JOHN F. DAUM (SB #52313)
FRAMROZE M. VIRJEE (SB #120401)
DAVID L. HERRON (SB #158881)
PETER L. CHOATE (SB #204443)
O'MELVENY & MYERS LLP
400 South Hope Street
Los Angeles, California 90071-2899
Telephone: (213) 430-6000
Facsimile: (213) 430-6407

Attorneys for Defendant
State of California

SUPERIOR COURT OF THE STATE OF CALIFORNIA

CITY AND COUNTY OF SAN FRANCISCO

ELIEZER WILLIAMS, et al., Plaintiffs, vs. STATE OF CALIFORNIA, DELAINE EASTIN, State Superintendent Of Public Instruction, STATE DEPARTMENT OF EDUCATION, STATE BOARD OF EDUCATION, Defendants.	Case No. 312 236
	Hearing Date: Sept. 17, 2003
	Time: 3:30 P.M.
	Department: 20
	Judge: Hon. Peter J. Busch
AND RELATED CROSS-ACTION.	

DECLARATION OF RICHARD A. BERK

DECLARATION OF RICHARD A. BERK

I, Richard A. Berk, say:

1. I am Professor of Statistics and Sociology at the University of California, Los Angeles, and Director of the UCLA Statistical Consulting Center. I have done teaching and research in statistics and sociology since 1970, with a special focus on methodological issues -- that is, what research techniques produce valid and reliable results and what techniques do not. I was given the Paul S. Lazarsfeld Award by the American Sociological Association for methodological contributions to the literature. My curriculum vitae is attached as Exhibit A.

2. I have been retained by defendants in this action and asked to analyze the scientific and statistical reliability and validity of a survey conducted by the Louis Harris Organization entitled "A Survey of the Status of Equality in Public Education in California -- A Survey of a Cross-Section of Public School Teachers." It is my understanding that this survey was carried out at plaintiffs' behest for the purpose of supporting their claims in this action. A copy of the expert report I prepared with respect to the survey is Exhibit B. The report accurately sets forth my opinions concerning the validity and reliability of the survey.

3. As more fully set forth in my report, it is my professional opinion that the Harris survey was not conducted in

accordance with sound statistical or sociological methods, and that no reliance can be placed on the conclusions drawn in the study. In particular, I am advised that plaintiffs' Memorandum in Support of Motion for Summary Adjudication of State's Duty to Ensure Equal Access to Instructional Materials cites the Harris survey for the proposition that 11.7% of California teachers do not have enough textbooks for use in class and 32% do not have enough textbooks to send home. As set forth in more detail in my report, no such conclusion may be validly or reliably drawn from the Harris survey. Given the errors and fallacies of the Harris survey, there is no way to determine from Harris's data what percentage of classrooms lack textbooks.

I declare under penalty of perjury that the foregoing is true and correct, and that this declaration was executed at Los Angeles, this 21st day of August, 2003.

Richard A. Berk

Exhibit A

Richard A. Berk
Department of Statistics
University of California, Los Angeles
Room 8130 Mathematical Sciences Building
Los Angeles, California 90095-1554
E-mail Adress: berk@stat.ucla.edu
Tel. (310) 206-8544
Fax: (310) 206-6658

Education

1960-1964: Yale University, B.A. in Psychology
1965-1969: The Johns Hopkins University, Department of Sociology Ph.D. 1970.

Honors and Awards

Elected to the Sociological Research Association
Elected Fellow to the American Association for the Advancement of Science
Paul S. Lazarsfeld Award for methodological contributions from the American Sociological Association

Employment

1970-1973: Assistant Professor of Sociology and Urban Affairs Northwestern University
1973-1976: Associate Professor, Department of Sociology and the Center for Urban Affairs Northwestern University
1976-1987: Professor of Sociology, University of California, Santa Barbara
1983-1987: Director, Social Process Research Institute, University of California, Santa Barbara
1986-1987: Professor of Sociology and Statistics, University of California, Santa Barbara
1988- : Professor of Sociology, University of California, Los Angeles
1990-1997: Director, Center for the Study of the Environment and Society, University of California, Los Angeles
1997- : Professor of Statistics and Sociology, University of California, Los Angeles
1997- : Director, UCLA Statistical Consulting Center

Ongoing Professional Activities

1976-Present: Editor, Evaluation Review (formerly Evaluation Quarterly)
1977-Present: American Statistical Association
1977-Present: Editorial Board, Social Science Research
1978-Present: Law and Society Association
1978-Present: American Association for the Advancement of Science
1986-Present: Institute of Mathematical Statistics
1993-Present: Scientific Advisory Council for the Climate System Modeling Program at NCAR
1995-Present: NRC Committee on Applied and Theoretical Statistics
2000-Present: Advisory Board for Center for Spatially Integrated Social Science

2000-Present: Visiting Faculty Member at the Los Alamos National Laboratories, Statistics Group.

