposed cultural pedagogy. On this point, one staff member commented, "They don't feel like it would have any impact on their students. They have a classroom full of students from different cultures and different backgrounds, and they don't see the benefit to that." However, it appeared that faculty and staff did not have a clear or meaningful understanding of the concept of cultural pedagogy. Moreover, the perceptions of faculty that Program A excluded all but Latino students confounded their already tenuous support for cultural pedagogy.

Statements of administrators and faculty confirmed the perceptions of program faculty and staff that the campus community did not value the efforts of program faculty. The following statements explain how the program did not achieve the full respect and confidence of faculty and administrators. One staff member remarked: "There are issues of, I'd say, with respect to this program and the college in general about the value of it, the need for it because I think there's a prevailing thought that we do already all we can for students of color just by default because we have such a diverse student population." Further, administrators and faculty reported that the program had not built connections with administrators from division deans to vice presidents. One administrator offered a solution to the program's perception as undervalued, saying, "They need to involve more people on campus and they haven't." However, some administrators did see some value in the program or at least the potential to serve students. Either way, program faculty and staff felt that they had not been given the chance to prove their program's worth.

From both student and staff perspectives, key program components like counseling and student services were highly effective. Among the services offered to students, counseling seemed to be the most influential and beneficial. Students reported that Program A counselor was critical to the success or development of their educational plans and, ultimately, to their goals of transferring to a university and entering health-care fields. Among all of the comments from students, one of the most important referred to how strongly the program had worked to involve their families, particularly their parents, in the program. Consequently, the family component of the program was integral to student success. To illustrate, one student summarized, "I think it's really, really nice of [the] program to have that for our parents and to invite them to all our events, because, like I said, it's very important for our parents to know exactly what we're studying." Here, this student reflected on how parent participation benefits both students and parents. Students reported how one or both parents would attend an orientation and, in some cases, subsequently enroll in a Counseling and Guidance course designed for parents of program participants.

In spite of the success of Program A, the program experienced increasingly more resistance on campus. Administrators reported that program staff had not prepared the program to succeed. Further, administrators contended that program faculty and staff had not sought the input of faculty and administrators from these fields when writing the grant and developing program components. On this point, one administrator recalled how a colleague "was not all that happy with some of the stuff going on because, again, some things were being done around her rather than in conjunction with her." The result was that key program components and theory were not clear to faculty and staff outside the program.

Among challenges that seemed most pressing were the gaps between what program faculty and staff said about why they could not secure the college administration's support and the latter's perception on why they thought program staff had not reached out to the campus community. Program staff found that leadership in key instructional divisions with which they worked most frequently had not been supportive of the program from the very beginning. Further, staff members perceived that faculty in those divisions were not open to participating in the program. In contrast, administrators thought program staff had not collaborated with faculty and staff. For example, they rarely invited them to program meetings, activities, or events. Indeed, one administrator said, "The only things I ever hear [from faculty] is why aren't we part of this." Yet another administrator commented on how the program had not approached faculty in the division, stating that there had "been very little outreach from the program to the division." However, program staff agreed that there was a need for the program in the first place, citing the lower achievement of Latino students in math.

Program B's Impact: Student Development and Achievement

Program B supports instruction for student transfer in the math, science, and engineering fields. Specifically, Program B's mission is to support students who have historically been underrepresented in higher education to excel in the math, science, and engineering fields so that they can transfer in these fields. As seen in Table 4, program participants generally mirrored the college's student population in fall 2004, with nearly identical percentages

Table 4. Selected Student Demographics, Program B (N=361), Math and Science* (N=671), and Program B's College (N=22,223), Fall 2004

	Percentage of Students Who Identified with a Demographic Category in Each Group		
Category	Program B	Math and Science	College
Latino/a	74.0	40.0	48.9
Female	43.0	55.0	43.6
Age 22-29	25.0	28.0	25.0
Enrollment Status: Continuing Student	94.5	72.0	76.2
Citizenship Status: U.S. Citizen	41.0	70.0	81.0

Source: Program B's College Research Department
Note: Some figures are rounded to nearest the whole number.
*Students enrolled in credit science or math courses at Program B's College

Table 5. Mean Local Cumulative GPA* by Gender. Program B, Math and Science, and Program B's College, Fall 2004

	Program B†		**Math and Science**		**College†**	
	n=150		n=671		N=30,662	
	Female	Male	Female	Male	Female	Male
Mean Local Cumulative GPA	2.84	2.83	2.79	2.76	2.63	2.76

Source: Program B's College Research Department
Note: Math and science refers to students enrolled in any transferable credit math or science courses at Program B's College
*GPA earned from courses at Program B's College
†Fall 2000-Fall 2004
‡All students enrolled in credit classes at Program B's College

