GENDER EQUITY PROGRAMS IN URBAN EDUCATION: REDEFINING RELATIONSHIPS BETWEEN FUNDING AND EVALUATION

Alice Ginsberg

Introduction:

This paper focuses on the evaluation of programs and policies designed to promote gender equity in schools and to raise awareness in general about gender issues in urban education. It raises the following problems:

- How to frame and define gender as an important issue in urban education
- How to decide where and by what criteria to distribute limited resources to gender-based educational programs in urban communities
- How to assess and evaluate the impact and importance of gender education work both locally and nationally

The paper has two main themes: The first is how such programs are typically assessed, meaning the different tools and criteria that are used to make judgments as to their worth and ultimate "success," or "failure." The second theme addresses the problems inherent in funding programs which are explicitly focused on gender and gender equity in urban education. Some recent studies and reports (Grady & Auburn, 2000; Mead 2001; Three Guineas Fund, 2001) suggest that it is still difficult to convince key educational stakeholders that gender equity in education is important - despite over a decade of research showing the impact of gender bias on both girls and boys throughout their education (AAUW, 1992; AAUW/Research for Action, 1996; Davis, 2000; Francis, 2000; Ginsberg, Shapiro & Brown, 2004; Leadbeater & Way, 1996; Orenstein, 1994; Sadker & Sadker, 1994; Shapiro, Sewell & DuCette, 1995; Ward, 2002). Skepticism as to the worth of gender equity programs is even greater in the case of urban education, as many of these students have been labeled "at-risk" and schools are encouraged to focus on teaching nothing but "the basics" (e.g., learning to read) ignoring differences in experience, culture, language, and/or community interests and values.

It is important to look at the issue of gender equity from the perspective of the teachers and educators actually working in classrooms and schools as well as that of the funders, administrators and policymakers who are often considered "outsiders" to the reform process. Ultimately, these groups must work together, share common goals and language, and produce evaluations that support innovative and effective programming for urban education. This is not, however, an easy process.

Recent studies confirm that gender equity is not a priority, or even a visible issue, for many school reform advocacy groups, despite an interest in broader issues of equity and racism (National Center for Schools and Communities, 2002). The general public is also extremely confused about just what gender equity in education means, and, in particular, conflicted about whether paying closer attention to girls somehow means shortchanging boys (Grady & Auburn, 2000). Moreover, it has been found that most educational foundations are apt to seek the least controversial funding criteria, that is, to fund programs that are broadly considered to be "universal," rather than targeting one specific group of children (Mead, 2001).

While educational program developers must constantly try to "prove" their programs' worth, many program funders are also looking for ways to justify funding decisions that target scarce resources for school reform work with gender at the center (Mead, 2001). In the current educational climate of high-stakes testing and standardized curriculum, increased accountability, and widespread systemic reform, many innovative reform programs designed to raise awareness about gender issues in schools are short-lived, under-funded, and isolated from other reform initiatives (Ward, 2002). This is an especially relevant issue when thinking about evaluation, as program evaluations are (often) fueled by the desire to gain future funding, and funding agencies routinely use evaluations to make important decisions about what kinds of programs and which programs specifically, they will support.

This paper addresses some of the ongoing questions and concerns that foundations and other educational funding agencies grapple with as they try to support such work. After briefly exploring some of the different models of gender equity programs in **Section I**, I look broadly at the issues of educational evaluation and of accountability (**Section II**). Though the questions raised in this section are not all specific to gender, they do raise important questions about why we place so much emphasis on certain kinds of evaluation, and how we use such evaluations, often narrow in scope, as a measures of "success" and "failure".

Section III explores the ways in which gender equity is not viewed as a priority for school reform, and the resulting ultimate bind funding agencies and policymakers find themselves in, even those that are already committed to supporting this issue. I look at the different ways in which funding agencies define and evaluate the impact and importance of this work both locally and nationally and, in doing so, investigate more intensively how foundations and other educational funding agencies frame and define gender in education work. For example, do a majority of educational foundations believe that gender is not an important enough priority to target limited resources for? Do foundations see gender as a synonym for "girls," and thus believe that it is not "inclusive" or "democratic" enough to merit special funding status (Mead, 2001)? Even if foundation staff members recognize that gender issues are worth paying close attention to, what is the impact on the foundation of supporting an issue that the general public and other important stakeholders do not yet recognize? Should foundations try to fund programs that address gender issues without calling attention to the gendered aspects of the programs? What are the advantages and disadvantages of this approach? For example, is there a danger that gender will become subsumed into other educational problems?

Section IV considers more broadly some of the pitfalls of traditional evaluations when applied to issues of gender and urban education. For example, interviewing or shadowing participants can be very demonstrative, but are rarely cost-effective. Likewise, many funders and policymakers do not value qualitative evaluation at all, believing that it is not scientific or systematic enough.

Section V presents some case study examples of alternative approaches to evaluations designed by the Ms. Foundation for Women which are specifically geared towards looking at gender.

Finally, in **Section VI**, I consider how funding organizations decide how and where to distribute scarce resources, recognizing that all girls are not equally disadvantaged, and moreover, that poor and minority *males* are often identified as the group *most at-risk* (Bierda, 2000; Connell, 1993; Davis, 2000; Flood & Dorney, 1997; Obgu & Simons, 1998). Should foundations reach out to all children equally, or try to concentrate their resources on those groups most in need? How do foundations decide which kinds of gender issues should be their primary focus? What kinds of "proof" do foundations need that their money is being well-spent and that their resources are being distributed to those most in need or most worthy? Are the voices of the program participants themselves the most important voices to listen to? What other kinds of outside measures are necessary (e.g., test scores, numbers of participants reached, materials generated, etc.)? How can evaluations be methodical and intentional, while also being flexible and authentic to those involved? What are the other considerations that drive foundations' decision-making processes? In other words, how relevant, important, or useful is program evaluation at all, in face of political relationships, stakeholder priorities, and long-term funding histories and cultures (Mead, 2001)?

I. Focusing on Gender in Urban Education: Types of Programs and Research

Before I discuss how gender programs are funded, compared, and assessed, I believe it is useful to look briefly at some of the different kinds of programs that fall under the rubrics of "gender and urban education." It is important to note that definitions of terms like *gender equity*, gender awareness, *gender bias*, and *gender studies* are not universally agreed upon in education. Nor, for that matter, are the programmatic elements which comprise them. While some define gender equity in urban education as treating boys and girls exactly alike, or at least giving them equal attention, feedback, and resources (Sadker & Sadker, 1995), others advocate teaching to students' (real or perceived, innate or socialized) *differences* (Gilligan, 1982; Gurian, 2003). This may mean that girls are encouraged to work collaboratively while boys are still working competitively; or that girls pay more attention to language arts while boys are strongly encouraged in the fields of math and science, etc. Others still see gender equity as a form of affirmative action or remediation, as has become most apparent in attacks on policies such as Title IX. Title IX, designed as an equal opportunity in education law, is most notable for its mandate to assure that girls' sports get the same funding as boys' sports, often necessitating a shift of limited funds and resources from boys to girls.

It is also worth noting that gender equity programs differ considerably in size, as well as in scope, content, length and goals. Programs geared specifically for teachers in the form of *professional development* or *curriculum development*, usually include a component where participants are asked to engage in self-reflection on their own gender biases, teaching histories, and cultural identities. As Smith (2000) writes in reference to the long-running Seeking Educational Equity and Diversity (SEED) program, "All conversations begin with teachers reflecting on their own schooling and life experiences in order to think about the way in which school curriculum powerfully shapes human identity and social identity" (p. 139). Similarly, Shapiro, Sewell & DuCette (1995) write, "We have to realize that who we are greatly affects our thinking about categories such as ethnicity, social class, gender and other areas of difference" (p. xiii).