Past Professional Activities of Special Note

1975-1977: Editorial Board, American Sociological Review
1977: Program Committee, Pacific Sociological Association
1977-1980: Executive Council, Criminology Section of the American Sociological Association
1979-1982: Executive Council, Methodology of the American Sociological Association
1979-1983: SSRC Committee on Social Indicators
1980-1982: NAS Committee on Sentencing Research
1982-1985: Board of Trustees, Law and Society Association
1982-1983: Review Panel Member, National Science Foundation, Sociology Program
1984-1985: Visiting Scholar, General Accounting Office, Program in Evaluation and Methodology
1985-1986: California Attorney General's Commission on Drug and Alcohol Abuse
1987: Board of Overseers, General Social Survey
1984-1987: American Sociological Association representative to Section U (Statistics) of the American Association for the Advancement of Science
1986-1987: NAS Working Group on Experimentation in Criminal Justice
1986-1987: Chair, Methodology Section of the American Sociological Association
1985-1991: California Attorney General's Advisory Committee on Data Use and Publications
1985-1991: Board of Directors (85-91), Vice Chairman (89-91), Social Science Research Council
1989-1991: NAS Committee on the Social Consequences of AIDS
1989-1991: NAS Working Group on Global Environmental Change
1991: NSF Task Force on Reorganization of the Biological, Behavioral and Social Sciences Directorate

List of Books

1. Between White and Black: The Faces of American Institutions in the Ghetto, in Supplemental Studies for the National Advisory Commission on Civil Disorders (with P. H. Rossi, D. P. Bordua, B. K. Eidson and R. Grove). Washington D.C.: Government Printing Office, 1968.

2. Collective Behavior. Dubuque, Iowa: W. C. Brown, 1974.

3. The Roots of Urban Discontent (with P. H. Rossi and B. K. Eidson). New York: John Wiley and Sons, 1974.

4. A Measure of Justice: An Empirical Study of Changes in the California Penal Code, 1955-1971 (with S. Lesser and H. Erackman). New York: Academic Press, 1977.

5. Corrections Reform and State Elites (with P. H. Rossi). Boston: Ballinger Press, 1977.

6. Crime as Play: Delinquency in a Middle Class Suburb (with P. Richards). Boston: Ballinger Press, 1979

7. Labor and Leisure at Home: The Organization of the Household Day (with S. F. Berk). Beverly Hills: Sage Publications, 1979.

8. Money, Work and Crime: Experimental Evidence (with P. H. Rossi and K. Lenihan). New York: Academic Press, 1980.

9. Water Shortage: Lessons in Water Conservation Learned From the Great California Drought, 1976-1977 (with T. Cooley, C.J. LaCivita and K. Israel). Cambridge, MA: Abt Books, 1981.

10. The Social Impact of AIDS in the U.S. Cambridge, MA: Abt Books, 1988.

11. Thinking About Program Evaluation (with Peter H. Rossi). Newbury Park: Sage Publications, 1990 (Second Edition, 1999).

12. Just Punishment: An Empirical Study of the Federal Sentencing Guidelines (with Peter H. Rossi). New York: Aldine de Gruyter, 1997.

Contributions to Professional Journals

1. "White Institutions and Black Rage" (with P. H. Rossi, D.P. Boesel, B. K. Eidson and W. E. Groves). Trans-Action, March 1969, pp. 24-31.

2. "Establishing Rapport with Deviant Groups" (with J. Adams). Social Problems, Fall 1970.

3. "Local Political Leadership and Popular Discontent in the Ghetto" (with P. H. Rossi). The Annals of the American Academy of Political and Social Science, September 1970.

4. "The Status of Women in Modern Language Departments: A Report" (with F. Howe and L. Morlock). Proceedings of the Modern Language Association, May 1971.

5. "Race and Class Differences in Per Pupil Staffing Schools, 1970-1971" (with A. Herriman). Integrated Education, January 1972.

6. "Patterns of Vandalism During Civil Disorders as an Indicator of Selection of Targets". (with H. Aldrich). American Sociological Review 37(5): 533-546, 1972.