Table 6. Term-to-Term Persistence Rates, Program B, Math and Science, and Program B's College, Selected Terms

	Percentage of Students Who Enrolled in the Next Term in Each Group					
	Program B*		**Math and Science**		**College**	
	N	%	N	%	N	%
Number of Students Who Enrolled in Any Credit Course in the First Term	361	87.0	671	75.0	30662	63.0

Source: Program B's College Research Department
Note: Math and science refers to students enrolled in any transferable credit math or science courses at Program B's College
*Fall 2000-Fall 2004

of students who are female and ages 22 to 29. However, the program had a higher percentage of students who identify as Latino and who are continuing from a previous term. In addition, the program had a lower percentage of students who were U.S. citizens.

Results from the evaluation of Program B revealed a program that was well developed and supported. Indeed, program staff built a strong, adaptable network of resources for students. The issue that program staff identified as the most important to examine in the evaluation was program effect or impact in terms of student achievement and development. Thus, one staff member said that he simply wanted to see "data that support what we do." In terms of student achievement, program participants generally experienced higher persistence rates, higher cumulative GPAs, and higher successful course completion rates than their counterparts in math and science courses. With respect to cumulative GPA, Table 5 demonstrates that both male and female program participants had higher mean, local cumulative GPAs in fall 2004 than their math and science counterparts and all students at the college. Indeed, in fall 2004, female program participants recorded a mean GPA of 2.84 compared to female students who attempted a transferable math course (2.79) and all female students enrolled as of first census (2.63).

In addition, Table 6 shows that program participants had higher persistence rates than students in math and science courses at the college and all students enrolled in at least one course at the college between fall 2000 and fall 2004. In fact, during the period, program participants registered a persistence rate of 87%. By contrast, three-quarters (75%) of students enrolled in credit math and science courses and just under two-thirds (63%) of all students at the college persisted to the next terms during the period.

One area to which program staff pointed was coordination and communication with colleagues. One administrator characterized the program as an "organized system" developed from a network of counselors and instructors. For program staff and faculty, monthly meetings, email exchanges, team teaching, and informal contact cemented their already strong organizational relationships. Thus, one student wrote in his journal about how program staff kept him informed. He said, "The program keeps in contact with the students via email. The emails [usually relate] information about scholarships, workshops, research programs, [and] school activities." Consequently, program staff constructed a triangle of communication between themselves, instructors, and students, setting students up to approach whomever they needed to address their needs and resolve their issues.

Program B's evaluation revealed two areas of possible improvement. Students seemed most troubled by the absence of a program component that invited their parents and siblings to participate. Students most frequently reported not understanding why program staff had not reached out to their parents. For their part, program staff filed two criticisms. First, they perceived a programmatic need to increase engineering course offerings. Second, program staff felt that scheduling conflicts between courses in the Freshman Experience/ Learning Communities seminar had been associated with lower persistence rates for program participants because they had to chose between courses offered within and outside of the program.

Discussion

The purpose of this study was to examine responsive evaluation theory as an alternative model for community college program evaluation. The study sought not only to uncover how effective responsive evaluation theory is to examine community college instructional support programs, but also to understand how responsive evaluation theory articulated with more systematic evaluation theories. Here, an evaluation of the research questions and a consideration of how the politics of evaluation plays a role in the process and outcomes of such efforts are discussed. In addition, the implications of using responsive evaluation in community colleges are considered.

Responsive Evaluation as a Model for Community Colleges

The first research question in the study related to how well responsive evaluation supports the evaluation. In more nuanced ways, responsive evaluation seemed more effective in identifying and evaluating programmatic issues. In three key steps of the evaluation—the identification of issues, resolution of issues, and the presentation of results—Program A seemed to benefit from the process. These three steps are at the center of a responsive evaluation model, and for Program A to have gained the most from these steps supports the observation that responsive evaluation was effective for the program. For Program B, the first two steps seemed to be less effective in identifying and resolving the issue that they identified, but seemed to be missing the dialogue, or "consultation and negotiation," characteristic of responsive evaluation.

In interviews with Program A students, the two key issues identified by program staff seemed to take on a life of their own. They readily responded to questions about the two issues, articulated an understanding of the issues, and contributed to resolution of the issues. Similarly, administrators responded to the interview questions and confirmed claims of program faculty and staff. In this way, subjects in Program A evaluation *identified, investigated, and resolved the issues*. In contrast, there were no real issues to resolve in Program B's evaluation. The key issue—program impact—was first examined through descriptive sta-

tistical data, which were supplied from the college's district research department, and second through a series of interviews about student experiences.

