Disrupting Dialogue:
Envisioning Performance Ethnography for Research and Evaluation

**By Sheryl Conrad Cozart,
Jenny Gordon,
Michael G. Gunzenhauser,
Monica B. McKinney,
& Jean A. Petterson**[1]

Sheryl Conrad Cozart and Jenny Gordon are assistant professors in the Graduate School of Education at George Mason University, Fairfax, Virginia. Michael G. Gunzenhauser is an assistant professor of educational studies in the College of Education at Oklahoma State University, Stillwater, Oklahoma. Monica B. McKinney is an assistant professor in the School of Education at Meredith College, Raleigh, North Carolina. Jean A. Patterson is an assistant professor of educational leadership at Wichita State University, Wichita, Kansas.

I'm concerned about what I saw. I'm not happy, not thrilled. It's not what you did but what you said. It should be of deep concern to people who are honestly trying to be A+. (Audience member, June 1998)

Most researchers who have extended ethnographic fieldwork into public performance will experience resistance and hostility from audiences from time to time. This disquieting antagonism, however, more than the audience approval, signals most clearly that ethnographic performance is a form of conduct deeply enmeshed in moral matters. (Conquergood, 1985, p. 2)

For the 1998 American Educational Research Association (AERA) meeting in San Diego, the authors of this paper created a performance ethnography that

came from our work on the research team evaluating the North Carolina A+ Schools Program. The performance was a play on performance itself. We dramatized how schools participating in the A+ Program (an arts integration reform) performed the program to parents, to their communities, to the program evaluators, and to themselves as they learned about the program and attempted to convince others that they were doing it well. Using a genre that we describe retroactively as a cross between ethnodrama (Mienczakowski, 2001) and readers' theater (Donmoyer & Yennie-Donmoyer, 1995), we dramatized scenes we had observed in schools and shared interview data organized around themes (Cozart, Gordon, Gunzenhauser, McKinney, & Patterson, 1998).

Rich with several years of field data, this performance was well received in the context of AERA as we portrayed ethnographic evaluation data in an alternative format before an audience of educators and researchers. We then performed the same material before the A+ Fellows, a group of arts and classroom teachers from schools participating in the A+ Program along with artists and higher education faculty. What was intended to give voice to a polyphony of perspectives and present data in a theoretically playful manner became problematic in front of this audience. While we intended to engage the audience in playing along with our metaphor of performance, instead we encountered resistance to our themes, challenges to our accuracy, and suspicions about our motives.

Instead of validating our point of view, the A+ Fellows forced us to consider other perspectives and rework our own, a process that has led to this paper. In Bakhtin's (1986) terms, the encounter with the audience of A+ Fellows ended our monologue and forced a dialogue for which we were unprepared. For Bakhtin, monologic thought is a closed circuit, which admits nothing new and exists in search of validation. In contrast, dialogue invites new thoughts and welcomes change. Our performance began as a monologue and became dialogic only at the insistence of the A+ Fellows audience as they reclaimed ownership of the data and reformulated our analysis and representation of it.

We construe our performance as failing to grasp the relational basis of dialogic encounter. Through reflection on our experience, we explore the implications of performance ethnography as a powerful tool for representing themes that emerge from evaluation research. Using Bakhtin's (1986) distinction between monologue and dialogue, we develop a rationale for the use of performance ethnography that may foster reflection between evaluators and program participants and enrich the meaning generated from the analysis of qualitative evaluation data.

Performance ethnography is best understood as a dialogue in which performers and audience engage in an equal exchange (Bakhtin, 1981, 1986; Conquergood, 1985; Mienczakowski, 2001). We contrast Bakhtin's conception of dialogue (which generates new meanings) with everyday conversation (in which new meanings may or may not be generated). We further propose that a dialogic approach to program evaluation can lead to a productive relationship between evaluators and

participants that results in richer and more complex understanding of the program under evaluation.

Transforming the Data into a Performance Text

The data for this performance came from a collaborative, mixed-methods evaluation of the North Carolina A+ Schools Program, which was created by members of the arts community and organized by the Kenan Institute for the Arts, a non-profit foundation (Corbett, Wilson, Noblit, & McKinney, 2001). Twenty-five schools participated in the pilot phase of the program, which we studied for four years, from its inception in 1995 to 1999. The pilot A+ schools were mostly elementary and middle schools located throughout the state. The A+ Program has three primary components: (1) integrated thematic units incorporating the arts and hands-on learning experiences; (2) instruction in a variety of art forms such as visual art, dance, drama, and music; and (3) partnerships with parents, area cultural resources, and local colleges and universities. The A+ Program was developed without rigid implementation guidelines, and schools were expected to interpret and implement A+ for themselves.