7. "The Structuring of Normative Judgements Concerning the Seriousness of Crimes" (with P. H. Rossi, E. Waite and C. Bose). American Sociological Review 39 (2): 224-237, 1974.

8. "A Gaming Approach to Crowd Behavior," American Sociological Review, 39(3): 355-375, 1974.

9. "Champagne's Assessment of Legal Services Programs: An Evaluation of an Evaluation," Urban Affairs Quarterly, Spring 1974.

10. "Generalized Performance Measures for Urban Political Systems" (with P. H. Rossi), Social Science Quarterly, Spring 1974.

11. "Performance Measures: Half Full or Half Empty?" in Social Science Quarterly, Spring 1974.

12. "Determinants of White Collar Income: An Evaluation of Equal Pay for Equal Work" (with W. Bridges), Social Science Research, Fall 1974.

13. "Doing Good or Worse: Evaluation Research Politically Re-Examined" (With P. H. Rossi), Social Problems, February 1975.

14. "Descriptive Distortions in Covariance Based Statistics" (with M. Hennessy and R. McCleery), Social Science Research 5(2): 107-126, 1976.

15. "Household Work in the Suburbs: The Job and its Participants" (with C. Berheide and S. F. Berk), Pacific Sociological Review, October 1976.

16. "Broken Homes and Middle Class Delinquency: A Re-Evaluation" (with M. Hennessy and P. Richards), Criminology, February 1977.

17. "The Vagaries and Vulgarities of Scientific Jury Selection" (with M. Hennessy and J. Swan), Evaluation Quarterly 1(1):143-158, 1977.

18. "Discretionary Methodological Decisions in Applied Research," Sociological Methods and Research 5(3):317-334, 1977.

19. "Sex, Earnings and the Nature of Work: A Job Level Analysis of Male-Female Income Differences" (with W. F. Bridges), Social Science Quarterly, February 1978.

20. "A Time Series Analysis of the Impact of a Water Conservation Campaign" (with J. E. Maid and D. M. Hoffman), Evaluation Quarterly 2(1):107-118, 1978.

21. "A Simultaneous Equation Model for the Division of Household Labor" (with S. F. Berk), Sociological Methods and Research 6(4):431-468, 1978.

22. "The Wheels of Misfortune: A Time Series Analysis of Bicycle Accidents on a College Campus" (with N. Juste, M. Johnson and T. Rose), Evaluation Quarterly 2(4): 608-619, 1978.

23. "Doing Good Well: The Use of Quantitative Social Science Data in Advocacy Proceedings" (with J. Oppenheim), Law and Policy Quarterly 1(2): 123-146, 1979.

24. "An Introduction to Estimation Procedures for Pooled Cross-Sectional and Time Series Data" (with D. Hoffman, J. Maki, D. Rauma and H. Wong), Evaluation Quarterly 3(3):385-410, 1979.

25. "A Dynamic Decision Theoretic Perspective on Modeling the Performance of the Criminal Justice System" (with T. Cooley), Social Sciences Research 8:265-286, 1979.

26. "Price and Prejudice: A Variance Components Analysis of Some Causes and Consequences of Regulating Chicago Storefront Banks" (with J. Oppenheim), Law and Society Review 14(1):7-26, 1980.

27. "Reducing Consumption in Periods of Acute Scarcity: The Case of Water" (with T. Cooley, C. J. LaCivita, R. Parker, K. Stedl and M. Brewer), Social Science Research 9(2): 99-120, 1980.

28. "Bringing the Cops Back In: A Study of Efforts to Make the Criminal Justice System More Responsive to Incidents of Family Violence" (with D. R. Loseke, S. F. Berk and D. Rauma), Social Science Research 9(3): 193-215, 1980.

29. "Crime and Poverty Among Ex-Offenders: Some Experimental Evidence" (with P. H. Rossi and K. J. Lenihan), American Sociological Review 45(5): 766-786, 1980.

30. "Does IQ Really Matter: A Study of the Use of IQ Scores for Tracking the Mentally Retarded" (with A. Shih and W. Bridges), American Sociological Review 46(1): 58-71, 1981.

31. "On Non-Linear Hierarchical Log Linear Models and the Iterative Proportion Fitting Algorithm" (with J. Magidson and J. Swan), Sociological Methods and Research 10(1):5-49, 1981.

32. "On the Compatibility of Basic and Applied Sociological Research: An Effort at Marriage Counseling," The American Sociologist, November 1981.

