

## Judgments of Salary Bias and Test Bias from Statistical Evidence

MICHAEL H. BIRNBAUM AND LINDA G. HYNAN

*University of Illinois, Urbana-Champaign*

Academics examined correlation scatter plots and made judgments of group bias. Half of the subjects judged bias in salaries from graphs that showed salaries plotted against merit with a separate correlation ellipse for each of two groups. The others judged test bias from graphs that showed job performance vs test scores. Relative positions of the centroids for the two groups and the within-group correlation were systematically varied. Judgments of bias were not consistent with either "forward" or "reverse" regression definitions of bias. Judged bias is not a monotonic function of the group difference in salary between persons of equal merit nor is it a monotonic function of the group difference in merit between persons of equal salary. Instead, judgments in both tasks were consistent with a difference of differences model; for example, when the group difference in standard deviation units is equal on both salary and merit, judges say the situation is unbiased. The results are consistent with a one-mediator model of group equity, which assumes that both measures are imperfectly and equally correlated with the mediator. © 1986 Academic Press, Inc.

Statistical definitions have been proposed to measure group-related bias in salaries and in test scores. In the test bias literature, a number of different definitions of test bias have been proposed (e.g., Darlington, 1971; Petersen & Novick, 1976). Possibly the most-often-applied definition of test bias is the regression definition, which states that there is no bias if the slopes and intercepts of the regression lines predicting the criterion (e.g., job performance) from the test are the same for all groups. For example, if blacks can perform the job better on the average than whites with the same test score, the regression definition would state that the test is biased against blacks because it underpredicts their job performance.

In the salary bias literature, some investigators have proposed that sex bias in the salaries should be defined as the difference in salary between men and women with the same measured qualifications (years of experience, productivity, merit, etc.). According to this definition, sometimes called "forward regression," salaries are biased against women if women

We thank Stuart Nagel, Fritz Drasgow, Lloyd Humphreys, Kevin Trenberth, and Barbara A. Mellers for helpful comments on an earlier draft. Requests for reprints may be sent to Michael H. Birnbaum, Department of Psychology, California State University, Fullerton, CA 92634.

are paid less on the average than men with the same qualifications. Another definition, sometimes called "reverse regression," states that if women are more qualified on the average than men who are paid the same salary, then the situation is biased against women. However, it has been shown that these two definitions do not in general yield the same conclusions. In particular, it is often the paradoxical case that women are paid less than men with the same qualifications and in the same population, women are *less* qualified on the average than men with the same salaries (Birnbau, 1979a, 1979b, 1981, 1983).

Birnbau's (1979a, 1981) one-mediator model of equity explains how these paradoxical group differences occur in the absence of bias. The mediated path model includes the regression definitions as limiting special cases. The four rival models considered by Darlington (1971) for the test bias literature can also be interpreted as special cases of the general mediated model (see Appendix).

Because different mathematical definitions of test bias and salary bias do not agree, and do not follow from axiomatic systems, it has been recognized that the issue of bias cannot be resolved in a purely statistical fashion (Petersen & Novick, 1976). Choice among definitions of bias is thus partly an issue in the psychology of judgment. The purpose of the present research is to obtain judgments of group bias or equity in order to compare human judgments of the degree and direction of bias with different models of bias.

### *Models of Bias*

Suppose two groups, A and B, have been measured on variables  $X$  and  $Y$ , and the correlation between  $X$  and  $Y$ , is the same within each group. For example,  $Y$  and  $X$  might be salary and measured qualifications, scaled in standard units. In the present experiments, the group centroids and the within-group correlation are manipulated, and the within-group variances are equal. Let  $Y_A$ ,  $Y_B$ ,  $X_A$ , and  $X_B$  represent the group means on  $Y$  and  $X$ , respectively. Let  $r$  represent the correlation/regression coefficient, which is the same for each group ( $0 < r < 1$ ).

The forward regression model can be written as follows:

$$b = (\bar{Y}_A - \bar{Y}_B) - r(\bar{X}_A - \bar{X}_B) \quad (1)$$

where  $b$  is the measure of bias against group B. This model implies that the difference in  $Y$  and the difference in  $X$  should contribute additively to the judged bias, but the correlation should multiply the effect of the difference in  $X$ . For example, forward regression implies that bias will be zero when the mean difference in salary equals the difference predicted on the basis of merit.