The five authors of this article served as graduate research assistants on the evaluation team while doctoral students at the University of North Carolina at Chapel Hill. We described how individual schools implemented the program and attempted to capture the diversity of implementation efforts across schools. Our method for the evaluation study was predominantly ethnographic case study, and we collaborated with other members of the team in data collection, analysis, and representation. Throughout the process, we struggled with the tension between qualitative evaluation and qualitative research.[2] In schools, we downplayed our evaluative role, presenting ourselves as documenters of the schools' efforts. Even in team meetings, we resisted viewing ourselves as evaluators and frequently had emotionally charged conversations about whether or not we should (or even could) judge or rank schools in one way or another (Cozart, 1997; Gerstl-Pepin & Gunzenhauser, 2002). Preliminary reports spoke generally about progress, recommendations, and "good" practices, but avoided labeling the schools themselves.

For our presentation, we focused on the metaphor of "performance" — which had emerged from the data at one particular school where a teacher compared the school's A+ implementation to a Broadway production (Gordon, 1998). We explored the many ways that the metaphor could be applied to other schools and the implementation of the program in general. The main thrust of the performance was to show how schools performed the A+ Program, or in other words, how they performed reform. We demonstrated several different themes of performance: we explored how schools performed the reform for constituents to gain support, how they performed their participation and interpretation for visitors such as reformers and researchers, and how they performed for each other in order to learn how to

implement the program. We also showed how we as researchers performed for the schools.

For many of the schools we studied, the A+ Program was a vaguely communicated reform that was open to interpretation, invited experimentation, and proved to be difficult to implement. The open-ended nature of the reform may have contributed to schools' tendency to perform. Furthermore, the research team may have exacerbated the schools' tendency to perform for us by asking them early on in our evaluation to provide us with exemplary A+ activities.

The performance text contained nine scenes, some composed of reenactments of classroom observations and student focus group interviews. In other scenes, we juxtaposed quotes on a theme, using data from interviews with teachers, administrators, parents, and others affiliated with the schools. Interspersed were scenes in which we spoke from our perspectives as members of the research team to provide transition and analysis. Throughout the performance, we read from scripts. Each scene incorporated movement, and many incorporated props. The blocking and props were kept as simple as possible. We recruited several colleagues who were members of the evaluation team but who had not participated in writing the text to help us perform roles. We also included several audience members by asking them to read brief portions of the script we provided for them in advance of the performance. The intention was to evoke polyphony and to highlight dynamic tension in the data.

The performance opened with a slide show depicting student artwork, while a tape of children singing played. The two opening scenes established the rich context of implementation and interpretation in which performance flourished. In the first scene, we stood with our backs to the audience and took turns turning and reading data excerpts in which interviewees explained to us what the A+ Program meant to them. These "definitions of A+" introduced the notion of multiple interpretations of the program among teachers, administrators, parents, and students, indications from our data that the program was understood differently and creatively in different contexts. This was reinforced in the second scene, in which we created a scenario in which a new graduate student joined the evaluation team. The other team members explained the program to the new member by reading from program documents, all from different passages and at the same time. They then advanced their own interpretations. The new evaluator was confused by the cacophony and disagreement, finally resolving, "I guess I'll know when I get to my school."

The third scene incorporated four classroom dramatizations of arts integration. These classroom lessons, which varied widely in instructional quality and innovation, included two social studies lessons (one involving visual art and the other music), a dance lesson, and a math lesson that included music. The fourth scene explored performance between the researchers and the schools through a series of overlapping vignettes. In one vignette, a researcher discovered a memo from a principal that asked teachers to display artwork and plan creative activities for the

56

researcher's visit. In another, a researcher made repeated phone calls over several months before the school agreed to be the subject of a case study, finally gaining access after making a formal presentation to the school's leadership team.

The fifth scene depicted three student focus group interviews from third, fifth, and seventh grades. Third graders gave their interviewer a lengthy list of activities they had done and told him what they had learned. The fifth graders spoke of music, fun, and projects in other classrooms but said that these things did not occur in their own classroom. The seventh graders told their interviewer that the integrated lesson he had just witnessed was rare, that over the course of the school year activities had decreased, but "every day the A+ people come, we do something like this [activity]."

In the sixth scene, researchers presented their understandings of the meaning of the A+ Schools Program after collecting data in schools. One researcher then summarized her argument from a conference paper about the researchers' multiple views on the meaning and value of the program and its effect on their interpretations (Cozart, 1997). In the seventh scene, we depicted a musical performance from a parent-teacher meeting that served the dual purpose of showing parents what children had learned and giving the principal a chance to explain and showcase the A+ Schools Program. To transition into the next scene, we incorporated teacher and parent comments about student performances, their value for student learning and expression and for publicity for the program, and comments about the time it takes to stage them. In the eighth scene, we assumed our researcher roles again, theorizing about performance in the comfortable confines of a coffee shop. In this scene, we advanced an argument about the educative function of performance as a way for teachers to become comfortable with arts integration and school reform. We closed the performance with a ninth scene, a series of comments from interviews responding to the question of what had changed in schools since the implementation of the program. We ended the performance as we began it, by standing with our backs to the audience.