33. "An Overview of Evaluation Strategies and Procedures" (with P. H. Rossi), Human Organizations, Winter, 1981.

34. "A Test of The Stability of Punishment Hypothesis: The Case of California, 1851-1970" (with D. Rauma, S. Messinger and T. Cooley), The American Sociological Review 46(6): 805-828, 1981.

35. "Throwing the Cops Back Out: The Decline of a Local Program to Make the Criminal Justice System More Responsive to Incidents of Domestic Violence" (with D. Rauma, D. R. Loseke and S. F. Berk), Social Science Research 11(3): 245-279, 1982.

36. "Selection Biases in Sociological Data" (with S. C. Ray). Social Science Research 11(4): 351-398, 1982.

37. "Crime and Poverty: Some Quasi-Experimental Evidence" (with David Rauma). Social Science Research 11(4): 318-351, 1982.

38. "Beyond Nine to Five: Sexual Harassment on the Job," editor with M. B. Brewer of a special issue of The Journal of Social Issues 38(4), 1983.

39. "Capitalizing on Nonrandom Assignment to Treatments: A Regression Discontinuity Evaluation of a Crime Control Program" (with David Rauma), Journal of the American Statistical Association 78(381): 21-27, 1983.

40. "An Introduction to Sample Selection Bias in Sociological Data," American Sociological Review 48(3):386-397, 1983.

41. "Prisons as Self Regulating Systems: The Case of California from 1851 to 1980" (with S. Messinger, D. Rauma and J. Berecochea). Law and Society Review 17(4): 547-586, 1983.

42. "The Specific Deterrent Effects of Arrest for Domestic Assault" (with L. Sherman). American Sociological Review 49(2): 261-271, 1984.

43. "Cops on Call: Summoning the Police to Domestic Violence Incidents" (with S. F. Berk, P. J. Newton and D. L. Loseke), Law and Society Review 18(3):479-498, 1984.

44. "The Foundations of Parole in California" (with S. L. Messinger, J. E. Berecochea and D. Rauma), Law and Society Review 19(1): 69-106, 1985.

45. "Does Arrest Really Deter Wife Battery? An Effort to Replicate the Findings of the Minneapolis Spouse Abuse Experiment" (with P. Newton), American Sociological Review 50(2): 253-262, 1985.

46. "Social Policy Experimentation: A Position Paper" (with R. Boruch, D. Chambers, P. Rossi and A. Witte), Evaluation Review 9(4): 387-430, 1985.

47. "Varieties of Normative Consensus" (with P. H. Rossi), American Sociological Review 50(3): 346, 1985.

48. "What a Difference a Day Makes: An Empirical Study of the Impact of Shelters for Battered Women" (with P. J. Newton and S. F. Berk), Journal of Marriage and the Family, June 1986.

49. "Remuneration and Recidivism: The Long-Term Impact of Unemployment Compensation on Ex-Offenders" (with D. Rauma), Journal of Quantitative Criminology 3(1): 3-28, 1987.

50. "How Applied Research Can Save Sociologists from Themselves," The American Sociologist, Summer, 1987.

51. "Anticipating the Social Consequences of AIDS: A Position Paper," The American Sociologist, Fall, 1987.

52. "Errors in Forecasting Social Phenomena," (with Thomas R. Cooley), Climatic Change 11:247-265, 1987.

53. "Dangerous People: A Review of the NAS Report on Career Criminals" (with S. Messinger), Criminology 25(3): 767-781, 1987.

54. "Police Responses to Family Violence Incidents: An Analysis of an Experimental Design with Incomplete Randomization" (with L. Sherman), Journal of the American Statistical Association 83(401): 70-76, 1988.

55. "The Role of Subjectivity in Criminal Justice Classification and Prediction Methods," Criminal Justice Ethics 7(1), 1988.

56. "When Random Assignment Fails: Some Lessons from the Minneapolis Spouse Abuse Experiment" (with G. Smyth and L. Sherman), Journal of Quantitative Criminology 4(3): 209-224, 1988.

57. "Thinking About Hate-Motivated Crime," Journal of Interpersonal Violence 5(3): 334-349, 1990.

58. "An Evaluation of an Undergraduate Course on AIDS" (with Paul Abramson, Joan Sekler, and Monique Cloud), Evaluation Review 13(5): 516-532, 1990.