The reverse regression model can be written

$$b = r(\bar{Y}_A - \bar{Y}_B) - (\bar{X}_A - \bar{X}_B) \quad (2)$$

which implies additive effects of the differences on the two dimensions, but an interaction between the correlation and the difference in  $Y$ . For example, this model implies that bias is zero when the average difference in merit equals the difference predicted from salaries.

The difference of differences model can be written

$$b = (\bar{Y}_A - \bar{Y}_B) - (\bar{X}_A - \bar{X}_B) \quad (3)$$

This model implies there will be no effect of the within-group correlation on the bias judgments. For the test bias situation it is necessary to reverse the sign of  $b$  in all of the equations.<sup>1</sup>

The three equations, as written, can be considered as statistical definitions that derive as special cases from a more general conception (Appendix). However, the three equations can also be interpreted as psychological models. Psychological versions of the models would replace  $\bar{X}_A - \bar{X}_B$ ,  $\bar{Y}_A - \bar{Y}_B$ , and  $r$  with subjective scale values, to allow for the possibility that the subjective correlation, for example, is a function of the objective correlation but may not equal the objective correlation. Similarly, subjective differences may be a nonlinear function of objective differences.

In extreme cases, all three models will agree on the sign of the bias (for example, Fig. 1). In general, however, the three models make different predictions (Birnbaum, 1985). The three models differ with respect to the effect of the correlation between  $X$  and  $Y$ . Equation (3) states that the correlation has no effect. In Eq. (1), correlation multiplies the difference in  $X$ , whereas in Eq. (2),  $r$  multiplies the difference in  $Y$ . A further distinction can be seen from the equations when  $\bar{X}_A - \bar{X}_B = \bar{Y}_A - \bar{Y}_B > 0$ . In this case, Eq. (3) implies  $b = 0$  whereas Eq. (1) implies  $b > 0$  and Eq. (2) implies  $b < 0$ . Furthermore, Eqs. (1) and (2) imply that the absolute bias is greater when both (equal) differences are larger and when the within-group correlation is smaller.

<sup>1</sup> It is interesting that the terms "forward" and "reverse" regression are used in opposite meanings in the salary and test bias literatures. In the test bias literature, "forward" regression means the variable assumed to be unbiased (job performance) is predicted from the variable suspected of bias (test performance), whereas in the salary bias literature these psychometric considerations are reversed. To avoid confusion, the present experiment uses the term "forward" for the prediction of  $Y$  from  $X$ , without any statement concerning the locus of bias. In the salary bias case  $Y$  is salary; in the test bias case,  $Y$  is job performance. For consistency, the sign of  $b$  must be changed between the two situations.

## METHOD

One of two questionnaires was sent to 208 members of the academic community (mostly faculty) of the University of Illinois who appeared on a mailing list for notices of statistical lectures and seminars. The list included faculty in mathematics, psychology, economics, political science, sociology, educational psychology, physical sciences, etc.

Half of the questionnaires dealt with the issue of salary bias and half of the questionnaires dealt with the issue of test bias. At least 42 completed questionnaires of each type were returned, a return rate of about 44% (76 faculty, 13 graduate students, 1 anonymous). A good number of the respondents were highly trained, as evidenced by their publications in mathematical statistics, multivariate statistical methods, equity, or fairness of test usage.

*Instructions*

The instructions for the salary bias questionnaire stated that the ellipses represented scatter plots showing salary vs merit (including experience and productivity) with a separate ellipse for each group, as in Fig. 1. The points inside each ellipse were said to represent centroids for groups A and B. The solid lines show regression lines predicting salary from merit for each group; the dashed lines show regression lines predicting merit from salary for each group. Instructions specified that the units were equal for all figures and were in equal standard units for group A for abscissa and ordinate in all panels.