Detached Response vs. Emotional Engagement

As we planned our performance ethnography, we lacked the precise language to describe what we were doing. We designed our AERA presentation so that it might do justice to the complexity of our data and elicit creative responses to our multiple plays on performance. We took seriously calls from Eisner (1997) and Denzin (1997) to explore alternative forms of analysis and representation. Our research did not lead to singular, tidy interpretations, and we wished to avoid condensing the data into a traditional, authoritative report. We wanted to generate for the audience what Eisner refers to as "productive ambiguity" (p. 8), which he defines as "a potential source of insight, a way of keeping the door open for fresh insights and multiple interpretations" (p. 9). In addition, we hoped that our creative approach to depicting the data would mirror the content of the arts-based A+ Program itself.

Our performance ethnography did not fit neatly into any one tradition (e.g.

readers' theater, ethnodrama, critical ethnodrama, performance text, and mystories) but rather incorporated aspects of several different traditions. Some portions of our performance were similar to what Donmoyer and Yennie-Donmoyer (1995) call readers' theater. As they describe it, "staging is simple; scenery is normally limited to stools and ladders; props are used sparingly, if at all … performers hold scripts, and any acting out of a piece is limited and highly stylized" (p. 406). Drawing from theater scholars Kleinau and McHughes (1980, cited in Donmoyer & Yennie-Donmoyer, 1995), they identify this approach as "presentational," meaning that readers' theater does not attempt to create an illusion of reality. "The audience … is invited to create meaning from what is suggested rather than from what is literally shown" (p. 406).

Readers' theater has become associated with the work of Brecht (1964) and his attempt to elicit critical response through creating a distance between the audience and the text. The relationship between the readers' theater actors and the audience is characterized by distance. This distance is meant to elicit a particular kind of meaning making, one designed to provoke an intellectual and critical response. As Denzin (1997) puts it, "Brecht's experiments were intended to create a thinking audience, an audience with detachment, an audience who could think and act critically" (p. 106). For Brecht, the meaning created by the audience is designed to inspire political activism that will lead to change, but it is born of critical, detached response to the performance. In various ways, our performance was directed toward a detached audience. Toward that end, we wrote the performance solely with the AERA audience in mind.

At the same time, in other aspects of our performance, we came closer to Mienczakowski's (1995, 2001) ethnodrama approach, in which scenes are portrayed more realistically. Some of the recreated scenes were more realistic in nature, serving as "plausible accounts of the everyday world" experienced by the A+ participants (Mienczakowski, 2001, p. 469). Mienczakowski uses the term "vraisemblance" as the quality of portrayals that are as close as possible to how they originally occurred. This close representation of important events is meant to elicit emotional responses from the audience.

Through this emotional engagement a dialogue emerges and meaning is constructed. As Mienczakowski (1995) puts it, "the ethnodrama consensual process, extended through the Bakhtinian (1984) dialogical interactions of the informant group's struggle to create and share meaning, are formally structured through the group's discussions" (p. 364). Research participants and/or the audience co-construct the performance itself with the playwrights. Mienczakowski stresses that ethnodramas (as dialogues) are essentially never finished. The scriptwriter has a responsibility to facilitate dialogue with the audience, and portrayals are continuously negotiated and re-negotiated before and after each performance. For Mienczakowski (2001), ethnodrama is an exchange that can place audiences at risk by raising intimate issues and invoking emotional responses.

Our Performance as Dialogically Problematic

In both the writing and performing of our performance ethnography, we underestimated the importance of dialogue. We had in mind the notion that our performance could lead to the creation of novel meanings, but only for our audience and not for ourselves. We had less of a sense of how that would occur. In this section, we describe how that process unfolded and unraveled.

When we presented *Performing Reform: Presenting the Polyphony* at AERA, we distributed a program prior to the performance that included the following statement about our intentions:

> By voicing the many perspectives on A+, we hope to present the complexity of how reform is constructed by the various constituents of each school community. We explore the relationship between performance and change in schools. In choosing this experimental format, we raise questions about the nature of knowledge and its construction. Through this novel method of presentation, we hope to learn more about the implications of our data. (Cozart, Gordon, Gunzenhauser, McKinney, & Patterson, 1998)

Elsewhere in the printed program, we described our performance as "playful," invoking some of the performance ethnography literature. Despite the small audience and the fact that many of them were well known to us, we were both nervous and excited to be sharing our work because of its alternative format and the risks we perceived to be attached to presenting data in an innovative way.

