

Analysis of Report of Wood & Sherman, Addendum to National Association of Scholars Amicus Brief

By

Ewart A.C. Thomas and Richard J. Shavelson
Stanford Institute for Higher Education Research
Stanford University

June 12, 2001

I. Introduction

A. Qualifications

Ewart A.C. Thomas is professor of Psychology at Stanford University, past dean of the Stanford School of Humanities and Sciences, and past chairman of the Stanford Department of Psychology. He received his Ph.D. in statistics from Cambridge University, England, in 1967.

Richard Shavelson is a professor of Education at Stanford University. He served as immediate past dean of the Stanford School of Education, as well as chair of the National Academy of Sciences' Board on Testing and Achievement. His research examines accountability and the assessment of student outcomes. Professor Shavelson earned his Ph.D. at Stanford University in Educational Psychology.

B. Information Considered in Forming Opinions

The Wood & Sherman Report (hereafter referred to as "Wood & Sherman"), attached as an addendum to the *amicus curiae* Brief of the National Association of Scholars ("NAS Brief"), examines the research conclusions about the effect of affirmative action on student outcomes from a number of studies including the "Expert Witness Report of Patricia Y. Gurin" (Gurin, 1998, hereafter referred to as the "Gurin Report").

In this report, we analyze the Wood & Sherman critique of the Gurin Report, and the theoretical and empirical analyses of the CIRP data in the Gurin Report. In preparation for this report, we have reviewed: (1) the Wood & Sherman Report, (2) the Gurin Report, (3) the Gurin supplemental responses in the Gratz litigation, (4) the Gurin Response to the Critique by the National Association of Scholars of the Expert Witness Report of Patricia Gurin posted on the University of Michigan website on May 30, 2001; (5) the district court's decision in *Gratz v. Bollinger*; (5) Justice Powell's opinion in *Bakke*; and (6) Argument I of the NAS Brief.

C. Summary

Our analysis is principally devoted to examining the fundamental conceptual difference between the methodologies of Gurin and those of Wood & Sherman—Wood & Sherman’s contention that indirect effects are irrelevant to a social scientific analysis of the effects of a diverse student body in an academic setting containing programs designed to enhance learning from a diverse student group. We conclude by assessing the scientific merits of the Gurin Report and the other principal elements of the Wood & Sherman critique.

We conclude that the Gurin Report follows an appropriate methodology for analyzing the effects of student racial (structural) diversity on educational outcomes in a higher education setting, and that this approach is fully consistent with conventional academic approaches to measuring the effects of such a variable on educational outcomes. It presents methodologically valid evidence that structural diversity, when combined with appropriate campus experiences, is a necessary condition to producing certain improved educational outcomes.

By contrast, the Wood & Sherman Report incorrectly rejects Gurin’s analysis of indirect effects, dismisses the relevance of important variables that she studies, and takes issue with standard social scientific methods that Gurin follows. As a result, they fail to address Gurin’s major point, and fail to consider relevant information for assessing whether structural diversity, in the context of programs at the University of Michigan, affects campus outcomes. These errors are central to Wood & Sherman’s attack on Gurin’s methodology, and fatally undermine their claim that Gurin has failed to provide sufficient evidence to support the district court’s factual findings concerning the benefits of diversity.

II. Analysis

A. Whether It Is Appropriate to Analyze Indirect Effects of a Diverse Student Body

The core of the methodological disagreement between the Gurin Report and the Wood & Sherman Report concerns whether it is appropriate to analyze indirect effects of a racially diverse student body. The Gurin Report examines how the University of Michigan’s affirmative action program operates in practice on campus. If structural diversity is present, it becomes an educational resource that needs to be translated into action to produce educational benefits. However, structural diversity is a necessary condition for those benefits.

The Gurin Response to the NAS Critique found on the University of Michigan’s website offers an instructive analogy in this regard. Cigarette smoking causes cancer through its effect on lung tissue. However, if one conducts a [regression] analysis that takes destruction of lung tissue into account as an independent variable controlling for all other variables, the direct effect of cigarette smoking on cancer disappears. Similarly, in

Gurin's analysis of education data, the direct effect of structural diversity on outcomes disappears when diversity experience enters the model because diversity experience appears as the causal mechanism. Structural diversity is a necessary condition that gives rise to the possibility of diversity experiences, and these experiences in turn affect outcomes. Simply having greater diversity alone does not guarantee positive outcomes. What is important is that racially diverse students (structural diversity) have common, positive experiences (diversity experience) and these experiences then produce positive educational outcomes.