Programs developed directly for students themselves take the form of mentoring, building self-esteem, or creating "safe spaces" for students to voice their opinions and discuss different models of masculinity and femininity (e.g., The Girls' Action Initiative, The Alice Paul Leadership Center, The Girl Scouts, Girls Inc., etc.).

More *research-based* programs systematically explore specific questions about classroom/school dynamics such as why girls actually or seemingly "allow" boys to harass them, or why boys are more inclined to take upper level math and science courses. Often referred to as *action research* or *teacher research*, it has been noted by Cohen and Manion (1984) and many others that this type of research is *situational* (concerned with diagnosing a specific problem in a specific context); *collaborative* (teams of researchers and teachers build inquiry communities and work together, bringing diverse experiences and perspectives to the research), and *self-evaluative* (modifications are made continually throughout the project and is thus a process evaluation as well as a product evaluation). Many times students themselves are asked to help develop, conduct, and interpret this research as an integral part of the curriculum itself (Cochran-Smith & Lytle 1993; Ginsberg, Shapiro & Brown, 2004). When involving students in research it makes sense to choose topics that are pertinent to issues in their lives, such as sexual harassment, career counseling, self-esteem, social justice and community activism.

These are just some examples of the kinds of educational programs that could be considered gender-focused. Such programs are still rare, and getting rarer still as the possible punishment inherent in *No Child Left Behind* (e.g., losing federal funds if schools don't improve test scores) become more and more real. As *No Child Left Behind* appears to be highly concerned with issues of evaluation and accountability, it cannot (and should not) be ignored; however, as the next section of the paper points out, there are different kinds of assessments, many of which are far more effective and revealing than high-stakes testing.

II. Evaluation and Accountability: Co-Joined Twins?

Educational program *evaluation* usually has one (or more) of the following purposes:

1) *Assessment*: in order to improve programs in process or to make changes in future programs;

2) *Research*: to make comparisons of the relative success of different program models, or to study the impact of programs at different sites and with different constituent groups; and to increase the state of knowledge in the field; and

3) *Accountability*: to measure how goals of different stakeholder groups, particularly policymakers and funding agencies, were met. (Council on Foundations, 1993)

It is no secret that the third motivation, *accountability*, is often the one that drives program evaluation. Indeed, accountability (Shapiro, Sewell & Ducette, 1995; Ginsberg, Shapiro & Brown, 2004) is not only the dominant reason given for program evaluation but plays a significant role in shaping the *ways* that programs are evaluated. Accountability is intricately related to the measures, methods, and tools that are considered legitimate markers of success, as well as the form and content of the final report(s) these evaluations take and who reads them. As Shapiro, Sewell and DuCette (1995) rightly note, "Assumably…all the significant outcomes of education can be objectively measured….implicitly or explicitly, assessment continues to drive the curriculum" (p.86). Yet they go on to note that "accountability and diversity tend to go in opposite directions" (p. 87) because accountability leads to uniformity and standardization, while diversity leads to a unique curriculum and to the reflection of individual learning styles in assessment techniques (p.87).

There are some very understandable reasons why both program developers and program funders need to be accountable for what they are doing - not the least of which - because limited and coveted resources are at stake. Nonetheless, it is the premise of this paper that an over-emphasis on accountability has the potential to skew other program goals -- such as assessment and research -- in ways that can ultimately make program evaluation both less authentic and less useful for everyone involved.

Throughout the paper I suggest that the current models of school reform and program evaluation, and the traditional markers of "success," (e.g., test scores) may not be especially useful when looking at gender-based educational programs. For example, an analysis of students' standardized test scores is unlikely to reveal whether teachers are giving boys and girls equal amounts of attention in the classroom or whether girls are taking greater leadership positions, have increased self-esteem, or are considering a broader array of career options (Ms. Foundation for Women, 2000).

Similarly, the overall "cost-effectiveness" of a program (e.g., in terms of its ability to be replicated and numbers of students served) is not necessarily the most important indicator of a worthwhile gender equity program. To be effective, gender programs need to do much more than, as the saying goes, *add women and stir*. Just inserting women into the curriculum is a relatively inexpensive way to make learning seem more equitable, but this method does not address the critical reason *why* women were absent in the first place, and *how* women are usually represented as compared to how men are represented. As Martinez (1995) notes, " a numerical increase in textual references and images doesn't promote multiculturalism if the content leaves a fundamentally Eurocentric worldview in place" (p.101). The same could be said of a patriarchal world view. Enid Lee (1995)

underscores that "if we don't make clear that some people benefit from racism, then we are being dishonest" (p. 13). And Patrick Finn (1999) reminds us that "if we teach children to critique the world but fail to teach them to act, we instill cynicism and despair" (p. 185). Children are highly aware of bias and inequality. In short, these issues of power and inequity need to be discussed and critiqued; it is not enough to simply try to balance them out in the classroom.

Alternative approaches (discussed at length later in this paper) are often very costly in terms of both financial and human resources and require the sustained commitment inherent in long-term components like mentoring and working closely with parent and community groups. These programs are also not easily replicated because different groups of children, in different schools, with different sets of resources available to them need different kinds of support (AAUW/Research For Action, 1996). When designing any gender-based educational program, it is especially important to take into account the intersections among gender, race, ethnicity, and class (McIntosh, 2002). As Ward (2002) notes after conducting a series of focus groups with equity consultants, "Gender Equity initiatives should be specific to and relevant within a context of a child's racial and ethnic community" (p. 4). Ward also found that, "Teachers were also described as stereotyping students by race, and the charge that teachers hold lower expectations for boys of color was heard across the focus groups" (p. 10).

Evaluators may find it difficult to isolate those students and practitioners who are being affected by gender programs, given the extreme state of flux in urban education. Teachers and administrators have noted that they are continually faced with multiple and competing reform mandates that make it extremely difficult to focus their energies on one particular set of goals (Ginsberg, Shapiro & Brown, 2004, p.147). Moreover, due to heavy dropout rates in urban schools, it is usually not the same group of students that are exposed to constant reforms year after year. With an average student dropout rate of as much as 60% at many urban high schools, and a teacher and administrator turnover and vacancy rate that is equally disruptive and alarming, those who participate in these "demonstration" programs are unlikely to be a stable group. For these reasons, gender equity programs (like many other reform programs) -- whether professional development programs for teachers or direct support services for students -- are unlikely to be good candidates for longitudinal, quantitative evaluation in urban education. It is also worth noting that most gender programs tend to be small "demonstration" projects disconnected from larger school programs and policies (Ward, 2002). Thus, when the grant money runs out, the program is quickly forgotten as teachers and administrators are bombarded with new mandates.