At the end we were delighted by how well the audience received our performance. Several audience members commented on the appropriateness of the form of our presentation for the arts-based program we were studying. Aside from minor comments, the audience seemed to receive the ideas we presented and our interpretations of the data without question. One audience member questioned our stereotypical portrayals of children's actions, others acknowledged similar experiences with the wide range of interpretations that schools have of reform efforts, and one person made a prescient comment, asking if we had presented our performance to any of the schools we had studied. Largely, the response was supportive of our attempt at an alternative form of data representation.

After our presentation at AERA we were invited to do our performance again, this time to a group of about 50 A+ Fellows, who were meeting in the summer to plan an instructional and planning institute for the next year. The A+ Fellows were arts teachers, classroom teachers, artists, and higher education faculty selected to conduct professional development for the A+ schools, and many of them were teachers at the schools participating in the A+ Program. While some of us were concerned about the response we might get from such a highly invested group of stakeholders, we were flush with our success at AERA and agreed to present our performance ethnography to them. Our concerns at that time centered on portrayals

of some teachers as ineffective and even incompetent. We intended to portray them as trying to implement a complex educational reform, but out of context, our portrayals made it seem like we were making fun of teachers. Further, we worried that teachers we portrayed might be identifiable in the relatively small A+ circle. These concerns led us to rework two of our original classroom scenes. Otherwise, we presented our performance as written. In contrast to our AERA performance, we did not provide the A+ Fellows audience with a program explaining our intentions, and we launched right into the performance without framing it or providing an introduction.[3]

The A+ Fellows responded with laughter at appropriate places throughout much of the performance. At the end of the performance the A+ Fellows applauded enthusiastically, and we invited their comments and reactions. Their initial comments were supportive of our attempt at an artistic representation of the data. Several commented that we had captured the various struggles that they and other teachers had undergone to implement the program and to make it spread throughout their schools.

However, ensuing comments made clear that the audience was split in its reaction. Some took issue with the playfulness of our analysis. Much consternation seemed to result from our choice to focus on the performance metaphor. Several members of the A+ Fellows audience inferred from our choice of the performance metaphor that we believed that they had been dishonest with us and had deliberately tried to misrepresent themselves. For example, one member of the audience said, "When you visit, of course we're trying to put our best foot forward. We're trying to share our best stuff when you come because you're not there every day to see it . . . . It's not about trying to fool you" (A+ Fellow, 1998).

Many saw our presentation less as a play on the metaphor of performance and more as a *critical* (and definitive) evaluation of their implementation of the A+ Program. One commented, "If all we are doing is just 'performance,' then let's stop pretending to be an innovative reform initiative and tell [them] to keep their money. If we are not making effective/affective change, then maybe we should put our energies into something else" (A+ Fellow, 1998).

In the sometimes-emotional conversation, several of the A+ Fellows articulated their dismay. One said, "If I'd never heard of A+ before, I'd say this was a clear presentation of ambivalence and confusion as the reality. I was surprised at the ending, being involved I wanted it to end more positively" (A+ Fellow, 1998). Another commented, "I enjoyed it, but I felt mocked. I understand this is a slice, but if it's the only slice others see, if this is their only taste, they need to see Act II" (A+ Fellow, 1998).

Some of the A+ Fellows' additional comments were critical of the performance. Analysis of their comments points to several significant issues: (1) some were unsettled by the suggestion that the implementation of A+ was flawed; (2) some were disturbed that a prior audience at AERA had seen what they considered to be a negative portrayal of A+; (3) some maintained that the research team had misunder-

stood the A+ Program; (4) some challenged our methodology, suggesting our misunderstanding of the program was the result of having spent inadequate time in schools; and (5) some critiqued the quality of the performance itself and offered to help make it more polished.

The critical nature of some of the comments made it difficult for us to engage in conversation with the A+ Fellows. Their responses were unexpected, and we were shocked. In our responses back to them, we spent considerable time in a defensive posture attempting to explain our reasoning and restating our purposes. At the same time, we demonstrated our willingness to listen to their interpretations and to reconsider our own. Afterwards, we stopped at a restaurant on the way home to process the interaction and console each other, beginning the multi-year conversation that led to this paper. Later we found out that the Fellows argued about our performance for the rest of their weekend meeting. Some found that it resonated with their experience, while others remained convinced that we had misunderstood and misrepresented the program.