The Wood & Sherman Report challenges this approach, and examines only whether racial diversity, viewed in isolation and without consideration of contextual factors, affects academic outcomes. This critique does not undermine the value of the evidence offered through the Gurin Report.

1. Analysis of Indirect Effects Is Consistent with Justice Powell's Opinion in *Bakke*

Wood & Sherman make a central, and in our view, incorrect assumption that evidence that student racial diversity (structural diversity) is necessary to achieve a compelling governmental interest can be established only through examination of its direct effects on educational outcomes. They claim that "the Powell diversity rationale asserts that, holding all other variables constant, students benefit by attending racially diverse colleges." (Wood & Sherman, at 53). By "holding all variables constant," they exclude and ignore effects on educational outcomes produced by students' diversity experiences and other educational activities that depend on the existence of structural diversity.

As we demonstrate below, examining indirect effects of a variable is an entirely legitimate social science methodology. Moreover, although Wood & Sherman call their approach "the Powell diversity rationale," their assumption that indirect effects should not be considered appears inconsistent with the language of Justice Powell's opinion in *Bakke*.

Justice Powell's opinion nowhere indicates that indirect effects of structural diversity are less relevant than direct effects, much less that indirect effects should not be considered. In fact, it appears particularly concerned with what social scientists term indirect effects. Wood & Sherman quote selectively from *Bakke*, but do not include in their report a footnote in which Justice Powell quotes at length the President of Princeton University describing "some of the benefits derived from a diverse student body." These examples all concern informal learning and student interaction.¹

¹ "[A] great deal of learning occurs informally. It occurs through interactions among students of both sexes; of different races, religions, and backgrounds; who come from cities and rural areas, from various states and countries; who have a wide variety of interests, talents, and perspectives; and who are able, directly or indirectly to learn from their differences and to stimulate one another to reexamine even their most deeply held assumptions about themselves and their world. As a

For this reason, social scientific methodology such as the one in the Gurin Report examining how structural diversity combines with campus experiences to affect educational outcomes appears entirely consistent with Justice Powell's opinion.

2. Diagram Demonstration

A convenient starting point for understanding both the Gurin Report and the Wood & Sherman critique is the Gurin Report's Figure 1--the "path diagram" (slightly revised in our Figure 1 below). This diagram depicts the hypothesized relationships among background and institutional variables, diversity variables, and student outcomes.