In considering how staff from both programs reacted to the evaluation results, it is clear that Program A staff benefited more than Program B staff. For Program A staff members, responsive evaluation meant that they were able to see the program from the broader programmatic and institutional contexts. In the case of Program B, the program director found the results meaningful but in a limited sense, citing how the results seemed to confirm perceptions. Although this was the case, responsive evaluation as a model seemed to support the program director's interest in the primary issue of program impact. Indeed, program staff and faculty were consulted early and identified one key issue that they wanted to see addressed and resolved. Throughout the evaluation, it seemed as though a more systematic model—one that uses a mixed-methods approach—could have provided similar results.

Responsive and Systematic Evaluation Models

The second research question in this study considered how responsive evaluation articulated with more systematic models. The results of the evaluation of both programs and a discussion of how the steps of responsive evaluation unfolded in the context of each program revealed that responsive and systematic evaluation models articulated well. That is, at one or more points throughout the evaluations, either the results of a past systematic evaluation were discussed, like in the case of Program A, or quantitative data were used in the evaluation, as was the case for Program B. Whatever the case, both responsive and systematic models seemed to be necessary for the completion of the evaluation of Program B, given the concerns with program outcomes that program staff expressed early in the evaluation. For Program A, subjects criticized the program in relation to quantitative data previously reported, whereas in the evaluation of Program B staff members had identified outcomes-oriented data as one of the issues that they wanted to see resolved.

Program A's use of systematic, outcomes-based evaluations over the previous three years had colored their experiences with program evaluation. The program director reported that, in contrast to systematic evaluations required under the terms of the federally funded grant, responsive evaluation found a "face" of the program that had not been known from results of past evaluations. On this point she remarked, "I think it gave us more that kind of face that you don't get from quantitative, you know, numbers, which is important because a lot of what we do is that piece." She reported that, in contrast to the "getting that 'A' in the class" concern that characterized evaluations in the past, she embraced responsive evaluation as a way to respond to the question: "How do you enrich somebody"? She concluded that the results of the evaluation would supplement what the program faculty and staff had already known about program impact in order to address issues related to communication and coordination. Further, the results allowed her to understand how others perceived her and the program; and, subsequently, how to continue the dialogue begun in the evaluation.

Responsive Evaluation and the Politics of Evaluation

A closer examination of the results of the evaluations of both programs reveals that evaluations are not immune from politics that frequently accompany program development and resource allocation. Here, politics refers to the strategies that program staff, faculty, and administrators employ to gain power and control over program resources and the dynamics that are indicative of relationships where scarce resources are at stake. In prior evaluations of Program A, the questions of which evaluation approach to use—more systematic or more participatory—often included intense debates. Program A faculty, staff, and administrators had not agreed to an evaluation model or criteria to evaluate the program. For example, some administrators cited data from prior systematic evaluations to support their view that the program was failing.

The politics of evaluation seem to be tied to the larger dynamics on campus, including the demographic composition of faculty and staff and the ways that administrators, faculty, and staff relate to each other. Thus, institutional characteristics and culture appeared to play a role in the evaluations in this study. For example, Program A staff was up against college staff skeptical of the program's effects on students. In fact, some administrators reported that such opposition was unavoidable on a campus where most faculty and administrators were white, middle class, and wedded to conventional teaching methods but the program that they were expected to support targeted Latino students and featured cultural pedagogy. At the college, administrators seemed to look suspiciously at the results of previous evaluations that suggested key program components had a positive impact on participants and cited a lack of communication of concepts like cultural pedagogy as a reason for refusing to support program development. In contrast, at the college there was a mix of ethnicities among faculty and administrators. Further, the college had a practice of an inclusive shared governance structure. Against this backdrop, administrators strongly supported Program B and program staff actively coordinated with faculty.

In addition to the demographic composition and institutional culture at each college, the programs' prestige seemed to affect the ways that administrators were able to

navigate the evaluation process and how they chose to use the results. From the beginning, Program A had to fight for programmatic survival as an upstart, Title V program. Program faculty and staff had to build support for the program to be seen as a worthwhile investment of the efforts of faculty and administrators. Conversely, Program B came with all the weight of the University of California Office of the President (UCOP), local UC campuses, and science programs at area high schools. The institutional backing and prestige with which Program B came bundled could have masked politics that otherwise would have been present if the program had been an upstart, categorically funded instructional support program like Program A.

For both programs, the responsive evaluation model used in the evaluation uncovered issues that likely would not have been addressed in a systematic evaluation. Although political and power dynamics certainly exist when institutions and programs use more systematic models, they are often not brought to the surface or included in the results. With responsive evaluation, the issues that Program A's evaluation, for instance, uncovered were underlying political concerns that could prove critical to the survival of Program A and the college. Responsive evaluation is well-suited to uncover the politics of evaluation because the model (a) *asks* program staff, faculty, and administrators to identify issues that may be political, and (b) broadens the scope of those invited to participate in the evaluation so that more perspectives—that complicate the results—are included.