The Makings of a Dialogic Encounter

Perhaps because the conversation between the A+ Fellows and us was so unexpected, it became a dialogic encounter that forced us to reconsider our stance as evaluators and our intentions in putting together the performance. In addition, the conversation following our performance led the A+ Fellows to consider the implementation of the A+ Schools Program in new ways and therefore advanced their thinking as well. Through our interaction with the A+ Fellows, we encountered something like Mienczakowski's (2001) ethnodrama, more radically engaging than readers' theater. In a significant contrast with ethnodrama, in our performance we did not envision or prepare a specific role for ourselves as participants in a dialogue.

When we developed our performance text, we were less concerned with the dialogic aspects of our work and more interested in presenting our data playfully and creatively, in a manner that retained its complexity. Focusing on an AERA audience, which we presumed to be distanced from ourselves and the schools we studied, we paid less attention to the relational aspect of our work, which hit us with full force with the audience of A+ Fellows. To use Mienczakowski's (2001) terminology, we placed our A+ Fellows audience at risk without realizing it.

For example, our representations were often exaggerated and unflattering, and teachers and students came across as caricatures. Even our efforts to show "good" A+ lessons instead became obnoxious depictions of students and teachers. We portrayed them as flat characters whose understandings were fixed. In contrast, we portrayed ourselves as round characters coming to new understandings of our experience. For instance, we depicted ourselves pondering the many productive meanings of performance, but we did not show our respondents reflecting with similar sophistication. This is evident in the following excerpt

61

from our performance, which is the third of three re-enactments of student focus group interviews.

Stage directions: (Bell rings. JENNY takes off hat, stands up. Others switch hats so bills are facing front. Some sit still while others fidget and sit up in chairs with legs underneath them on the seat. MIKE takes JEAN's hat off, JEAN puts it back on.)

Jenny (to others): So, what grade are you in?

Mike, Sheryl & Jean (Not in unison, as they fidget): Third grade!

Jenny: When I was in your classroom today, I saw volcanoes you made out of clay. What are some other things that you've done this year that are kind of like that?

Mike: We made brownies. We learned how to make them.

Sheryl: We made peanut butter and jelly sandwiches. And we compared them to the world, like the magma, the inner crust, and the outer crust.

Jean: When we made icebergs, we put water in a bag and filled it up with water and put it in the freezer back there, and then it froze. We let it freeze for a week.

Mike: A week? A day.

Jean (To Mike, slowly for emphasis): A week. *(Turning to Jenny, then back out)* We didn't learn anything. We learned how to make icebergs.

Sheryl: Today when we done the liquid measurement.

Mike: We made these volcanoes. We went outside and made it erupt.

Sheryl: The molasses cookie. How to cook. How to measure.

Mike: We tasted molasses. It tastes awful.

Jean: We made icebergs, and we learned about disasters. We learned about the Titanic. It hit the iceberg, and now we learned the bottom of icebergs are real sharp.

Mike: They take more time until they melt.

Stage directions: (Bell rings. They relax in their chairs, take off hats, and sit back.)

This scene is an example of the kind of representation that was troubling to the A+ Fellows. As we analyzed this scene against the A+ Fellows' responses to our performance, it became clear that the multiple interpretations possible from this

scene were left unexplored. In the excerpted section, a child says about the iceberg activity, "We didn't learn anything." We represented the child in an ambiguous way that could be interpreted as disparaging the teacher and the meaning of the activity. Additionally, the scene can be interpreted as dismissive of the school's efforts to implement the program. We selected this particular scene to demonstrate a classroom of children who had consistently done interesting activities integrated into their regular classroom curriculum. As third graders, they were often unable to explain exactly why they did the activities, but that interpretation of the children's responses need not have led to an interpretation that the program was ineffective or that the children were not learning anything.

Instead, presenting the scene differently (and participating in an ensuing dialogue), we could have invited discussion about the multiple meanings presented. Issues of concern to the A+ Fellows could have more easily been brought forth through dialogue, such as the difficulty of assessing student knowledge through means other than standardized testing or as one A+ Fellow suggested, getting teachers to focus on arts integration as a process rather than a product.

In retrospect we have come to recognize that the flattering conversation after our performance at AERA was monologic, essentially confirming our point of view and offering no challenge to it. We left the conference quite exhilarated with the success of our piece. However, our performance lacked qualities of a Bakhtinian dialogue or what Conquergood (1985) calls a genuine conversation that is characterized by mutuality and engagement in multiple perspectives. Conquergood argues that performative ethnographers should adopt a stance he calls dialogical performance, which he explains:

> This performative stance struggles to bring together different voices, world views, value systems, and beliefs so that they can have a conversation with one another. The aim of dialogical performance is to bring self and other together so that they can question, debate, and challenge one another. It is a kind of performance that resists conclusions; it is intensely committed to keeping the dialogue between performer and text open and ongoing. (p. 9)

For Conquergood, performing is a moral act, and the performer establishes a complex stance in relation to the audience, achieving a balance between detachment and commitment on one axis and identity and difference on another. This delicate balance is difficult to achieve in qualitative research and is even more problematic for program evaluators who are supposed to remain detached in order to maintain their "objectivity."