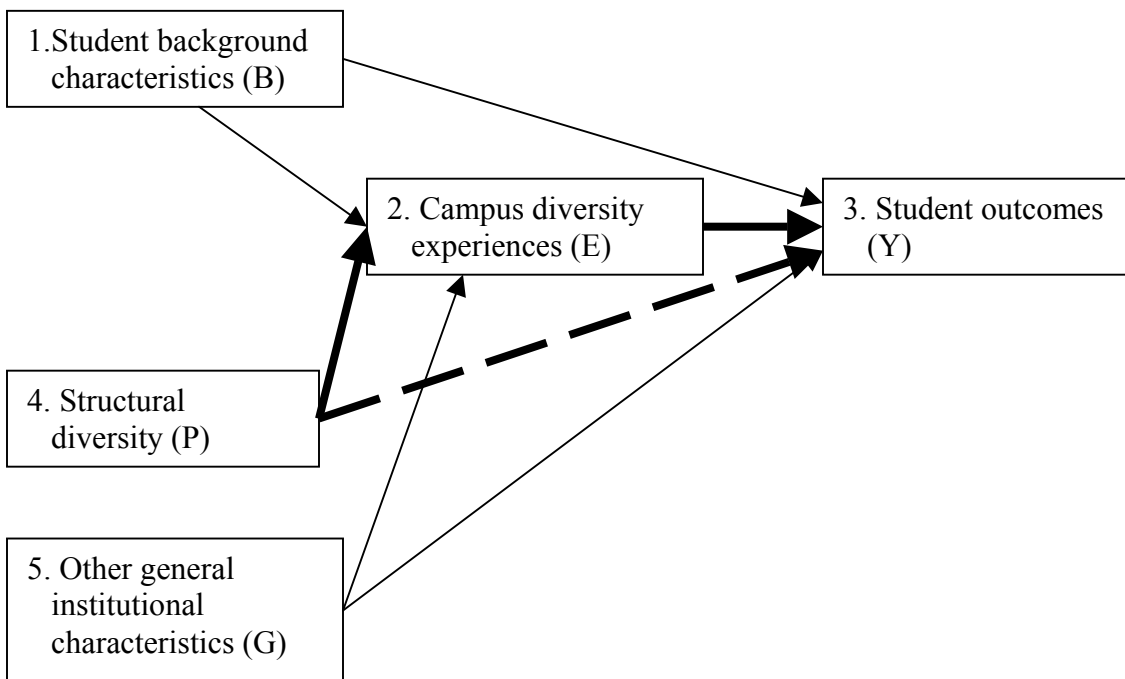


Figure 1. Hypothesized relationships among background and institutional variables, diversity variables, and student outcomes.

wise graduate of ours observed in commenting on this aspect of the educational process, 'People do not learn very much when they are surrounded only by the likes of themselves.'

* * *

In the nature of things, it is hard to know how, and when, and even if, this informal 'learning through diversity' actually occurs. It does not occur for everyone. For many, however, the unplanned, casual encounters with roommates, fellow sufferers in an organic chemistry class, student workers in the library, teammates on a basketball squad, or other participants in class affairs or student government can be subtle and yet powerful sources of improved understanding and personal growth." Bowen, Admissions and the Relevance of Race, *Princeton Alumni Weekly* 7, 9 (Sept. 26, 1977).

Regents of the University of California v. Bakke, 438 U.S. 265 (1978), at 312 n.48.

The variables in boxes 1, 4 and 5 are presumed by Gurin to influence the variables in boxes 2 and 3. As noted above, much of the disagreement between Wood & Sherman and the Gurin Report centers around the role of direct and indirect effects. In Figure 1, the following are *direct effects*:

- 1 → 2
- 1 → 3
- 4 → 2
- 4 → 3
- 5 → 2
- 5 → 3
- 2 → 3

We depicted the direct effect of structural diversity (4) on student outcomes (3), the focus of Wood’s and Sherman’s critique, by a broken bold line in Figure 1. Wood & Sherman assert that “[s]ince the racial diversity of the student bodies in the CIRP database varies from institution to institution, it is possible to use the database to test the Powell rationale empirically using multivariate regression analysis. If the regression coefficient for racial diversity is statistically significant in a model that controls for all explanatory variables, one can say that racial diversity is associated with certain outcomes and is plausibly a cause of these outcomes” [path 4→3 in Figure 1]. (Wood & Sherman, at 58)

However, student background characteristics (1), structural diversity (4), and other institutional characteristics (5) may also have *indirect effects* on student outcomes (3) through the effect of diversity experiences (2) on student outcomes (3):

- 1 → 2 → 3
- 4 → 2 → 3
- 5 → 2 → 3

Thus, for example, structural diversity (4) may have a direct effect on student outcomes (3) (4 → 3), an indirect effect through diversity experiences (2) on outcomes (4 → 2 → 3), or both. We highlighted the paths in Figure 1 from structural diversity (percent minorities (4), to diversity experiences (2), to student outcomes (3) because this is the major hypothesis set forth in the Gurin Report. Structural diversity is hypothesized to act through campus experiences to affect student outcomes, a hypothesized indirect effect, *not* a direct effect.

Wood & Sherman assume that in Gurin’s report, diversity experiences (2) are merely a “proxy for racial diversity.” (Wood & Sherman, at 82-83). This is a fundamental misinterpretation of Gurin’s model because the Gurin Report does not claim that diversity experiences serve as a *proxy measure* for (that is, serve in place of), structural diversity. Rather, the Gurin Report’s theory is that campus diversity experiences transform structural diversity, an essential “resource variable,” into

educational outcomes. In other words, the structural diversity is a necessary condition for these educational benefits, but the causal effect of structural diversity on student outcomes works through the diversity experiences students have in an institution. This is entirely different than diversity experiences serving as a proxy measure for diversity.

As Figure 1 shows, in Gurin’s model, diversity experiences (2) are viewed as intervening causal variables; not as “proxy measures of racial or ethnic ‘diversity,’” as the NAS Brief claims. (NAS Brief, at 7). The fact that Wood & Sherman incorrectly regard diversity experiences as proxies may be the reason for their claim that “[a]s proxies for racial diversity, these variables are woefully inadequate.” (*Id.*)

3. Algebraic Demonstration

Through an algebraic depiction of the Gurin and Wood & Sherman models, one can pinpoint the fundamental difference between the arguments even though there is substantial overlap between the two models. This analysis shows that other than the methodological error in the Wood & Sherman model described above, the models are in fact highly similar.