In spite of these inherent differences, many funding agencies continue to hold grantees "accountable" to these kinds of measures of success. In fact, the stakes may be even higher for such programs, because of the fact that gender equity remains highly contested and a low priority in most school reform initiatives, and funders and policymakers therefore want to be able to point to immediate and dramatic changes. Unfortunately, this is not an easy task. Many educational stakeholders believe that gender equity is not an important part of educational practice (see Section III) and refuse to prioritize it in any way by giving the participants the needed on-going support and resources. Thus, the emphasis on accountability has the strong potential to impede the process of change. In the case of one urban reform program designed to reduce dropout rates for inner-city children in New York City, developed by an agency called Cities in Schools (CIS), the evaluator reported that:

> CIS has become burdened with accountability. Reports, meetings, schedules, and agency mandates have taken precedence over children and their needs. All three schools visited had principals extremely supportive of CIS. But the program is not a part of the school. The teaching and CIS staff are separate and distinct from each other and their attitudes are often competitive and adversarial. (Council on Foundations, 1993, p. 116)

Moreover, programs which are slow to show change can present a dilemma for foundation staff who fear that "negative" evaluations may "yield information that could reflect negatively on staff judgment" (Council on Foundations, 1993, p. 17). Although it also should be noted that evaluating grants and programs can potentially provide a positive opportunity for foundations to evaluate their own priorities and practices (Council on Foundations, 1993).

III. Making the Case for Gender Equity in Urban Education

- The issue of gender in school reform work: Not even on "The List"

In a recent study conducted by the National Center for Schools and Communities (NCSC, 2002), fifty-one diverse community organizations from across the country were given an extensive list of questions regarding their perceived role in local school reform. The study was designed to "identify shared priorities and issues" (p.1). For the purposes of discussion, an *issue* was defined by the NCSC as "a problem that people understand as being susceptible to policy change and around which they are willing to organize" (p. 1). For example, nearly half the groups surveyed were concerned with supporting 1) after-school enrichment opportunities in their communities; 2) facilitating parent involvement in schools, and 3) increasing funding and accountability.

In its analysis of the interviews, the National Center for Schools and Communities (2002) was able to identify twenty-five "categories" of topics that could be used for inter-group comparisons. Among these twenty-five categories, it was notable that: "*Girls were not to be seen or heard. No interview defined issues, information needs, or context in terms of female students*" (p. 13). This finding is especially interesting given that issues of "equity," "safety," and "racism" were among the categories of topics raised by a large number of respondents.

Given the wealth of contemporary educational research citing gender bias and inequities *across* the curriculum, along with high incidence of sexual harassment and gender violence in schools, and the double discrimination faced by poor girls and girls of color, one would think that *girls*, or at the very least *gender*, would appear somewhere on this list of concerns. But they do not. The question is, *why not*?

When I posed this question to the Center's Executive Director,[1] he responded that although he wasn't sure, he suspected that many people still thought of gender issues -- such as sexual harassment, or girls' limited enrollment in upper level math and science courses -- as an "individual" kind of issue. In other words, they did not see gender as something that needed to be *systematically* and *institutionally* addressed.

- **Public Perceptions of Gender Bias in Education: Not Even on the Radar Screen**

A completely different set of studies conducted by the Frameworks Institute and commissioned by the Caroline and Sigmund Schott Foundation to gauge *public* perceptions of gender equity in education, may offer us some additional clues as to why gender is not seen as a priority in school reform work. Through a series of interviews and focus groups, authors Grady and Auburn (2000) concluded that gender bias in schools was still a largely invisible issue which did not show-up on the American public's radar screen. They furthermore deduced that many Americans still tend to see gender discrimination as a problem in the workplace rather than in the classroom (Grady & Auburn, 2000).

The Frameworks Institute also uncovered a number of other important findings regarding *how* the general public "frames" and understands the issue of gender equity in education. Embedded in these findings are significant tensions and dilemmas which help to explain why gender equity is often overlooked and undervalued in school reform work. For example, the dominant rhetoric that education is the key to social mobility; that schools treat all children equally; and that school systems are not influenced by larger social, cultural, and political issues, is clearly alive in the public's belief that the classroom is an "ideal, controlled environment" (Grady & Auburn, 2000, p. 6) where students are protected from rather than exposed to discrimination and bias.

Likewise, the fear that paying closer attention to girls necessarily means taking something away from boys (many of whom are also "at risk") is also evident in the finding that many people resist focusing on the "specific disadvantages faced by girls" (Grady & Auburn, 2000, p. 6). In other words, people are likely to resist gender equity if they see equity as a metaphor for remediation or affirmative action. This is most currently obvious in the attacks on Title IX, which many people perceive to be unfairly and unnecessarily draining resources from boys' sports in order to provide more resources for girls.

Indeed, as another important finding suggests, many people believe that teachers actually *favor* girls. This may well be because girls are generally quieter, less disruptive, more compliant than boys, and in many cases, get better grades and are considered to be more "mature" students (AAUW, 1992; Grady & Auburn, 2000; Orenstein, 1994; Sadker & Sadker, 1995). Yet as these same studies underscore, the end result is often that boys get more attention, and girls get less feedback. Moreover, it is difficult to talk about boys and girls as discrete categories, given that girls living in poverty and girls of color often experience school in extremely different ways than those from middle-class white families and communities. As Orenstein (1994) astutely observed in her ethnographic study of working-class, minority girls at an urban middle school

> In the classrooms at Audubon, issues of gender are often subsumed by issues of basic humanity, often secondary to enabling a student - any student - to go through the school day without feeling insulted, abused, or wronged by her peers or by her teachers. (p. 137)

It is important to underscore that the word *gender* is often synonymous with *girls*, and therefore that gender equity is somehow a specialized concern that benefits one group over another. Further, there is a concern that these affirmative actions are often without merit, a concern which has become central to educational funding agencies which seek to distribute limited funds in the most democratic and responsible manner. Evaluation of such programs thus has sought to prove not only that the programs themselves are well designed and effective, but that the entire topic is worthy of **concern.**

- **Funding for Gender in Education Programs: Addressing "The Bottom Line"**

A number of compelling studies underscore that gender is as divisive an issue in the funding world as it is in the school reform and public arenas. In *Gender Matters: Funding Effective Programs for Women and Girls,* for example, Mead (2001) reports that:

> The bottom line is that funders have a strong preference for funding so-called universal (or coeducational) programs, and [have] little awareness of the need to consider gender when setting grant making priorities or allocating funds to grantees. (p. 3)

In her study of funders in the Greater Boston area, Mead identified a number of different rationalizations foundations use to resist focusing on gender. These include: *efficiency* --wanting limited resources to reach the broadest possible audience; *democracy* -- wanting programs to be as inclusive as possible; and*relevance* -- gender is not the most "critical criteria" for school reform (2001, p. 10-11).

Yet when Mead (2001) studied twenty-five so-called "universal" co-educational youth-development programs for urban teenagers, she found that gender was, in fact, an extremely relevant category. Mead found that these programs did not pay close enough attention to the different life experiences of boys and girls, and the ways that these experiences are shaped by gender norms (albeit these "norms" were further shaped by issues of race, ethnicity, and class). For example, Mead notes that women are significantly more likely than men to be living in poverty due primarily to "labor-market segregation and women's significantly greater role in raising children" (p.17). Thus a program that is concerned with issues of poverty or that seeks to assist poor people must consider these gendered components.

Mead (2001) also notes that because women and girls are socialized differently than men and boys, in mixed gendered groups they may be inclined to talk less, be "reluctant to engage in verbal conflicts" (p.17), or less likely to take leadership roles. Mead concludes by making a case for more *gender sensitive* programming rather than *universal gender-blind* programs. Programs may continue to be co-ed, although Mead suggests that in single-sex programs girls may avoid feeling like the "other," may feel more safe, and may have greater opportunities to exercise leadership abilities. Although Mead is not arguing that gender differences are innate or immutable, her findings underscore the conclusion that, "to be effective for women and girls, programs need to take gender into account" (p. 4).