In our case, we made several errors related to our performative stance. To use Conquergood's (1985) terminology, we displayed an "enthusiast's infatuation," whereby we took on a stance suggesting we understood the program participants and could portray their experiences. The stance was not dialogical, however; instead, we communicated a superficial understanding of their perspectives and

trivialized their complex realities. At the same time, our stance took on elements of detachment, whereby we distanced ourselves from the program participants and displayed moments of cynicism toward their efforts to understand and implement the program. Our performative stance exuded greater confidence than was warranted. To put it differently, it was a closed rather than an open stance.

Conquergood's notion of dialogue comes again from Bakhtin (1986). Through dialogue, participants communicate with each other equitably and productively. It is a relation with ethical and epistemological aspects. As Lynch (1993) points out, the ethical aspect of Bakhtin's conception of dialogue means that no one is subordinated to anyone else. "Each [person] must deal with the others (and be reciprocally treated) as real people" (Lynch, 1993, p. 100).

Short of the dialogic ideal Mienczakowski (1995) attributes to Bakhtin, our work instead fell more to Bakhtin's notion of monologue. Missing in monologue is "answerability," which implies continual communication with and responsibility to others. Answerability requires that participants be seen as real and embodied people in order for dialogue to take place. In contrast, monological discourse recreates unequal power relations by maintaining the primacy of the researcher. Participants are relegated to a second-class status, a problem common to program evaluation research in particular and social science research in general.

According to Gardiner (2000), disengagement between researchers and participants reproduces hierarchical power relationships and supplants equal participation in making meaning. As such, certain ideas (notably the researchers') are given voice, while others (those of the participants) are silenced. By rendering participants subordinate, all old ideas are protected, and new ideas that might have arisen through an equal exchange are stifled. Bakhtin (1986) likens these ideas to "a fish in an aquarium [that] knocks against the bottom and the sides and cannot swim farther and deeper" (p. 162). The ideas become dogmatic because there is no possibility for new ideas to emerge.

Attending to Relation in Performance Ethnography

Had we been prepared to engage in a dialogic encounter with the A+ Fellows, our actions would have been more consistent with the relationships that we had developed with the schools during the three previous years of data collection. As evaluators charged with capturing the diverse implementations of the program, we had been careful in our relationships with schools to be accepting of their choices and to resist efforts to categorize or rank schools in their success. During our visits to the schools, our approach had been to clarify our interpretations and conduct ongoing member checks of our analyses of the cultural changes schools were undergoing. We did this on an individual basis with key informants at the schools in which we conducted case studies.

In yearly reports about the program, we attempted to capture the mixture of

successes and struggles; the reports were designed to summarize A+ implementation across sites and were not specific to particular schools. As individuals and in groups, we wrote conference papers about themes that had emerged from our fieldwork, but these analyses were cloaked by pseudonyms and presented in venues that made the analyses anonymous and generally unavailable to the participants in the A+ Program. Consequently, our performance was a radical departure from what they had come to expect from the evaluation team. For some of the A+ Fellows, particularly those who were not teachers at case study schools, it was the first time they had heard anything concretely critical about the program.

Our playfulness and naiveté placed us into a struggle over ownership of the data. In constructing our performance, we created new contexts for the scenes and filtered them through our interpretive lenses. As they watched our performance, the A+ Fellows heard us repeating their own words in the unfamiliar context of our script. The immediacy of performance ethnography made explicit our appropriation of the words of the participants. The audience members reacted with justifiable indignation, an indignation that signaled a reclaiming of their words for themselves.

In an email written to us after the performance, one A+ Fellow shifted the focus of the dialogue from our metaphorical play to the work schools do to educate children. She reframed the dialogue, and in doing so, she challenged the importance of the context we had created:

> When I juggle for students and friends, what may be a five-minute display of a skill actually took HOURS AND MONTHS of practice to make those five minutes seem fun and interesting. What your presentation seemed to do was to show the "five-minute snapshot," but not the hours and months of work which goes into what we do in the schools to help our kids learn. (A+ Fellow, 1998, electronic communication, capitalization in original)

This A+ Fellow's reclaiming of the data and repositioning of the dialogue forced us to question our goals and how we represented both the program and those implementing it. By articulating other points of view, this comment (and other comments made by A+ Fellows) launched the dialogic process and led us to reconsider issues associated with performance as a form of representation.