In this algebraic demonstration:

- B will denote the student background characteristics (*e.g.*, SAT composite, self-reported high school grades),
- P, the percentage of undergraduates who were students of color, will denote “structural diversity,”
- G will denote the other general institutional characteristics (*e.g.*, selectivity, type, control),
- E will denote “diversity experiences” (or “socialization & discussion” in CIRP; *e.g.*, enrollment in an ethnic studies course, discussed racial issues), and
- Y will denote a “final student outcome.”

The substance of the arguments in the Gurin Report and the Wood & Sherman critique is exemplified by the following two excerpts.

1. **Gurin** asserts that “[a]n important question to examine first is whether [P] shapes classroom diversity and opportunities to interact with diverse peers...In order to examine [this question], I examined the relationships between [P] and [E]...[P] had significant positive effects on [E] among all students...In addition, the fact that these relationships are significant creates the possibility that [P] will also affect [Y] (not just [E]) in indirect ways (*e.g.*, through [E]). These indirect effects can only occur if the measures of [E] are significantly related to [Y], which is the major focus of the results in the next sections.” Gurin Report, at 115 Empirical Results, part 1.

2. By contrast, **Wood & Sherman**'s methodology is set out in their discussions of Astin, Chang and Gurin. In the section of their report discussing Chang and Astin that immediately precedes their analysis of the Gurin Report, they assert that:

- (a) "So far as the diversity rationale is concerned, the correlation between [E]...and [Y] are of interest only to the extent that there is a positive synergy or interaction (called an 'interaction effect' by statisticians) between [P, E and Y]." Wood & Sherman, at 62; and
- (b) "...as far as the Powell rationale is concerned,...what one really wants to know from the CIRP studies is how much of a difference, if any, [P] has on [Y], given that [P] has a statistically significant impact on [E]...[T]he really crucial data...is the way that [E] interact[s] with [P and Y]. An interaction effect is essentially a measure of the positive or negative synergies between (in this case) [P, E, and Y]...[T]he interaction effects are certain to be very weak, since correlations that would be used in the calculations are very small, and because one gets the interaction effects essentially by multiplying probabilities." *Id.* at 68-69.

A comparison between the quotes in (1) and (2) offers a distillation of differences between Gurin's and Wood's and Sherman's theoretical rationales and statistical methodologies. As noted above, Gurin examines the effect of diversity *in context* on educational outcomes. Wood & Sherman incorrectly assume that Justice Powell's opinion allows diversity to be a compelling interest only if it has a direct effect on educational outcomes regardless of the educational programs and opportunities for informal interactions between students that a university program may provide.

One can present the two rationales in simple algebraic form. For purposes of this analysis, let Y_{ic} , E_{ic} and B_{ic} be the levels of a final outcome variable, of the diversity experiences, and of the background characteristics, respectively, for the i 'th student at the c 'th college, and let P_c and G_c be the levels of structural diversity and other general institutional characteristics, respectively, at the c 'th college.

a. The Gurin Approach

Gurin tests the following model:

$$E_{ic} = a_0 + a_1P_c + a_2B_{ic} + e_{ic}. \quad (1a)$$

$$Y_{ic} = b_0 + b_1B_{ic} + b_2E_{ic} + b_3G_c + e'_{ic}. \quad (1b)$$

P = structural diversity

E = diversity experiences

Y = final student outcome

B = student background characteristics

G = other general institutional characteristics

As stated above, Gurin first shows that the effect of structural diversity P (percent minorities) on (diversity experiences) E is significantly positive, *i.e.*, that $a_1 > 0$ in Equation (1a). Next, Gurin shows that diversity experiences are correlated with outcomes, *i.e.*, that $b_2 > 0$ in Equation (1b). Then Gurin shows² that, when E (diversity experiences) is in the regression, P (percent minorities) has a negligible effect on Y (educational outcomes), *i.e.*, that $b_4 = 0$ in Equation (1c). However, it does not change the relationship between P (structural diversity) and E (educational diversity), which is positive in all cases. P is a necessary, but not a sufficient condition to produce positive academic outcomes.