The instructions for the salary bias condition described the situation in the left of Fig. 1 as biased against group B (note that both groups have

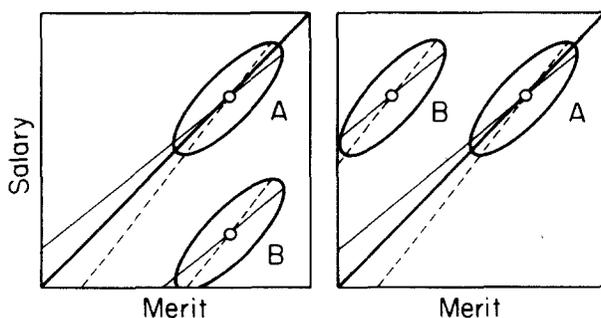


FIG. 1. Scatter plots showing salary vs merit for groups A and B. Bivariate ellipses and centroids are shown for each group. Solid lines show regression lines predicting salary from merit for each group. Dashed lines show regression lines predicting merit from salary for each group. Diagonal lines shows identity line for standard scores for group A. All three models describe the situation on the left as biased against B, and they describe situation on the right as biased against A.

equal average merit but group A has higher salaries). The right of Fig. 1 was described as biased against group A (note that both groups have equal average salaries but group A has higher merit). The instructions for both tasks were neutral with respect to the models, as they mentioned centroids and all four regression lines; the extreme cases in Fig. 1 would be biased in the same direction under all three definitions.

For the test bias questionnaire, the ordinate was labeled job performance and the abscissa was labeled test score. The solid lines show prediction of job performance from test score and the dashed lines show prediction of test score from job performance. Subjects were asked to judge how fair it would be to hire on the basis of test score, without taking group membership into account. The situation on the left was described as biased against group A (both groups have the same average test score whereas group A performs the job better on the average). Similarly, the situation on the right was described as biased against group B (both groups perform the job equally well but group B has lower test scores; hence, members of group B would not be hired even though they can perform the job equally well).

In both forms of the questionnaire, participants were asked to examine scatter plots as in Fig. 1 and make judgments on a scale that ranged from  $-9$  to  $9$ , with category ratings varying from  $-8 = \textit{very very unfair to group B}$ , to  $8 = \textit{very very unfair to group A}$ , with  $0 = \textit{equal for both groups}$ . Even integers between  $-8$  and  $8$  were labeled with intermediate categories.<sup>2</sup>

### *Experimental Design*

Figure 2 illustrates the experimental design that was used for both salary and test situations. The ellipse for group A was always in the location shown, and there were 13 different possible locations for the centroid of group B, indicated by the open circles in Fig. 2. Within each figure there were two ellipses, both representing the same correlation. There were two values of the within-group correlation, high =  $.8$  and low =  $.48$ . The ellipses and the regression lines were drawn as in Fig. 2.<sup>3</sup> Each figure was  $2.5 \times 2.5$  in. A faint  $10 \times 10$  grid of horizontal and vertical lines was also visible.

<sup>2</sup> Note that  $b$  in Eqs. (1), (2), and (3) represents bias against group B, whereas the dependent variable uses positive scores for bias against A. Thus, the judgments can be interpreted as  $-b$  in the case of salary bias and  $b$  in the case of test bias.

<sup>3</sup> The relationship between the ratio of major to minor axis of the ellipse and the slope of the regression line is as follows:  $m/n = \sqrt{(1 + \rho)/(1 - \rho)}$ , where  $m/n$  is the ratio; and  $\rho$  is the correlation coefficient (slope of regression line on standard score plot).

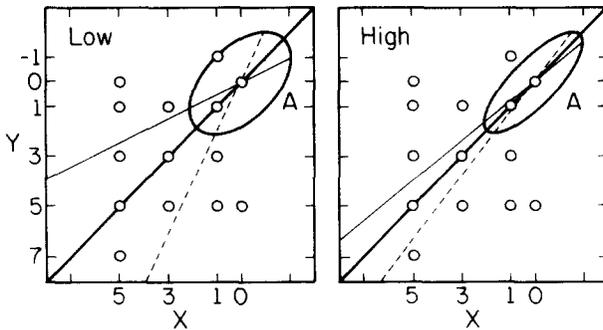


FIG. 2. Experimental design. Open circles show loci of centroids for group B. Group A was fixed in location shown. Ellipse for low correlation (left) and ellipse for high correlation (right) are shown.  $Y$  and  $X$  were labeled salary and merit or job performance and test score, respectively. Values show differences in  $X$  and  $Y$  between group A and B for reference with Fig. 3.