Looking back, we see that our interest in playful representation outweighed our commitment to engage in dialogue and that our naiveté resulted in the A+ Fellows' sense of betrayal. In crafting our performance ethnography we engaged in an intellectual exercise and constituted a monologic form of knowledge (Bakhtin, 1986) in which we expounded on our interpretation of the A+ Program without regard for alternative views. While we professed to be interested in productive ambiguity, we only anticipated that ambiguity would be explored on our own terms. Our performance lent our interpretation a finished quality that was inconsistent with dialogue and the contestable nature of our interpretation. While we thought that we

were "done" when the script was written, the A+ Fellows audience forced us to reexamine our roles and responsibilities as researchers and evaluators.

Because in performance ethnography performers and audience are brought together in the same space, the dialogic opportunities are perhaps more apparent than they are with other more traditional texts (Conquergood 1985; Mienczakowski, 2001). Audience members and performers engage in dialogue on Bakhtin's metaphoric level as well as on a literal level. This gives performance ethnography a particular intimacy and power.

Envisioning Performance Ethnography for Program Evaluation

In this paper, we depict performance ethnography as having multiple ethical dimensions and multiple implications for the relationships between the evaluator and program participants. We have found performance ethnography to be essentially an intimate form of representation that has tremendous implications for the ownership of qualitative data, particularly data associated with program evaluation.

As researchers and evaluators, we have a responsibility to those who participate in our research, regardless of whether it is for program evaluation or other purposes. This responsibility challenges the core of program evaluation. A traditional program evaluation hires "experts" to observe, analyze, and give feedback to program implementers. The expert stance positions us as evaluators who know more than the participants, which sets up a monologue rather than a dialogue. In assuming the traditional role of evaluator as expert, we become like Bakhtin's fish in the aquarium, knocking against the bottom and sides and unable to swim farther or deeper.

In this section, we address implications for qualitative researchers who wish to use performance as part of evaluation studies. We speculate on how to create a performance that not only gives justice to the multiple voices in qualitative data but that also invites dialogue that is meaningful, non-threatening, and ethically defensible. A dialogic approach to program evaluation research can lead to a more ethical and equitable relationship between evaluators and program participants, which in turn can result in richer understandings of program implementation. It also allows participants to claim ownership of the evaluation and can provide a vehicle for disseminating results.

Purpose

The main implication of our experimentation with performance ethnography is that evaluators need to invite the perspectives of the audience instead of approaching the viewers with a finished product. The performance is best seen as a means to further the exchange about the program under evaluation and to solicit alternative views. The ongoing nature of the exchange needs to be clear for both performers and audience. While the performance offers the possibility for dialogue,

the possibility for disagreement must be protected. Premature or artificial consensus risks returning to a monologic stance.

Collaboration

One approach Mienczakowski (2001) advocates is inviting participants to collaborate on writing the script. Involvement of participants in the program under evaluation in developing the performance can facilitate dialogue early in the creative process, provide shape and direction for the project, and give it life beyond the moment of performance. Collaboration makes ownership of data a less antagonistic issue.

Interestingly, Bakhtin speaks directly to the word polyphony, which appears in the title of our original performance, *Performing Reform: Presenting the Polyphony*. In admiring Dostoevsky's work, Bakhtin (1986) suggests that the authorial voice is not positioned to be more powerful than those of his characters. Instead Dostoevsky gives equal voice to his characters and resists settling their debates. As evaluators, through our performance to the A+ Fellows, we have learned not to assume an authorial position vis-à-vis the participants in our portrayals of them and ourselves.

Segments

If the ideal of creating a collaborative performance is not possible, parsing a performance into smaller segments, then allowing for dialogue between segments, may promote richer dialogue. In our case, we likely inhibited dialogue by overwhelming our audience with multiple images and ambiguous messages. Stopping periodically to take stock and promote dialogue could enrich our own as well as participants' understandings of subsequent scenes.

Interpretive Power

As performers of qualitative evaluation data, we also have the opportunity to limit our interpretations of the data and invite those of audience members. Portrayal of our respondents as static figures and ourselves as round characters locates agency and power in the evaluators and gives prominence to our interpretations. Genuine dialogue could lead to collective interpretations and new understandings.

Privilege

The detached playfulness that we cultivated came at the expense of the relation between our respondents and ourselves. Careful considerations of the power program evaluators have to claim the privilege of playful interpretation should direct us away from blithely portraying the words and lives of others through our own lenses. As we found at AERA, playfulness can lead to a successful presentation

and an exercise that is intellectually stimulating for the researchers. But as Eisner (1997) warned, "We also need to be sure, if we can be, that we are not substituting novelty and cleverness for substance. In other words, we need to be our own toughest critics" (p. 9).