The fact that the direct effect of structural diversity on outcomes drops out when diversity experience is added to the model, far from undermining Gurin's findings, strongly supports Gurin's contention that P has an indirect effect on Y. To a social scientist, Gurin's analysis shows that P (structural diversity) affects Y (educational outcomes) through E (diversity experience), as Gurin argues. This conclusion is carefully derived because equations 1a-c, shown on the preceding page, allow us to isolate the "causal" effect of structural diversity, through educational diversity, and is entirely justified by Gurin's analysis.

b. Similarities of Gurin and Wood & Sherman Approaches

Wood & Sherman appear to posit the following model:

$$Y_{ic} = b_0 + b_1B_{ic} + b_2E_{ic} + b_3G_c + b_4P_c + b_5P_cE_{ic} + e_{ic}. \quad (2)$$

P = structural diversity

E = diversity experiences

Y = final student outcome

B = student background characteristics

G = other general institutional characteristics

The "positive synergy or interaction" posited by Wood & Sherman is modeled by the term containing b_5 in Equation (2).³ This term is missing in Equation (1b, c) and this represents the formal difference between Wood's and Sherman's and Gurin's theoretical rationales.⁴

$$Y_{ic} = b_0 + b_1B_{ic} + b_2E_{ic} + b_3G_c + b_4P_c + e'_{ic}. \quad (1c)$$

² In her "step 3" regressions discussed in Wood & Sherman at 88.

³ The inclusion of the term with b_5 is the way in which most statisticians understand the concept of "interaction" within the methodology of multiple regression. It is possible that Wood & Sherman have some other formal model in mind, because, within this formal methodology, Wood's and Sherman's term, "synergies," has no place in the literature of social scientific methodology.

⁴ It is almost certain that the effects of P on E and Y are non-linear over the whole range of P values. However, it is likely that, for "small" values of P, *e.g.*, $0 < P < 0.2$, the range that includes most of the selective colleges and universities, these effects are approximately linear. We show linear effects of P in Equations (1) and (2) so as to keep the present comments simple.

Does this formal difference between Gurin's model (Equation (1c)) and Wood's and Sherman's model (Equation (2)) have any practical importance? The answer appears to be "No," because Wood & Sherman believe that "the interaction effects are certain to be weak." This means that Wood & Sherman would assume that $b_5 = 0$ in Equation (2), in which case Equation (2) reduces to become functionally identical to Equation (1c). In other words, Gurin and Wood & Sherman are using essentially the same formal model in thinking about the relationships among P, E and Y.

Another similarity between Gurin's and Wood's and Sherman's arguments is that Wood & Sherman appear to take it as "given that [P] has a statistically significant impact on [E]" (*Id.* at 68), which, as already noted, was shown by Gurin.

c. Fundamental Difference Between Gurin and Wood & Sherman Approaches

The fundamental difference between Gurin's and Wood & Sherman's approaches is that: (a) Gurin hypothesizes, and then demonstrates through the regressions, that P affects Y *indirectly*, and that the mechanism of this indirect effect involves E. By contrast, (b) Wood & Sherman argue that only the demonstration of a *direct* effect of P on Y would suffice (*i.e.*, be "the really crucial data") for an application of the Powell rationale. The latter argument flies in the face of a large body of social science research showing that institutional variables, such as P, have their effects on individual level variables, such as Y, through other mechanisms. This approach is firmly established in research regarding effects on higher education outcomes, in addition to the psychology sources cited by Gurin.⁵

Our primary interest is, and the primary interest of a social scientist should be, to understand what "causal" effects exist between these variables and how they operate. However, Wood & Sherman, rather than examining thoroughly whether there is a positive relationship between the variables studied, diverge from accepted social scientific method. Instead, they write that the Gurin approach is "absurd, and . . . a devious attempt to distract attention from . . . [the] finding that there are no educationally significant correlations between [P] and [Y]" (Wood & Sherman at 79), and that her approach "makes absolutely no sense as a way of addressing the question whether [P] might produce educational benefits" (*Id.* at 81).

Based upon our review of Gurin's and Wood & Sherman's analyses for this report, we believe that Gurin's approach is scientifically sound, and that these and other attacks in the Wood & Sherman Report are misplaced and unwarranted. Through careful statistical analysis, Gurin has shown that racial (structural) diversity operating through students' diversity experiences in college produces educational benefits to the students involved and to the larger society.