### Procedure

All of the booklets contained the two extreme cases of Fig. 1 with instructions, followed by five warm-up trials, followed by the 26 experimental cases (2 correlations  $\times$  13 locations). The trial order was random, with the restriction that the same pair of centroids would not be repeated on successive trials.

The warm-up contained all regression lines and the identity line as in Fig. 1 for all subjects. For the experimental trials, half of the subjects in each situation received two ellipses with centroids, both regression lines for group A, and the identity line. The other half received ellipses with centroids and no regression lines or identity lines. This variable (lines vs no lines) had virtually no effect and is ignored in subsequent analyses.

## RESULTS

Figure 3 plots the mean judgments of group bias as a function of the difference in  $X$  (merit or test score) between group A and B with a separate solid curve for each value of the difference in  $Y$  (salary or job performance). The left panel shows the results for the low correlation and the right panel shows results for the high correlation. The upper panels show results for the 48 subjects who judged salary bias and the lower panels show the results averaged over the 42 subjects who judged test bias. The dashed lines connect points where the difference between the salary difference and merit differences are equal (i.e., the dashed lines in Fig. 3 connect judgments of points in Fig. 2 that are equidistant from the identity line).

The lower panels of Fig. 3 show that the test bias data resemble the salary data closely when the judgments are reflected about zero. Thus,

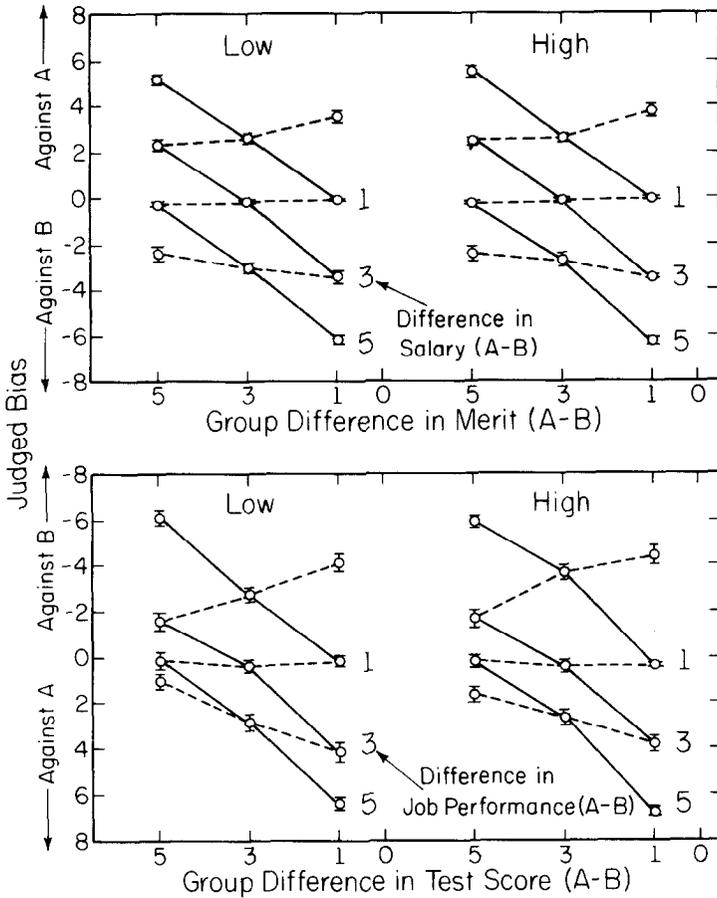


FIG. 3. Upper panels show mean judgments of salary bias, as a function of difference in merit of (A-B) with a separate solid curve for each level of salary difference. Positive scores indicate judgment of bias against group A. Dashed curves connect judgments of cases for centroid of group B equidistant from the identity line in Fig. 2. Judgments for low and high correlations are on the right and left respectively. Lower panels show mean judgments of test bias, plotted as above except ordinate scale has been reflected. Brackets show  $\pm 1$  standard error for each mean.

for a given graph, the judgment for the test bias situation can be well predicted from the salary bias judgment multiplied by  $-1$ . Note that the ordinate for the test bias data has been reflected to facilitate comparison of the two sets of data.