Yet even foundations that have specifically committed to using gender as a guiding focus find this work interrupted by nagging questions of how to achieve their goals in a climate of universality, invisibility, and resistance to using gender as a specific programmatic lens. In the spring of 2000, The Three Guineas Fund (2001) conducted interviews with thirty-one funders and girls' program staff, bringing individuals from each group together for discussion. The resulting report, *Improving Philanthropy for Women and Girls*, provides recommendations for both groups and addresses some of the contradictions and dilemmas also raised in Mead's (2001) research.

The Three Guineas Fund (2001) reports that, "Foundations often focus on numbers of girls served and cost per girl" (p. 7). Yet program staff dispute that this is the most "effective" measure for evaluation. Instead, program staff advocate for "fewer girls served, smaller staff-to-girl ratios, and more in-depth programming" (p. 7). As one staff member describes it, "Large numbers typically have no long-term impact. When you're reaching 500 people, the impact is superficial" (p. 7). The Three Guineas Fund report also underscores that funders often have unrealistic expectations of evaluation results, expecting change to happen much more quickly than it usually does and that "funders do not often accept qualitative, including anecdotal evaluation measures" (p.7).

There are, of course, some practical reasons for this. As Mead (2001) rightly notes, "Foundations are both rational and irrational in their decision-making: they are influenced not only by carefully presented research evidence but also by internal and external pressures" (p. 6). In other words, foundations recognize that they need some measure of public and political support for their work, as well as to satisfy board members, donors, and other important stakeholders demands that they are making a real difference and using their resources wisely and productively. In many instances, grant-making decisions are limited by what Mead describes as a foundation's "history and culture," noting that "prior decisions and standard operating procedures influence and constrain available options and choices in the present" (p. 43).

And it is worth adding that those *seeking funds* are not blind to this reality. As one Executive Director of a non-profit candidly told me (when I interviewed him for my dissertation2 several years ago):

> When you go to big funders, you don't go for one grant. You put yourself up for adoption. You set-up a long term funding relationship. So you can come up with any good idea and can count on them for money. One reason for accountability is not just that public funds are adequately spent, but to maintain the continuity of the relationship with the funder….So the adoption proceeding goes through.

Though it sounds crude, this "adoption process" is no joke for struggling non-profits that are competing with large numbers of other organizations for an increasingly smaller pool of resources. This ultimately means that program developers and development officers need to design and "sell" programs that can be *easily proven* to be "successful," and thus merit more and future funding. In the current educational climate this means programs that reach large numbers of students and schools, raise test scores, produce "packaged" curriculums, are easily replicable, and are not particularly controversial.

This is in contrast to programs that may "fail" to produce products and raise test scores, yet can succeed in other ways. For example, programs which raise important new questions and insights about how students and teachers are experiencing school; highlight diverse perspectives including formally "silenced" voices; and, perhaps most importantly, teach us about what kinds of educational changes are superficial and what kinds of changes are meaningful and sustainable.

IV. Evaluating the Results of Gender Equity Programs in Urban Schools: What Constitutes Success?

Studies such as those summarized above underscore the ways in which gender is still a largely invisible, uncomfortable, contradictory, and misunderstood issue in school reform. Yet this has not stopped a wide variety of organizations -- ranging from large urban school districts and state agencies to the smallest non-profit community groups -- from designing programs to address gender inequities and raise awareness of the importance of paying closer attention to gender in schools. Although these programs have received some public attention, most of them are short-lived and have few paper trails.

In a relatively recent literature review of best practices in gender equity and education, commissioned by the Caroline and Sigmund Schott Foundation, Dr. Janie Victoria Ward, Director of The Alliance on Gender, Culture and School Practice at Harvard University, and her co-authors noted that it was very difficult to gather information about community-based programs that do gender work in schools because the majority of such gender-based programs operate after-school hours and are not aligned with the school's official curriculum, culture, or policies (Ward, Rotehnberg, Benjamin & Feigenberg, 2002).

These findings echo those revealed in similar research conducted by the Ms. Foundation for Women (2000) a decade earlier, in which the Foundation sought to "understand what it was about effective programs serving girls and women that made them work," and, in fact, to "prove" that they were working to benefit both girls and their communities" (p.1-2). After conducting an overview of such programs and convening interested stakeholders, the Ms. Foundation for Women concluded that:

> [T]he reality for most girls' programs is that they are not part of an explicit and intentional evaluation process….Most youth programs do not even have a budget for evaluation, which is typically considered either a luxury item separate from the 'real' work of girls' lives or a seemingly meaningless task required by funders. (p. 1)

Yet even when educational programs do have an explicit and comprehensive evaluation plan and budget, these evaluations are often, as the Ms. Foundation for Women (2000) suggests, driven by *accountability to funders rather than authentic opportunities for learning*. And the kinds of questions that funders and other stakeholders want answered, such as proof of increased student achievement and sustained changes in school culture, are often difficult to measure, as evaluators are frequently stymied by the reality of life in public (particularly urban) schools.

For example, programs that aim to provide gender-focused professional development for teachers and administrators around gender inequities are often difficult to evaluate because of inconsistent participation and heavy turnover of those practitioners involved. Such was the case with the Gender Awareness Through Education Program (GATE) developed by the Pennsylvania Humanities Council and funded by the Annenberg Foundation, Core States Bank, and the Arco Chemical Company. Without further probing and without considering the larger context in which these programs take place, it would be easy for evaluators to conclude GATE's low participation levels were due to disinterest in the program. But, as was the case in one such program, low attendance was not simply a reflection on the worth of the program. In all of the schools involved, participants were constantly changing jobs (often in non-linear ways, such as an art teacher who became the Dean of Students), retiring, transferring to other schools, or were simply unable to find a common meeting time given the myriad of additional responsibilities each was saddled with. Thus, any sort of quantitative measurements about how many people the program reached and how committed they were to its goals, needed to be qualified by qualitative, anecdotal evidence.

Participants in the above-mentioned program, for example, reflected upon the value of the program in highly positive terms. One noted that, "It's almost as if a consciousness in every word I say and how I present material to my students, even physical eye contact and movement, has changed." Teachers noted repeatedly that one outcome of the project was that that they were much more sensitive to their students' viewpoints and perspectives, and more able to engage them in classroom discussions and learning. This is particularly significant given that the schools involved were primarily comprised of poor and minority students - those who research shows are often the most alienated from school, and the most likely to dropout. One teacher noted in his

final reflection that, through his participation in the program, "I got insights into the way kids think, their view of the world. This is very important. I know we need to personalize education, understand their thinking."

Teachers also measured success in terms of the kinds of communities that were formed with other educators and whether or not such communities can/will be sustained after the official program is over. When asked directly, "How do you measure success in reform programs?" one GATE participant responded, "Firstly, by my relationships with other people in the group. Did I keep in touch with anyone in the group? [Are we] still in contact?" Another had a similar comment when asked what was the most important aspect of the participating in the program, "Being able to share with other teachers. We get very isolated." Yet another responded, "I think GATE is a very successful program because we're still talking about it" (Ginsberg, Shapiro & Brown, 2004, ch. 6).