Continuing Dialogue

Performance ethnography constructed as a relational dialogue suggests creating mechanisms for continuing dialogue. We know that our performance encouraged an extensive dialogue among the A+ Fellows, but lacking in this process was an extended and sustained dialogue between them and the program evaluation team. While a few email exchanges between evaluation team members and A+ Faculty members occurred in response to our performance, the dialogue was not extensive; nor was it any longer face-to-face.

We maintain that once the formal, external evaluation component has concluded, the dialogic possibilities of program evaluation should continue. While an external evaluation component might have an ending date, in keeping with the spirit of Bakhtin's (1986) notion of dialogue, the evaluation is ongoing, never finished. The program evaluators can and should foster a dialogic relationship among those implementing the reform, who may then take ownership of the evaluation process. Bakhtin (1986) alludes to the ongoing nature of dialogue when he writes,

> There is neither a first nor a last word and there are no limits to the dialogic context (it extends into the boundless past and the boundless future). Even past meanings, that is, those born in dialogue of past centuries, can never be stable, finalized, ended once and for all) — they will always change (be renewed) in the process of subsequent, future development of the dialogue. (p. 170)

These alternatives attempt to destabilize the unequal power relationships that result from our roles as evaluators. As an embodied form of representation, performance ethnography is primarily a relational enterprise. It is a powerful means for engendering dialogue. Ideally, performance ethnography is a dialogic encounter in which all parties are willing to learn from each other.

All of the issues addressed in this paper speak to dialogue as a frame for performance ethnography and program evaluation. The challenge is to create a performance that serves multiple purposes, including doing justice to the complexity of the data, presenting enough ambiguity to invite multiple interpretations, and inviting dialogue around issues of significance to evaluators and program participants.

Notes

[1] The authors' names appear in alphabetical order to reflect their equal contributions to this

paper and the collaborative nature of the project. An earlier version of this paper was presented by the authors at the annual meeting of the American Educational Research Association, April 2003.

[2] Throughout this article, we make distinctions between our roles as evaluators and researchers to illustrate the ongoing tension we continue to experience in these roles.

[3] In preparing for this second performance, it did not occur to us to print more programs to distribute to the A+ Fellows audience.

References

Bakhtin, M. M. (1981). *The dialogic imagination*. Austin, TX: University of Texas Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays*. Austin, TX: University of Texas Press.

Brecht, B. (1964). *Brecht on theatre: The development of an aesthetic* (J. Willett, trans.). New York: Hill & Wang.

Conquergood, D. (1985). Performing as a moral act: Ethical dimensions of the ethnography of performance. *Literature in Performance, 5*, 1-13.

Corbett, D., Wilson, B., Noblit, G., & McKinney, M. (2001). *The arts, identity, and comprehensive education reform: A final report from the evaluation of the A+ Schools Program.* Winston-Salem, NC: Thomas S. Kenan Institute for the Arts.

Cozart, S. C. (1997, October). *Researching the researchers.* Paper presented at the annual meeting of the American Educational Studies Association, San Antonio, TX.

Cozart, S. C., Gordon, J., Gunzenhauser, M., McKinney, M. B., & Patterson, J. A. (1998, April). *Performing reform: Presenting the polyphony.* Performance presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Denzin, N. K. (1997). *Interpretive ethnography: Ethnographic practices for the 21st Century.* Thousand Oaks, CA: Sage.

Donmoyer, R. & Yennie-Donmoyer, J. (1995). Data as drama: Reflections on the use of readers' theater as a mode of qualitative data display. *Qualitative Inquiry, 1*(4), 402-428.

Eisner, E. W. (1997). The promise and perils of alternative forms of data representation. *Educational Researcher, 26*(6), 4-10.

Gardiner, M. E. (2000). *Critiques of everyday life*. London, UK: Routledge.

Gerstl-Pepin, C.I., & Gunzenhauser, M.G. (2002). Collaborative ethnography and the paradoxes of interpretation. *International Journal of Qualitative Studies in Education, 15*(2), 137-154.

Gordon, J. (1998, March). *A+ as Broadway: A play on metaphor.* Paper presented at the 19th annual Ethnography in Education Forum, Philadelphia, PA.

Lynch, R. A. (1993). Bakhtin's ethical vision. *Philosophy and Literature, 17*, 98-109.

Mienczakowski, J. (1995). The theatre of ethnography: The reconstruction of ethnography into theatre with emancipatory potential. *Qualitative Inquiry, 1*(3), 360-375.

Mienczakowski, J. (2001). Ethnodrama: Performed research — limitations and potential. In P. Atkinson, A. Coffey, S. Delamont, J. Lofland, & L. Lofland (Eds.), *Handbook of ethnography* (pp. 468-476). London, UK: Sage.