Figure 3 shows that the data are inconsistent with either forward or reverse regression definitions of bias. Forward regression (Eq. (1)) implies that for the salary condition (a) the center dashed lines in Fig. 3 should be below zero (bias against B); (b) the center dashed lines should both have positive slopes; (c) bias against B should be greater when the

correlation is smaller. Reverse regression (Eq. (2)) implies the mirror image; the center dashed line should be above zero (bias against A), the slopes should be negative, and the bias against A should be greater when the correlation is lower.

However, the center dashed lines in Fig. 3 (where the salary difference equals the merit difference) are near zero and nearly horizontal in all panels. Figure 3 also shows that there is virtually no difference between the high- and low-correlation conditions. Both regression definitions imply that for a given pair of centroids on the identity line, the amount of bias should be greater when the within-group correlation is lower. Figure 3 shows that when a pair of centroids falls on the identity line, the situation is judged to be unbiased, irrespective of the separation of the centroids or the correlation, contrary to either regression definition. However, this pattern is consistent with the difference of differences model (Eq. (3)).

Individual data were highly consistent with the averages shown in Fig. 3. For each subject, there are six trials with two centroids on the identity diagonal, forming a  $2 \times 3$  design. The mean, effect of correlation, and effect of  $X$  form three independent contrasts, which should obey different patterns according to the three models. Of the 48 respondents in the salary condition, all but 2 had mean judgments between  $-1$  and  $1$  (close to unbiased), including 38 judges who responded exactly "0" on all six trials, consistent with Eq. (3). The two others had negative means and their effects of correlation and  $X$  were in the direction predicted by forward regression. No other subjects had all three contrasts in the directions predicted by either regression model.

For the test bias judgments all but 10 of the 42 participants had means in the interval from  $-1$  to  $+1$ , including 23 who responded exactly "0" on all six trials, consistent with Eq. (3). Of the 10 outside this interval, 4 were negative and 6 positive, including 2 cases that had all three effects in accord with forward regression and two cases in accord with reverse regression. No other subjects had all three patterns consistent with Eq. (1) or (2). In sum, the vast majority of the individuals were in agreement with the pattern of the means in Fig. 3.

Analysis of variance was performed using the first 20 subjects for each condition of task and lines or no lines. All effects of task (after reflection) and lines vs no lines were negligible. Correlation had no significant main effects ( $F(1,76) < 1$ ) or interactions with  $X$  or  $Y$  ( $F(2,152) = 2.45$  and  $1.87$ , respectively), nor was correlation involved in any other significant interactions. The main effects of  $X$  and  $Y$  were statistically significant ( $F(2,152) = 965.6$ , and  $752.6$ , respectively), but their interaction was not,  $F(4,304) = 2.46$ , in agreement with the apparent parallelism of the solid curves in each panel.

The divergence of the dashed lines from left to right in each panel of

Fig. 3 shows that for a given vertical deviation from the identity line in Fig. 2, the judged inequity becomes more extreme as the centroid of B approaches A. The interaction between  $X$  and the deviation between  $X$  and  $Y$  (nonparallelism of dashed curves in Fig. 3) was statistically significant,  $F(4,304) = 45.6$ . Individual data were examined and the group averages are highly representative of individual data patterns.

The lack of main effects and interactions involving correlation is consistent with Eq. (3), but would not be expected from either regression definition. Equation (3) is also consistent with judgments of cases where both centroids fall on the identity line in Fig. 2. The parallelism of the solid curves in Fig. 3 is consistent with the assumption that the judgments of bias are a linear function of the difference between subjective differences in  $X$  and  $Y$ , as in Eq. (3). The divergence of the dashed lines in Fig. 3 indicates that subjective differences are negatively accelerated functions of the objective distances in the stimuli for both  $\bar{X}_A - \bar{X}_B$  and  $\bar{Y}_A - \bar{Y}_B$ . In summary, the data for both tasks are compatible with the difference of differences model of Eq. (3).