Another problem frequently encountered in the evaluation process concerns how to measure the impact of the program *on students*. In other words, it's all well and good to have more insightful, sensitive teachers, but how do gender programs actually improve student performance and future achievement? Again, in theory this seems like an easy question to answer: Isolate those students involved. Test them and compile data from them at the beginning of the program and at selected intervals throughout. The reality, however, is far harsher. Just as teachers and administrators frequently change positions and responsibilities, the core group of students being affected by a particular program may also be changing constantly. In the program mentioned above, for example, the high school dropout rate averaged 60%. In such circumstances it is difficult, if not impossible, to develop any sort of effective "longitudinal study." And those students who did stay at the school over the entire program, often had sporadic contact with participating teachers, as evidenced by the comment of one participating teacher who explained in a written reflection:

> That teaching year I had five different groups of students. Four classes that I taught were special education students whom I saw on alternative days for English and History (one day I taught English and the next History for double periods, alternating subjects and days for the schools year). The fifth class I taught was a regular English class that also met every other day for two periods. All these classes were grouped heterogeneously by grade and ability.3

The program developers' idea of tracking students longitudinally beyond the pilot phase, while extremely worthwhile, proved to be even more infeasible. It would have been impossible to isolate those students affected only by this particular program and compare them to other groups of students. Perhaps most importantly, the kinds of changes initiated -- such as changes in students' self-esteem and career choices -- would be extremely difficult to discern from these kinds of multiple choice tests, with so many competing factors and in such a short time frame.

Similar problems are brought to light regarding the issue of documenting "sustained" change in school culture. Most urban schools do not have a single culture; rather different students experience school very differently depending upon their race, class, gender, family support systems, academic ability, which teachers they have, and other factors. Moreover, urban schools are continually in the process of reform. As Hess (1999) suggests:

> Not only are districts pursing an immense number of reforms, they recycle initiatives, constantly modify previous initiatives, and adopt innovative reform A to replace practice B even as another district is adopting B as an innovative reform to replace practice A. (p. 5)

Just as it is difficult to isolate those particular students being affected by a special program, it is equally difficult to isolate the impact of one particular program within the context of a myriad of other changes. The atmosphere of instability that is common in urban schools cannot be taken for granted in the evaluation process. During the course of the program discussed above, for example, the District's Superintendent resigned almost immediately after the program began, and a new Superintendent was hired bringing an ambitious, district-wide schools reform plan of his own. The new Superintendent further made it known that he was not in favor of "pilots and demonstration projects." This not only meant that teachers were not rewarded or recognized for the extra time they essentially volunteered to the gender program, but also that they were faced with a multitude of other reforms and expectations -- some of which were directly at odds with gender awareness work. This Superintendent would leave a number of years later, and the district would eventually be taken over by the state, bringing the entire district into a state of flux and uncertainty.

Shapiro (2004) has suggested that accountability often becomes a question of who to *blame*, rather than how to find workable solutions and create collaborative coalitions. To a certain extent, any accountability system is flawed in that there are many factors that educators and students simply cannot control. These include poverty and racism, as well as the fact that different stakeholders are often working towards different overall goals and objectives. Hargreves (1994) has spoken of this as the difference between real collegiality and "contrived collegiality."

Although participating teachers' evaluations of the GATE program were highly positive, most indicated a lingering disappointment that the program did not accomplish something more "concrete," something more "replicable." Some participants lamented that they were not able to interest many other teachers and school administrators in the work they were doing around gender, and that while they themselves had changed considerably, the school as an institution remained basically the same. As one teacher said in an end of program interview, "Did it change the school? I would say on a scale of 1-10, maybe about a 3." Another noted similarly, "The discussion died with me at the end of the year. I couldn't communicate to other teachers how to talk about these issues."4

Program developers and funders expressed a variety of similar concerns about its ultimate "payoff." Some comments in this regard included, "I don't have a sense of how much was got out [of it]. I don't have a measure of translation to the classroom," and "We need experiments that can be more easily translated into replicable programs."5 The underlying messages, commonly heard in education and evaluation circles, are:

> 1) The most important role of evaluation is to measure the end product, as opposed to raising new insights and
> questions about important and complex issues, providing a forum for educators to discuss and debate; and
> 2) Programs that are not easily packaged and replicated are not worthwhile investments for funders.

Thus, the question remains: *What would a gender program evaluation look like where success was measured by the amount of reflection, inquiry, discussion, and genuine learning rather than simply "The Bottom Line"?*

V. New Models of Evaluation, New Definitions of "success," New Juxtapositions of Qualitative and Quantitative data

There is, of course, a rich history of qualitative evaluation, participatory evaluation, ethnographic evaluation, and teacher/action research (Anderson, Herr & Nihlen, 1994; Cochran-Smith & Lytle, 1993; Shapiro, Parssinen & Brown, 1992) which often stands in stark contrast to evaluation based only on "statistics" and "test scores." Although programs which address "women's" issues have become more plentiful since the Ms. Foundation for Women first began such funding in the early 1980's, there still isn't an extensive body of evaluation literature that focuses explicitly on gender issues in schools -- especially in ways that do not simplistically set boys and girls up in opposition to each other (Skelton, 2001).

This last point is particularly important, because, as Skelton (2001) has emphasized, "the complexity of gender and achievement cannot simply be 'read off' crude, basic data" (p. 165). In considering the experiences of girls and boys in schools, one must take a "relational" approach as opposed to an "essentialist" approach. A relational approach understands that boys and girls construct their own cultural identities, based on differences in race, class, ethnicity, and other differences, and that these identities are not fixed but rather different aspects are more prominent in different contexts. The "essentialist" notion that "boys are boys" (Gurian, 2003) does not work in practice. Educators need to consider complex examples of mascul*inities* and femin*inities*, meaning that: 1) boys and girls are studied in connection and interaction with each other, not as static beings, but as people who are always interacting with their environment; 2) issues of race, class, ethnicity, sexual identity, and religion are carefully considered as the data is collected and disaggregated; and 3) ideally the students themselves help to collect and analyze data based on their own (ever-changing) perspectives and experiences. This stands in stark contrast to both the testing approach and also to a hegemonic view which does not consider issues of socialization and power differences.

The question of what an alternative gender evaluation might look like was tackled by the Ms. Foundation for Women (2000) when they convened a group of like-minded foundations and donors to form the Collaborative for Healthy Girls/Healthy Women. The Ms. Foundation for Women set about creating a special fund for the development and support of programs to increase girls' self-esteem, leadership, community activism, and achievement in education. The Collaborative created a $4 million dollar fund to "provide resources over three years to new and existing organizations with programs focused on girls' empowerment and activism" (p.2). The extremely diverse groups that ultimately received funding included the Asian and Pacific Islanders for Reproductive Health (Long Beach, CA); the Center for Anti-Violence Education (Brooklyn, NY); Mi Casa Resource Center for Women, Inc. (Denver, CO); and Native Action (Lame Deer, MT), among others.

In March 1999, the Collaborative organized the Young Women's Action Team (YWAT), a representative group of girls drawn from six of the grantee programs, who worked alongside young women scholars to develop research questions and evaluation tools for the entire Collaborative. A central question that emerged from this group was, "*How does being in a girl-centered program impact girls' lives?*" This was a question that they wanted to answer on a number of different levels: 1) the individual level; 2) the social network level; 3) the community level; and 4) the institutional level. In other words, the focus was not exclusively on academic achievement as we have come to view it through the lens of standardized testing. While individual achievement was an important outcome to be evaluated, the group also wanted to understand what made girls leaders and what enabled them to pursue social justice policy and programs for the betterment of the entire communities and schools.

Using a *participatory evaluation research approach*, in which the girls themselves played a central and on-going role in evaluating the programs they participated in, several exciting new evaluation tools were developed and tested. Two of the most interesting, described herein, include the *Voice, Action, Comportment, and Opportunity Checklist (VACO)* and *The International Storytelling Measure (ISM)*.