⁵ See, e.g., E. Pascarella & P. Terenzini, *How College Affects Students* at 12 (1991); E. Pascarella, College environmental influences on learning and cognitive development: A critical review and synthesis. In J. Smart (Ed.), *Higher Education: Handbook Of Theory And Research* (1985); A. Astin, *What Matters in College?* at 101-103 (1993).

4. The Wood & Sherman Report Incorrectly Criticizes Gurin’s “In Turn” (Indirect Effects) Hypothesis

This fundamental difference in approach leads Wood & Sherman to an incorrect interpretation of Gurin’s step 3 regression results. Wood & Sherman correctly summarize Gurin’s results, that “[a]fter all these variables [E] have been controlled, the regressions fail to find significant correlations between [P] and [Y]” (*Id.* at 82). But then Wood & Sherman incorrectly argue: “[t]his means that Gurin cannot argue that [P] produces educational benefits even when it is conjoined with these other factors. This finding completely devastates Gurin’s ‘in turn’ hypothesis.” (*Id.*)

This claim is mistaken because of the same error that pervades the Wood & Sherman analysis—their decision to ignore indirect effects. At this point in Wood & Sherman’s argument, the “in turn” hypothesis refers to the correlation between P and E (*Id.* at 82, 2nd ¶), which Gurin established empirically, not “logically or mathematically,” as Wood & Sherman claim. (Compare Gurin Report, Table 1 at 116 with Wood & Sherman at 82.) Therefore, the “finding” that controlling all variables there is not a relationship between [P] and [Y] cannot devastate the “in turn” hypothesis because this result is fully consistent with the indirect effect that Gurin demonstrates. This is why the Wood & Sherman criticism of Gurin’s step 3 regression results is invalid.

B. Wood & Sherman Incorrectly Assert That in Gurin’s Model Diversity Experiences Can Have A Beneficial Educational Effect Without Structural Diversity

Wood & Sherman raise another superficially serious criticism of Gurin’s conclusion that is not related to the fundamental difference outlined above. They contend on p. 85, “that when [E]...is a statistically significant predictor of [Y]...then, according to [Gurin’s] model, the size of the effect is independent of the number of minorities on campus. Thus the claimed beneficial effects of her campus experience variables would remain statistically significant even if the number of minorities on campus were to drop.”

This criticism is flawed because as we show below an increase in diversity experience [educational diversity → change to diversity experience] *within* an institution is related to an increase in educational outcomes. Moreover, a change in structural diversity *between* colleges *is* associated with a change in average student outcomes. The Wood & Sherman criticism conflates two kinds of “beneficial effects”—those occurring *within the same* college as opposed to those occurring *between different* colleges. Within the same college, the percentage of minority students (P) is necessarily constant, so the relationship of interest is the one between diversity experiences (E) and educational outcomes (Y). This inquiry examines the effect of E (diversity experiences) on Y (educational outcomes) within a given college, *i.e.*, the increase in Y (educational outcome) for a unit increase in E (diversity experiences) at a given school. From Equation (1c), we see that this within-college effect of E (diversity experiences) on Y

(outcome) is simply the magnitude of, b_2 . The increase in educational outcome (Y) depends on the level of E multiplied by the effect b_2 (holding other variables constant).

$$Y_{ic} = b_0 + b_1B_{ic} + b_2E_{ic} + b_3G_c + b_4P_c + e'_{ic}. \quad (1c)$$

The second effect of E is the *between-college* effect, *i.e.*, the increase in average educational outcome (Y) associated with a unit increase in P (percent minorities) in going from one college to another, all other variables remaining constant. The only way to assess the effect of varying P (percent minorities) in the CIRP database is to compare different colleges. For this reason, Wood & Sherman must be referring to this between-college effect on average Y when they conclude that the benefits would remain significant when P drops. This argument can be analyzed formally by means of Equations (1a) and (1b).

In equations (1a) and (1b), diversity experience (E), student background (B) and outcome (Y) are measured for individual students at each college:

$$E_{ic} = a_0 + a_1P_c + a_2B_{ic} + e_{ic}. \quad (1a)$$

$$Y_{ic} = b_0 + b_1B_{ic} + b_2E_{ic} + b_3G_c + e'_{ic}. \quad (1b)$$

To analyze the relationships *between colleges*, we use the average E (E “bar” in equation 1a’), the average B (“B bar”), and the average Y (“Y bar”) for each college.⁶ We can now re-write the original equations in terms of averages, focusing here only on the key variables in Wood & Sherman’s argument: structural diversity (P), average diversity experience (“E-bar”), and average outcome (“Y-bar”) between colleges:

$$\bar{E}_c = a_0 + a_1P_c, \quad (1a')$$

(Expressed verbally, average college diversity experience is a function of structural diversity), and

$$\bar{Y}_c = b_0 + b_2 \bar{E}_c. \quad (1b')$$

(Expressed verbally, average student outcome is a function of average diversity experience.)