## DISCUSSION

The present results are not consistent with either "forward" or "reverse" regression definitions of bias. Instead, the situation is judged to be "fair" when the difference between the means in  $Y$  relative to the within-group standard deviation of  $Y$  equals the difference between the means of  $X$  relative to the within-group standard deviation of  $X$ . This result would be consistent with the difference of standard differences model, which follows from the one-mediator model when the correlation between observed salary and true merit equals the correlation between true merit and measured qualifications. For the test bias case, it would mean that both the test and job performance are regarded as equally correlated with true performance (Appendix).

The difference of differences model of group inequity is consistent with Mellers' (1982) findings for individual inequities. Mellers (1982) asked subjects to judge the bias of the salaries of individuals within a larger group. She found that the judgments could be represented by a difference of differences model involving three relational transformations. People appear to compare each salary with the distribution of salaries, compare each merit with the distribution of merits, compute the difference between these two relative values, and finally judge this difference according to its relation to the distribution of salary-merit differences. Although the present results do not provide a test of the relativity (distributional) part of Mellers' (1982) theory, the present finding of a difference of differences model fits nicely with her results, consistent with the hypothesis that people assess group bias and individual bias by the same process. In the present study, the units of  $X$  and  $Y$  were standardized,

with group A in a fixed location having a fixed type of distribution. Therefore, the present study does not distinguish among different theories of distributional relativity (Mellers, 1982; Mellers & Birnbaum, 1982).

When statistical definitions and human judgments do not agree, at least one of the two can be questioned. On some occasions, when statistical theorems arise from a mathematical theory that optimizes some goal, one would conclude that statistical computations should replace human judgments, and disagreement would be taken as evidence of human failure to process information correctly. On the other hand, if someone proposed a definition of "chair" that names desks, books, and lamps as "chairs" and fails to identify a reclining chair as a "chair," that definition might cause confusion because it departs from human usage. In such a case the definition would be questioned, rather than the ability of humans to identify chairs.

In the present case, it seems reasonable to question the regression definitions of bias because they do not represent the concept of bias as used by people who examine scatter plots. The difference of standard differences model of bias remains consistent with the judgments and has a mathematically consistent foundation, as it derives from the one-mediator model (Appendix). Therefore, the general mediated model of group equity appears to provide a more useful null hypothesis for the study of group equity than the special cases of forward or reverse regression.

It would be interesting to investigate how judgments of bias would be formed when the information is presented in different formats. For example, subjects might be asked to give judgments on the basis of verbal descriptions. In pilot work we noticed that graduate students and faculty are puzzled by verbal descriptions such as the following: "For this company it has been found that men receive \$2000 more salary on the average than women who have the same seniority and simultaneously women have 5 years less seniority on the average than men who receive the same salaries." People often asked how this can occur. By drawing two ellipses with centroids on the identity line and showing the two regression lines for each ellipse, the paradox could be quickly explained. Because such verbal phrases do not convey information about the positions of the centroids and the within-group variances and covariances, and because they often sound contradictory, the present study used graphical display.

When an investigator reports only a single regression equation predicting salary from merit in a study of real salaries, readers might reach the conclusion that salaries are biased against B if B has lower salaries than predicted by A's equation. However, the present data show that people may reach the opposite conclusion when all of the data are presented graphically, depending on the relative locations of the centroids and other information not available in a regression equation.