These were discrete, short-term programs that mostly took place outside of regular school hours, and that included a relatively steady group of girl participants and adults. These evaluations do not address the question of how to do similar work within the context of schools where just about the only thing that stays steady from year to year is the cafeteria food. Nonetheless, these evaluation tools may be widely considered as examples of evaluations that do more than measure superficial changes, and are, in and amongst themselves, opportunities for learning and reflection.

1) VACO

VACO was developed as a method of measuring and describing "incremental change in girls' leadership skills and qualities" (Ms. Foundation for Women, 2000). The Ms. Foundation for Women defines VACO as:

> **V Voice:** girls' ability to speak on their own behalf
> **A Action:** girls' ability to use their voices to act on behalf of themselves and others
> **C Comportment:** girls' ability to carry themselves with pride, respect and dignity
> **O Opportunity:** girls' ability to ask for and take advantage of new changes and experiences
> (p. 1)

The VACO evaluation includes both a pre-test and a post-test, although its primary purpose is to *chronicle girls' development as it happens day to day*. Staff members pick a minimum of six girls to observe during program activities, taking notes on the different ways that each girl uses her voice and interacts with others. At the end of each observation period, staff complete a VACO checklist for each of the four VACO categories being observed, noting the existence of certain behaviors such as:

- **Voice:** Challenged another girls' opinion; stated and defended a point of view or idea; expressed analysis of injustice, discrimination or prejudice; struggled to say something hard about herself in a group.

- **Action:** organized others to engage in activities without being told; stopped conflict between other girls; resisted pressure from others to go along with something.

- **Comportment:** looked directly at others; paid attention to the facilitator; listened to peers in the group.

- **Opportunity:** suggested ways to find more resources; volunteered to do something that is new or challenging; asked to have more responsibility.

The developers of VACO stress that the evaluation, though clearly not quantitative in nature, can provide "statistical documentation" of girls' development as leaders within the various programs they participate in. For this to work, however, certain guidelines must be followed. These include observing the same girls at different points in time and ensuring that the same staff member observes the same girls each time to maintain consistency, among other guidelines. They also stress that VACO should be used with girls in the first week that they begin attending the program in order to produce a clear measure of change.

What makes VACO especially exciting, however, is that the girls themselves are a critical part of collecting and triangulating (e.g., cross referencing) the data. For example, at the end of each observation period, staff and girls separately complete the same checklist. Not only does this provide a way for staff *and girls* to compare their observations (thus triangulating data and challenging discrepancies in language and awareness), *but it also provides an excellent tool for the girls' own self-reflection*.

What would it mean for an approach such as VACO to be incorporated into a school's portfolio of assessment tools, to be integrated into the classroom on an ongoing basis, much like standardized tests? The answer to this question is premised on the idea that actions can be as meaningful as "words", and further, that knowing a lot of "facts" does not necessarily translate into important skills like imagination, responsibility, organization, curiosity, risk-taking, open-mindedness, and leadership. Thus, as we evaluate students' achievement in school -- whether they be boys or girls -- we need to consider evaluation tools that "measure" these skills, and furthermore, that ask students to assess themselves as a means for further development. Clearly this is a much more subjective and time-consuming project than testing, but is nonetheless, critically important if we really want to understand how to empower youth to be social leaders of the future.

2) ISM

Another Ms. developed evaluation, ISM, stands for the Intentional Storytelling Measure. In some ways, this is a kind of standardized "test," though certainly not one in which there is only one "right" answer. The exercise, which involves girls reading and responding to a number of hypothetical problems, is designed to "see whether girls perceive themselves as capable of acting as agents for change in relationship to their peers, families and communities" (Ms. Foundation for Women, 2000). Students are asked to brainstorm a number of possible solutions, to choose the one that they think is "ideal" and to describe the one which they think they would most "realistically" pursue. As the girls come back to the same stories over time, the ISM helps answer the question, *What is the effect on girls and on their communities of their involvement in social change work (e.g., organizing, community service, policy advocacy, and community activism)?*

The stories themselves vary widely. One story focuses on a group of girls who notice that their friend Tina is being physically abused by her boyfriend. As the situation gets more out of control, Tina tells her friends that "it's no big deal," and to "mind their own business." Another story chronicles a group of girls who, after passing an empty lot filled with junk everyday in a neighborhood with no parks or playgroups, begin to think of ways that they could change the situation. A third story concerns a group of girls who go to a school board meeting to make a presentation about developing a new program in the school addressing sexual harassment. The girls are subsequently sexualized and belittled by the male head of the school board who says: "Thank you, beautiful girls, for providing us with such a treat" and then gives them advice about "looking for boyfriends" (Ms. Foundation for Women, 2000).

As previously noted, after reading each of the stories, the girls are asked to brainstorm possible actions and solutions, including those that are most ideal, and those that are most realistic. The girls then come back to the same stories later in the program, with the ultimate goal of seeing in what ways their responses changed as a result of their participation in the program. It is cautioned that the *same stories* must be used for pre- and post-testing, as different stories are not equivalent. It is also cautioned that when brainstorming, girls should not be prompted towards particular solutions.

A t the end of the program, the girls' responses to the stories are carefully coded according to pre-determined guidelines, depending upon what the program is trying to accomplish (e.g., the skills or qualities that it is hoped the program will impart to girls). The Collaborative suggests that, for the sake of consistency, it is useful to have two people code the same data to determine inter-rater reliability.

Students need to be able to make what are often difficult decisions about how to solve problems for which *there is clearly more than one right answer*. This means considering not only the ideal solution, but those that are most realistic, and those that are most ethically compelling for them personally. Moreover, this evaluation tool stresses the need for evaluation to be an ongoing process where students have the ability to change their answers without suffering penalties.

VACO and ISM are just two examples of an entire package of tools developed by the Collaborative. As noted above, these evaluation tools would need to be reconstructed if they were going to work within a larger and more chaotic school-based environment. Nonetheless, the key components inherent in each are critical. These include:

1) The idea that the participants (e.g., girls) themselves should be intricately involved in the evaluation process;

2) The importance of consistent and intentional gathering of evaluation data, as well as the triangulation of data; and,

3) Addressing and evaluating the real issues the programs are trying to address, and the real contexts in which the programs are taking place. In other words, the evaluation should serve a greater purpose than simply proving the successes or failures of a program, but should serve as an on-going and meaningful tool for self-reflection, problem solving, and relationship-building.

This approach may be termed "feminist assessment," which Shapiro (1995) defines as a form of assessment which "assumes questioning is expected regarding all forms of assessment and evaluation of their ultimate uses" (p. 98). In other words, assessment is part of an ongoing process of inquiry, through which learning generates new questions rather than simplistic answers.

While this process may appear to some critics to be too subjective to count as a real evaluation, as the Ms. Foundation for Women (2000) stresses, "scientific documentation and participatory self-reflection do not have to be at odds with each other" (p. 27). The important thing is to make the evaluation process an authentic one for those involved. According to Ms. and the

Collaborative:

> Evaluation research does not have to be academic, formulaic, or bureaucratic. Rather, it can be fun and engaging even as it legitimates and empowers our work. Evaluation research can provide us with the real and powerful results of working with young people to change the world. It is also a tool we can make our own and translate into a language that bridges the gaps between age, culture, and experience. (p. 27)

Even so, there are legitimate questions as to how such approaches will be received in an educational climate that is obsessed with numbers (e.g., numbers of participants served, numbers of dollars expended, numbers of test scores, etc.). Even those foundations and other funding agency staff that are interested in an alternative method of assessment of gender-based school reform must give practical consideration to the stakeholders who oversee and judge their work and decision-making.