59. "Drug Use, Prostitution and the Prevalence of AIDS: An Analysis using Census Tracts" Journal of Sex Research 27(4): 607-621, 1991.

60. "Estimates of AIDS Incubation Periods from Convenience Samples" (with Paul Abramson), New Directions in Program Evaluation 46 (Summer): 37-50, 1991.

61. "Toward a Methodology for More Morals," Sociological Methodology, 1991.

62. "A Bayesian Analysis of the Colorado Springs Spouse Abuse Experiment," (with Alec Campbell, Ruth Klap and Bruce Western), Criminal Law and Criminology, 83:170-200, 1992.

63. "The Differential Deterrent Effects of An Arrest in Incidents of Domestic Violence: A Bayesian Analysis of Four Randomized Field Experiments," (with Alec Campbell, Ruth Klap and Bruce Western) American Sociological Review: (5) 57:689-708, 1992.

64. "Between Two Extremes: An Examination of Efficiency and Effectiveness of Community Service Orders and their Implications for the U.S. Sentencing Guidelines," (with Malcolm Feeley and Alec Campbell), University of Southern California Law Review, 66(1): 155-207, 1993.

65. Comment on The Role of Offense and Offender Characteristics in Federal Sentencing. Southern California Law Review, 66(1): 401-404, 1993.

66. "Chance and the Death Penalty (with commentary and rejoinder)," (with Robert Weiss and Jack Boger), Law and Society Review,27: 89-110 1993.

67. "Measuring the Impact of Water Conservation Campaigns," (with Daniel Schulman, Matthew McKeever, and Howard Freeman), Climate Change, 24: 233-248, 1993.

68. "Preliminary Data on Race and Crack Charging Practices in Los Angeles," (with Alec Campbell), Federal Sentencing Reporter, 6 (1): 36-38, 1993.

69. "The Role of Water in Urban Metabolism," in Urban Policies, 94(1): 131-148, 1994.

70. "Three Trends in Evaluation Research," Evaluation Practice, Vol. 15, No. 3, October 1994: pp. 261-265.

71. "Public Perceptions of Global Warming," (with Daniel Schulman) Climate Change, Vol. 29, pp. 1-33. 1995.

72. "Statistical Inference for Apparent Populations," with commentary and rejoinder, (with Bruce Western and Robert Weiss), Sociological Methodology, Vol. 25, 1995. Cambridge, UK: Blackwell Publishing.

73. "Assessing the Capriciousness of Death Penalty Charging," (with Robert Weiss and Cathie Lee), Law & Society Review, Vol. 30, No. 3, pp. 607-626, 1996.

74. "Motivated by Hatred or Prejudice: Categorization of Hate-Motivated Crimes in Two Police Divisions," (with Elizabeth Boyd and Kari Hammer), Law & Society Review, Vol. 30, No. 4, pp. 819-850, 1996.

75. "Just Punishment: Guideline Sentences and Normative Consensus," (with Peter H. Rossi and Alec Campbell), Journal of Quantitative Criminology, Vol. 13, No. 2, pp. 267-290, 1997.

76. "The Ethics of Scientific Research: An Analysis of Focus Groups of Scientists and Institutional Representatives" (with Neil Wenger and Stanley Korenman), Journal of Investigative Medicine, Vol. 45, No. 6, pp. 371-380, 1997.

77. "Evaluation of the Research Norms of Scientists and Administrators Responsible for Academic Research Integrity" (with Stanley Korenman and Neil Wenger), Journal of the American Medical Association, Vol. 279, No. 1, pp. 41-47, 1997.

78. "Deployment of Violence: The Los Angeles Police Department's Use of Dogs" (with Alec Campbell and James J. Fyfe), Evaluation Review, Vol. 22, No. 4, pp. 535-565, 1998.

79. "News Media Portrayals and the Epidemiology of Homicide" (with Susan Sorenson), American Journal of Public Health, Vol. 88, No. 10, pp. 1510-1514, 1998.

80. "Punishment for Unethical Behavior in the Conduct of Research" (with Neil S. Wenger, Stanley G. Korenman, and H. Liu), Academic Medicine, Nov. 73(11): 187-1194, 1998.

81. "Water Conservation." The Southern California Environmental Report Card, Institute of the Environment, UCLA, 1998.

82. "Public Perceptions of Climate Change: A Willingness to Pay Assessment" (with Robert Povall), Climate Change, Volume 41, Number 3/4: 413-446, 1999.