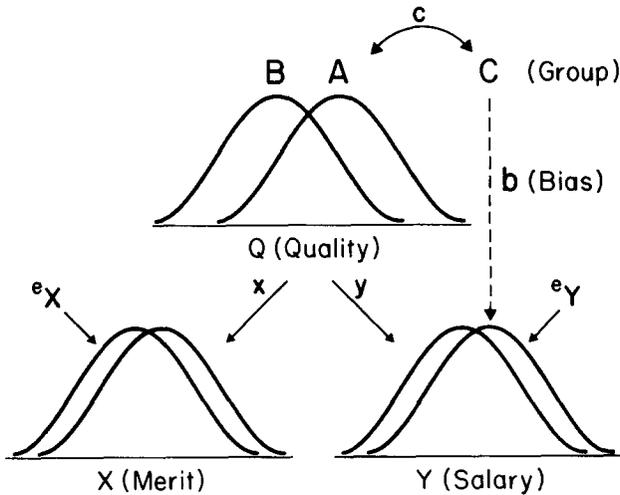


FIG. 4. Mediated model of bias. Groups A and B may differ on underlying quality ( $Q$ ). The correlation between groups ( $C$ ) and quality is denoted  $c$ . Variables  $X$  and  $Y$  (illustrated for merit and salary) are partially determined by Quality, and  $Y$  (salary) may also depend on bias ( $b$ ). The equations in standard score notation are:  $Y = yQ + bC + e_Y$  and  $X = xQ + e_X$ , where  $e_X$  and  $e_Y$  are mutually uncorrelated errors. For the test bias situation, the bias path would be from group membership to  $X$  (test score).

Because there is a loss of information when only a single regression equation is presented, it seems reasonable to request that data for real studies of salary and test bias should be presented graphically, as in Figs. 1, 2, and 4. From such graphs, one can construct regression equations, see the means, and examine the within-group correlation; however, given only a single pair of regression equations, one cannot reconstruct the graph.

When data are presented graphically, it should be clear whether or not the conclusion for a given set of data would depend on the choice of a definition of bias. In this way, computational issues can be separated from judgmental issues.

The present results suggest that statistically oriented academics do not make judgments consistent with either forward or reverse regression definitions of bias. Instead, their judgments are consistent with a special case of the mediated model of group equity in which both measured variables are assumed to be equally correlated with the mediator.

## APPENDIX

Birnbaum's (1979a, 1981) mediated path model, shown in Fig. 4, provides a convenient framework for the comparison of alternative definitions of bias. Let  $C$ ,  $X$ , and  $Y$  refer to group membership (dichotomous), and the measured variables, respectively. The latent variable,  $Q$ , is true quality which mediates the correlations among  $C$ ,  $X$ , and  $Y$  in the absence of bias. Bias, denoted  $b$ , is a direct path from group membership to  $X$  or  $Y$ . In the case

of salary equity, for example,  $X$  would be measured merit and qualifications,  $Y$  would be salary, and  $C$  might be 1 for male and 0 for female.

The errors in  $X$  and  $Y$  are denoted  $e_X$  and  $e_Y$  and are assumed to be uncorrelated with  $Q$ ,  $C$ , and each other. Thus, if bias is zero then people with the same true quality ( $Q$ ) will have the same average salaries, independent of group membership. The correlations between  $Q$  and  $X$  and  $Y$  are assumed to be positive but may be less than perfect. Thus, measured merit may be an imperfect assessment of true quality:  $X = xQ + e_X$ ; and salary may also vary with "luck" factors that are not due to true quality of work, and possibly bias:  $Y = yQ + bC + e_Y$ .

According to Birnbaum's general model, the correlations among the variables can be represented as follows:

$$\rho_{XY} = xy + xcb \quad (4a)$$

$$\rho_{XC} = xc \quad (4b)$$

$$\rho_{YC} = yc + b \quad (4c)$$

where  $\rho_{XY}$ ,  $\rho_{XC}$ , and  $\rho_{YC}$  are the three observed correlations;  $x$ ,  $y$ , and  $c$  are the correlations of  $X$ ,  $Y$ , or  $C$  with  $Q$ , respectively;  $b$  is the coefficient of bias. If bias,  $b$ , is zero, then the general model implies the following:

$$\rho_{XY} \leq \rho_{YC}/\rho_{XC} \leq 1/\rho_{XY} \quad (5)$$

There are four special cases of this model that are regarded as rivals (see, e.g., Darlington, 1971). These four models follow from assumptions concerning  $c$ ,  $x$ , and  $y$ .

1. If there are no group differences in  $Q$  ( $c = 0$ ), there should be no correlation between group membership and either  $X$  or  $Y$  (for  $b = 0$ ). If  $\rho_{XC} = 0$  then  $\rho_{YC} = b$ . This model can be rejected, for example, if men and women differ in average merit.