VI. Resurfacing Questions for Program Funders and other Stakeholders

This paper, while grounded in contemporary educational research, also reflects upon my own experiences as a program officer at a non-profit educational organization, as the director of a professional development program piloted in a large urban school district, as an evaluator of many different gender equity initiatives, and as a consultant to several national foundations seeking to support programs that address gender issues in schools. Indeed, the impetus for this work comes from attending many roundtables, focus groups, and advisory meetings comprised of educational and gender experts where I found that the same compelling questions kept resurfacing.

Just as program developers struggle with the language they use to describe their programs, most foundations also spend a considerable amount of time designing mission statements, funding goals, and request for proposals. Those foundations that want to help girls in particular, struggle with whether they should state this overtly in their mission statement and request for proposals. The alternative, to use more neutral language like "gender" (which is more inclusive of boys) or even better "diversity" (which is inclusive of everyone), is more likely to receive widespread public support and avoid criticism that the foundation is favoring one group over another.

On the positive side of switching to language from girls to gender or diversity is the recognition that biological sex is not the only issue at stake. Rather, the focus here is on the construction of norms and roles for girls and boys, and the different life experiences and opportunities that result from them. The negative side of this language shift, is that, as Mead (2001) suggests, programs that don't specifically target girls threaten to subsume their needs, voices and perspectives, within a broader more universal (e.g., white male) framework.

Another related question that foundations struggle with is what is meant by the term *equity*? Of course, once one starts down this road, the inevitable questions regarding the relationship among gender, race, class, ethnicity, and other differences arise. When we talk about girls as a "disadvantaged group" in need of special programs and consideration, this is, of course, a relative term. Not all girls are disadvantaged in the same way. Obviously girls of color and poor girls face more complex and daunting challenges than middle-class white girls. Many go as far as to argue, with good reason, that *African American* boys are actually the most at-risk group. As one participant in a professional development program designed to raise awareness about gender inequities admitted in an open-ended written reflection:

> An early struggle at my setting…was that many staff members believe that African-American males are so deeply at-risk that it is superfluous and academic elitism to devote professional development time to exploring sexism in our classrooms.

Indeed, in a series of recent interviews I conducted with education, gender, and policy experts across the country[6] -- such as presidents of schools of education and research institutes, executive directors of foundations and women's funds, and leaders of political advocacy organizations -- a significant number felt that gender issues paled in comparison to the much more tangible tragedies of racism and poverty. These leaders persistently pointed to the abundance of African American and Latino youth in poor urban communities with no positive role models, inadequate school funding, decrepit and unsafe school buildings, large numbers of uncertified teachers, and alarming dropout rates.

Of course, it is important to point out that at least 50% of these children referred to above are *girls*, and, moreover, that boys in these communities are also damaged and constrained by unhealthy and rigid definitions of masculinity, etc. Nonetheless, the

pervasive belief was that gender was simply not the dominant concern. This leads many foundation board members and staff members to question how they can support change and build broad-based advocacy around what many people feel is a low priority issue.

The following questions are raised repeatedly:

- Should we start where people (e.g., politicians, administrators, funders, teachers, parents, etc.) *are* in their thinking and understanding of the issues, or where we *want them to be*?
- If girls and gender remain the focus of the work, how much attention (if any) should one pay to gender in isolation from other forms of difference?
- In order to adequately and genuinely address issues of race and class, etc. should different groups of girls (or boys) be targeted as higher priorities for funding?
- Should we try to address all these issues at once and cater to all groups in need, or will we end up simply diluting our resources to the point in which we are effective for no one?

These are critically important and very difficult questions to answer.

Another series of questions frequently raised in these meetings concerns whether to *build on existing work or create new work* The continuing impetus to create new programs reflects not only a lack of awareness about the important work already being piloted by women's organizations, community groups, school districts, and other reform agencies across the country, but an impatience to see concrete and large-scale changes as soon as possible. In "A Conversation About Girls" (Valentine Foundation, 1990), a series of conversations among funders, researchers and gender equity advocates, organized by the Valentine Foundation and Women's Way, it is stressed that, "Many programs labeled a failure have simply not been funded long enough for the girls to achieve the program goals. This is a financing problem, not one of program design or inability of girls to respond" (p. 4). This dilemma thus circles back to the fundamental issue of this essay: gender and program evaluation.

- What kinds of measures of change do we look for and recognize as important and valid?
- How quickly do we need to see documentation of these changes?
- Who is the primary audience for these evaluations?

Since funding agencies are faced with the difficult question of how they will evaluate the impact of the programs they fund and make decisions about future funding based on internal and external evaluations, many program directors fear that any negative evaluation will automatically kill future funding of their programs. It is thus not uncommon for them to reduce the complexity and sophistication of the evaluation tools to ones in which only the positive (and universally recognized) outcomes are highlighted. Other programs do not have the staff or the know-how to conduct a formal evaluation of their work and are unable to find the necessary support (financial and human) to make this work a priority (Council on Foundations, 1993). And even for those groups that are developing and using cutting-edge evaluation tools, such as the Collaborative Fund for Healthy Girls, Healthy Women, it is not yet clear how these kinds of evaluations will hold up when push comes to shove (e.g., when large sums of money, resources, and political good-will are at stake).

This is one of the reasons why, when we think about the relationship between funding and evaluation, we need to think beyond the traditional models of philanthropy. In "A Conversation About Girls" (Valentine Foundation, 1990), the authors stress that one of the challenges for foundations supporting this kind of work is learning to work collaboratively. They also note that it is important to "include the young girls [or boys] themselves within the process of planning and designing programs," further underscoring that "We must not let ageism interfere with the design of programs that may be more effective than those we ourselves design" (p. 7). An example of a foundation acknowledging such work is as the Michigan Women's Foundation which sponsors young girls, training them in philanthropy, and then has them develop RFPs (Request for Proposals) and awards grants equaling $20,000 per year.

The "Conversation About Girls" (Valentine Foundation, 1990) raises another important point:

> The resources of private foundations alone, however, are not sufficient. The larger task lies with legislative bodies that must exercise corporate responsibility, addressing in concrete programmatic and financial terms problems that affect the life and health of the entire society. (p. 6)

In other words, while focusing on gender equity in urban education and while making sure this work genuinely includes boys as well as girls, we need to look beyond schools themselves, understanding that this is a social justice issue, and not a passing educational fad. Funding agencies interested in supporting equity programs are constantly being asked to explain and defend why such programs are important, and how they can make a real difference in real students' lives and in their urban communities

as well. This is most especially difficult when the pressure is to focus on the three R's as early as preschool, preparing children at younger and younger ages to pass the narrowly designed tests which have become so central to our entire educational system.

Finally, it is useful to point out and remember that addressing gender in education means bringing a variety of stakeholders to the table, not just in the early planning stages or as underutilized "advisors" but as real participants throughout the entire program; paying attention to *process* as well as *product*, meaning that teaching and learning go beyond memorizing facts, and that teachers and students have complex relationships which make gender study messier but significantly richer; and always having larger goals in mind, (e.g., moving beyond the "add women and stir" approach towards one where gender becomes an ongoing critical lens of analysis, as does race, class, ethnicity, and other differences in human identity and experience). The Council on Foundations (1993) further adds some very useful suggestions to this list, including that participants "start early," planning the evaluation as an integral part of the program itself; "[M]ake sure your evaluation plan has built-in flexibility," which "allows for change or expansion in midstream if the evaluation data begins to show an important new direction for inquiry"; "[h]ave respect for previous work," as "a good evaluation builds on existing knowledge," and whenever possible, "[i]nvolve several funders and pool resources" (p. 252-265).