It follows from (1a’) and (1b’) that:

$$\bar{Y}_c = b'_0 + a_1b_2P_c. \quad (3)$$

(Expressed verbally, average student outcome is a function of the percentage minority (P).)

⁶ We use average E and average B because the average of a campus-level variable based on all students in the sample at that campus is the value of the variable itself.

Equation (3) shows clearly that if P (percent minority students) drops, so does average Y (educational outcomes), *i.e.*, the benefits of structural diversity decline if structural diversity drops. This is why the Wood & Sherman conclusion that the beneficial effects of diversity would remain significant if P were to drop, is incorrect.

C. Other Related Wood & Sherman Criticisms of Gurin Methodology

The Wood & Sherman Report and the NAS Brief raise a number of other methodological criticisms of the Gurin study that fail to undermine its validity. In this section, we briefly discuss several of these other criticisms.

1. Gurin's Use of Self-Reported Data on Grades

Wood & Sherman criticize Gurin's statistical significance findings to the extent that they rely on self-reported data (Wood & Sherman, at 81 & n.66; NAS Brief, at 10–12).

This criticism is at odds with standard social scientific practice. Academic surveys are expensive to conduct and burdensome on respondents. Consequently, self-reports of academic achievement are often depended upon, *when there is evidence of the validity of their interpretation* (e.g., P. Ewell *et al.*, 1993). Gurin, in her "Response to the Critique by the National Association of Scholars" posted on the University of Michigan website, provides an accurate and relevant summary of what is known about the validity of self-reports of academic performance. Reliance on such data in the circumstances Gurin describes is standard practice among social scientists.

2. Gurin Reports Limited Correlations

The NAS Brief argues that Gurin relies upon Astin's work, but includes in her study only one of the variables of academic achievement "accepted" by Astin and finds "no statistically significant, consistent results" for this variable (grades). It asserts that she looks for 24 possible correlations between her proxies for diversity and grades, and finds a statistically significant correlation with regard to only six, and these correlations indicate that her factors have almost no effect on grades (NAS Brief, at 11 & n. 5).

We disagree with the assertion that diversity experiences (E) are a proxy for structural diversity (P). As shown in our report, they clearly are not. Second, Wood & Sherman have chosen one particular outcome variable, grades, to make their argument regarding "no statistically significant, consistent results." However, this ignores other categories of educational outcomes, such as engagement and motivation and democratic outcomes that are very important measures of educational outcome, as President Bowen's letter cited in *Bakke* indicates.

3. The Astin Study

The NAS Brief argues that Astin, whom Gurin cites for other reasons, studied one of the same data sets, and reached the opposite conclusions with regard to the effect of diversity on academic outcomes. (*Id.* at 13).

Astin and Gurin agree that there were very weak, *direct* effects of structural diversity on student outcomes. However, this in no way rebuts Gurin's point that the causal effect of structural diversity on student outcomes operates *indirectly* through the causal mechanism of campus diversity experiences. In fact, Astin agrees that structural diversity effects operate indirectly through experience. As we have shown above, her methodology is appropriate, and is no less so because, as expected, the direct effect drops out in the presence of diversity experiences, the mechanism through which structural diversity mainly works to produce enhanced outcomes.

V. CONCLUSION

For the foregoing reasons, it is our opinion that the Gurin Report provides methodologically valid evidence, consistent with standard social scientific method in our fields, of a relationship between affirmative action programs described in her data set and improved educational outcomes.

By contrast, the Wood & Sherman critique of the Gurin Report suffers from serious methodological short-comings. By refusing to consider highly relevant indirect relationships and raising other incorrect methodological objections to Gurin's evidence, Wood & Sherman reach the incorrect conclusion that the Gurin Report fails to demonstrate that increased racial diversity in education, combined with appropriate academic programs, improves student learning.

The issues raised by these contrasting approaches are methodologically important and intellectually stimulating. We are grateful for this opportunity to study them.