2. If  $X$  is assumed to be perfectly correlated with  $Q$  (i.e.,  $x = 1$ ) and if  $b = 0$ , it follows (from Eq. (4a), (4b), and (4c)) that the average value of  $Y$  should be the same for each group with  $X$  held constant. For example, men and women should receive the same average pay for equal measured qualifications. This implication corresponds to forward regression.

3. If  $Y$  is assumed to be perfectly correlated with  $Q$  ( $y = 1$ ) and if  $b = 0$ , it follows that there should be no difference in  $X$  with  $Y$  fixed, which is consistent with the "reverse" regression definition of equity. For example, men and women with the same salaries should be equally qualified on the average.

4. If neither  $X$  nor  $Y$  is assumed to be perfectly correlated with  $Q$  then the two group differences will go in (seemingly) opposite directions, when  $b = 0$ . For example, men will have higher average qualifications than women of the same salary, and simultaneously men will have higher average salary than women with the same qualifications. This paradoxical case has been discussed in detail by Birnbaum (1979a, 1979b, 1981, 1985).

If it is assumed that  $X$  and  $Y$  are equally correlated with  $Q$  (i.e., that  $x = y$ ), Eqs. (4b) and (4c) imply

$$b = \rho_{CY} - \rho_{CX} \quad (6)$$

These assumptions yield an expression for  $b$  that is independent of  $\rho_{XY}$ . From the definition of the correlation coefficient, this model can be rewritten

$$b = \sqrt{p(1-p)} [(\bar{Y}_A - \bar{Y}_B)/\sigma_Y - (\bar{X}_A - \bar{X}_B)/\sigma_X] \quad (7)$$

where  $p$  is the proportion of males;  $\bar{Y}_A$ ,  $\bar{Y}_B$ ,  $\bar{X}_A$ , and  $\bar{X}_B$  are the mean values of  $X$  and  $Y$  for the two groups; and  $\sigma_X$  and  $\sigma_Y$  are the standard deviations of  $X$  and  $Y$ , respectively. When the proportion and standard deviations are fixed, Eq. (7) is a difference of differences

model, like Eq. (3). Equation 7 implies that if  $b = 0$ , the sex difference in salary in standard deviation units should equal the sex difference in merit in standard units. The judgments of bias, in Fig. 3, are consistent with this prediction. However, the present experiments did not manipulate the proportions or the distributions; therefore, the present data do not permit a test among alternative versions of the difference of differences model, such as Mellers (1982) and Gollob (1984).

## REFERENCES

- Birnbaum, M. H. (1979a). Procedures for the detection and correction of salary inequities. In T. R. Pezzulo & B. E. Brittingham (Eds.), *Salary equity*. Lexington, MA: Lexington Books.
- Birnbaum, M. H. (1979b). Is there sex bias in salaries of psychologists? *American Psychologist*, *34*, 719–720.
- Birnbaum, M. H. (1981). Reply to McLaughlin: Proper path models for theoretical partialling. *American Psychologist*, *36*, 1193–1195.
- Birnbaum, M. H. (1982). On the one mediator model of salary equity. *American Psychologist*, *37*, 1146–1147.
- Birnbaum, M. H. (1983). Perceived equity of salary policies. *Journal of Applied Psychology*, *68*, 49–59.
- Birnbaum, M. H. (1985). Relationships among models of salary bias. *American Psychologist*, *40*, 862–866.
- Darlington, R. B. (1971). Another look at “cultural fairness.” *Journal of Educational Measurement*, *8*, 71–82.
- Gollob, H. F. (1984). Employee worth and indices of sex discrimination in salaries. *Proceedings of the Social Statistics Section of the American Statistical Association*.
- Mellers, B. A. (1982). Equity judgment: A revision of Aristotelian views. *Journal of Experimental Psychology: General*, *111*, 242–270.
- Mellers, B. A., & Birnbaum, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 582–601.
- Petersen, N. S., & Novick, M. R. (1976). An evaluation of some models for culture-fair selection. *Journal of Educational Measurement*, *13*, 3–29.

RECEIVED: November 12, 1984