Despite not making the official list of reform priorities, showing up on the public radar screen, or addressing the "bottom line," the very fact that these questions are continuing to be raised and taken seriously provides a measure of hope that the issue of gender equity in education will not simply "go away."

References

American Association of University Women. (1992). *How schools shortchange girls: A study on major findings on girls and education.* Washington, D.C.: Educational Foundation and National Education Association.

American Association of University Women/Research for Action. (1996). *Girls in the middle: Working to succeed in school.* AAUW Report: Washington, D.C.

Anderson, G., Herr, K., & Nihlen, A.S. (1994). *Studying your own school.* Thousand Oaks Press, CA: Corwin.

Bierda, M. (2000). The myth of the African American male. *WEEA Digest, 9(10).*

Cohen L., & Manion, L. (1984). Action research. In J. Bell, T. Bush, A. Fox, J. Goodey and S. Goulding, (Eds.), *Conducting small-scale investigations in educational management* (pp.41-71). London: Harper and Row.

Connell, R.W. (1993). Disruptions: Improper Masculinities and Schooling. In Weiss, L. & Fine, M. (Eds.), *Beyond Silenced Voices: Class, Race, and Gender in United States Schools* (pp.191-208). New York: State University of New York Press.

Council on Foundations. (1993). *Evaluation for foundations: Concepts, cases, guidelines and resources.* San Francisco: Jossey-Bass Publishers.

Davis, J.E. (2000). Mothering for manhood: The (re)production of a black son's gendered self. In C. Brown & J. Davis (Eds.), *Black sons to mothers: complements, critiques and challenges for cultural workers in education* (pp. 51-70). New York: Peter Lang.

Finn, P. (1999). *Literacy with an attitude: Educating working class children in their own self interest.* Albany: State University of New York Press.

Flood, C. & Dorney, J. (1997). Breaking gender silences in the curriculum: A retreat intervention with middle school teachers. *Educational Action Research, 5(1),* 71-86.

Francis, B. (2000). *Boys and girls and achievement: Addressing classroom issues.* London: Routledge/Falmer.

Hall, P. (1997). "Epilogue: Schooling, gender, equity and policy" In Bank & Hall, (Eds.), *Gender equity and schooling.* New York: Garland Publishing.

Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the post-modern age.* New York: Teachers College Press.

Ginsberg, A., Shaprio, J.P. & Brown, S.P. (2004). *Gender in urban education: Strategies for student achievement.* Portsmouth, NY: Heinemann Press.

Grady, J. & Auburn, A. (2000). Talking gender equity and education: A FrameWorks message memo. FrameWorks Institute. Report commissioned by the Caroline and Sigmund Schott Foundation: Cambridge, MA.

Hess, F. (1999). *Spinning wheels: The politics of urban school reform.* Washington, D.C.: Brookings Institution Press.

Leadbeater, B. & Way, N. (1996). *Urban girls: Resisting stereotypes, creating identities.* New York: New York University Press.

Martinez, E. (1995). Distorting Latino history: The California textbook controversy. In Levine, Lowe, Peterson & Tenorio (Eds.), pp. 100-108. *Rethinking Schools: An Agenda for Change.* New York: The New Press.

Mead, M. (2001). *Gender matters: Funding effective programs for women and girls.* Tufts University: Medford, MA.

Ms. Foundation for Women. (2000). *The new girls' movement: New assessment tools for youth programs.* Report and evaluation tool kit.

National Center for Schools and Communities. (2002). Unlocking the school house door: The community struggle for a say in our children's education. Fordham University: New York City, NY.

Ogbu, J. (1987). *Minority education and caste: The American system in cross-cultural perspective.* New York, NY: Academic Press.

Obgu, J. & Simmon, H. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology of Education Quarterly 29(2),* 155-188.

Orenstein, P. (1994). *School girls: Young women, self-esteem and the confidence gap.* New York: Doubleday.

Research for Action. (1996). *Girls in the middle: Working to succeed in school.* Washington, D.C.: American Association of University Women Educational Foundation.

Sadker, M. & Sadker, D. (1994). *Failing at fairness: How our schools cheat girls.* New York: A Touchstone Book.

Shapiro, J., Parssinen, C. & Brown, S. (1992). Teacher-Scholars: An action research study of a collaborative feminist scholarship colloquium between schools and universities. *Teacher and Teacher Education, 8(1),* 91-104.

Shapiro, J.P., Sewell, T.E., & DuCette, J.P. (1995). *Reframing diversity in education.*
Lancaster: Technomic Publishing Company.

Skelton, C. (2001). Typical boys?: Theorizing masculinity educational settings. In B. Francis, & C. Skelton (Eds.), *Investigating gender: Contemporary perspectives in education.* pp. 164-176. Philadelphia: Open University Press.

Three Guineas Fund. (2001). *Improving philanthropy for girls' programs.* Three Guineas Fund: San Francisco

Ward, J.V., Rothenberg, M., Benjamin, B.C. & Feigenberg, L. (2002). *Results from focus groups and interviews with gender*

equity consultants. Report commissioned by the Caroline and Sigmund Schott Foundation: Cambridge, MA.

Valentine Foundation. (1990). A conversation about girls. *Conference Proceedings.* Philadelphia, PA.

Endnotes

1-These interviews were part of a consulting assignment for a New England Foundation that was in the process of refining its gender equity initiatives. I interviewed approximately twenty top educational policymakers from diverse organizations across the country. back

2-Collected during research for dissertation: Ginsberg, A. (1999). *When Policymakers and Practitioners Partner: A Stakeholder Analysis of an Urban School Reform Program.* Philadelphia: University of Pennsylvania, Graduate School of Education. back

3-From GATE participants' monthly seminar reflections. back

4-From GATE participants' monthly seminar reflections. back

5-Collected during research for dissertation: Ginsberg, A. (1999). *When Policymakers and Practitioners Partner: A Stakeholder Analysis of an Urban School Reform Program.* Philadelphia: University of Pennsylvania, Graduate School of Education. back

6-see footnote 1 back

Alice Ginsberg, Ph.D.

Alice Ginsberg is an education and gender equity consultant whose recent client list includes The Caroline and Sigmund Schott Foundation (Cambridge, MA) and The Ms. Foundation for Women (New York, NY). From 1990-1998 she was a program officer at the Pennsylvania Humanities Council, where she developed and directed the GATE (Gender Awareness Through Education) program - a three-year professional development program in the School District of Philadelphia. Ginsberg holds a B.A. in women's studies from Temple University and a Ph.D. in Education from the University of Pennsylvania. Her dissertation, When Policymakers and Practitioners Partner: A Stakeholder Analysis of an Urban School Reform Program, explored the process and politics of a (gender-based) school reform initiative from the diverse perspectives of the teachers, administrators, school district officials, parents, students, academic facilitators, program developers, and funding agencies involved. In addition to her other publications in journals such as *Teaching and Learning*, *Current Issues in Comparative Education (CICE)*, *Women's Studies Quarterly*, *Lilith*, *New Directions for Women,* and *The Temple Review*, she is the first co-author of *Gender in Urban Education: Strategies for Student Achievement* (Heinemann, 2004), and is currently collaborating on an anthology of writings on gender and educational philanthropy. She can be reached at aliceginsberg@yahoo.com.

Report accessibility issues and request help

Source URL:https://urbanedjournal.gse.upenn.edu/node/168