Conclusion and Implications

In this study, using the prescriptive steps of responsive evaluation theory, the results demonstrated that responsive evaluation theory in concert with systematic evaluation theories—in the case of Program B—and without the inclusion of a systematic evaluation approach—as was the case with Program A—proved to be effective in varying degrees to the evaluations of the two programs. First, responsive evaluation seems best suited for formative evaluations of ongoing, institutionalized programs or, alternatively, programs that are in the middle of a categorically funded grant period. These programs are ideal candidates for responsive evaluation because there is time to improve the program. Conversely, programs that are near or at the end of a grant may want to look to more systematic evaluation theories. Second, the areas that staff members identify as the issues of the evaluation are critical to how effective responsive evaluation can be for a program. In Program A's case, the concerns were less with program components than with the larger political issues that had dogged the program since the very start. This is not to say that the use of responsive evaluation should be restricted only to programs with critical, externally driven issues. Rather, evaluators must decide, based on the issues that programs identify, whether responsive evaluation would be effective. If Program B had identified more specific issues related to program impact or implementation, responsive evaluation would probably have been able to resolve them. Program faculty, staff, and participants have a distinct role in responsive evaluation. In this study, they proved to be a critical component in resolving the issues identified and guiding the direction of the evaluation. From the very start of the two evaluations, program staff drove the evaluation in that they identified issues that became the focus of the evaluations.

The results of this study are encouraging for community colleges. Responsive evaluation seems best suited for institutions that, from their charge, have to respond to the needs of their students and the community. Indeed, there is perhaps no better place for an evaluation model that responds to members of the community than the community colleges, where faculty and staff play important roles in facilitating student success. In such contexts and under the conditions described above, responsive evaluation can be an appropriate and useful approach for community colleges

References

Alkin, M.C. (2004). Comparing evaluation points of view. In M.C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 3-11). Thousand Oaks, CA: Sage Publications.

Bennet, J. (2003). *Evaluation Methods in Research.* London: Continuum.

Cohen, A.M., & Brawer, F.B. (1996). *The American community college.* San Francisco: Jossey-Bass.

Cousins, J.B., & Earl, L.M. (1992, Winter). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14(*4), 397-418.

Creswell, J.W. (2003). Research design: Qualitative, quantitative, and mixed methods approaches. *Thousand Oaks, CA: Sage Publications.*

Green, K.C. (1981, Jan-Feb). Program review and the state responsibility for higher education. *The Journal of Higher Education, 52*(1), 67-80.

Guba, E.G. (1978). *Toward a methodology of naturalistic inquiry in educational evaluation.* Los Angeles: Center for the Study of Evaluation, University of California Los Angeles.

Guba, E.G., & Lincoln, Y.S. (1989). *Fourth generation evaluation.* Newbury Park, CA: Sage Publications.

Guba, E.G., & Lincoln, Y.S. (1981). *Effective evaluation.* San Francisco: Jossey-Bass.

House, E.R. (2001, Winter). Responsive evaluation (and its influence on deliberative democratic evaluation). *New Directions for Evaluation, 92,* 23-30.

Patton, M.Q. (2002). *Qualitative research & evaluation methods.* (3rd ed.). Thousand Oaks: Sage Publications.

Patton, M.Q. (1990). *Qualitative evaluation and research methods.* (2nd ed.). Newbury Park, CA: Sage Publications.

Patton, M.Q. (1980). *Qualitative evaluation methods.* Beverly Hills, CA: Sage Publications.

Perry, R. R. (1972, December). Institutional research: Vital third force in higher education. *The Journal of Higher Education, 43*(9), 737-753

Rossi, P.H., & Freeman, H.E. (1993). *Evaluation: A systematic approach* (5th ed.). Newbury Park, CA: Sage Publications.

Royse, D., Thyer, B.A., Padgett, D.K., & Logan, T.K. (2006). *Program evaluation: An introduction* (4th ed.). Belmont, CA: Brooks/Cole.

Shaddish, W.R., Cook, T.D., Leviton, L.C. (Eds.). (1991). *Foundations of program evaluation: Theories in practice.* Newbury Park: Sage Publications.

Stake, R.E. (1995a). *Standards-based & responsive evaluation.* Thousand Oaks, CA: Sage Publications.

Stake, R.E. (1995b). *The art of case study research.* Thousand Oaks, CA: Sage Publications.

Nathan R. Durdella, Ph.D., is an assistant professor in the Department of Educational Leadership and Policy Studies at California State University, Northridge. He can be reached at nathan.durdella@csun.edu.