83. "Death Penalty Decisions in Los Angeles County: An Illustrative Analysis Using Skeptical Priors" (with Robert Weiss, Wenzhi Li, and Margaret Farrell-Ross), Sociological Methods and Research, Volume 28, Number 1: 92-116, 1999.

84. "Reporting Unethical Research Behavior" (with Neil S. Wenger, Stanley G. Korenman, and Honghu Li), Evaluation Review, Volume 23, Number 5: 553-568, 1999.

85. "An Evaluation of California's Inmate Classification System Using a Generalized Regression Discontinuity Design" (with Jan de Leeuw), Journal of the American Statistical Association, Volume 94, Number 448: 1043-1052, 1999.

86. "Young guns: An Empirical Study of Persons Who Use a Firearm in a Suicide or Homicide." (with Susan Sorenson), Injury Prevention, Volume 5: 280-283, 1999.

87. "Just Punishments: Federal Guidelines and Public Views Compared," (with Peter Reuter), Federal Sentencing Reporter, Volume 12, Number 1: 27-29, 2000.

88. "Measuring Consensus about Scientific Norms, "with Stanley G. Korenman, and Neil S. Wenger), Science and Engineering Ethics, Volume6, No 2000: 315-340, 2000.

89. " Approximation and Inversion of a Complex Meteorological System via Local Linear Filters" (with Frederic Schoenberg, Robert Powell, Cheng Li, Rong Li, and Robert Weiss). Journal of Applied Meteorology, Volume 40, Issue 3, 2000.

90. Households that Pollute, Southern California Environmental Report Card, Institute of the Environment, 2000.

Contributions to Books

1. "White Institutions and Black Rage" (with P. H. Rossi, D. P. Boesel, B. K. Edison and W. Eugene Groves), in Otherto Revolts, Peter H. Rossi (ed.), MIT Press, 1970.

2. "The Emergency of Mutual Violence in Crowd Behavior: A Case Study of an Almost Riot." in Collective Violence, James F. Short, Jr. and Marvin E. Wolfgang (eds.), Aldine Press, 1972.

3. "The Controversy Surrounding Analyses of Collective Violence: Some Methodological Notes." in Collective Violence. James F. Short Jr. and Marvin E. Wolfgang (eds.), Aldine Press, 1972.

4. "Characteristics of Public Issues as Determinants of Political Behavior" (with P. H. Rossi), in Cities in Change, John Walton and Donald Carns (eds.), Allyn Bacon, 1973.

5. "Data Requirements for Evaluation Research: A Plea to Plan Ahead," in Criminal Justice Research, Emilio Viano (ed.), D. C. Heath, 1975.

6. "Social Science and Jury Selection: A Case Study of a Civil Suit," in Psychology and the Law, Gordon Bermant, Charlan Nemeth and Neil Vidar (eds.); D. C. Heath, 1976.

7. "Some Ethical Implications of the New Jersey Pennsylvania Income Maintenance Experiment" (with P. H. Rossi and M. Boeckman), in Ethics and Social Intervention, H. C. Kelman and D. P. Warwick (eds.), Washington D.C.: Hemisphere Press, 1977.

8. "The Politics of Corrections Reform" (with P. H. Rossi), in Corrections and Punishment: Structure, Function, and Process, Volume VIII, David Greenberg (ed.), Beverly Hills: Sage Publications, 1977.

9. "Feet of Clay in Hobnail Boots: An Assessment of Statistical Inference in Applied Research" (with M. Brewer), in Evaluation Studies Review Annual, Volume III, Thomas Cook (ed.), Beverly Hills: Sage Publications, 1978.

10. "The New Home Economics: An Agenda for Sociological Research," in Women and Household Work, Sarah F. Berk (ed.), Beverly Hills: Sage Publications, 1980.

11. "Money Work and Crime: A Summary of Findings from the Transitional Aid Research Project" (with P. H. Rossi and K. J. Lenihan), in Problems in American Social Policy Research, C. C. Abt (ed.). Cambridge, MA: Abt Books, 1980.

12. "A Didactics Review of Some Recent Statistical Developments with Implications for the Evaluation of Criminal Justice Programs," in Handbooks of Criminal Justice Evaluation, M. Klein and K. Teilman (eds.), Beverly Hills: Sage Publications, 1980.

13. "Evaluating Criminal Justice Legislation" (with F. Borstein and I. Nagel), in Handbook of Criminal Justice Evaluations, M. Klein and K. Teilman (eds.), Beverly Hills: Sage Publications, 1980.

14. "A Retrospective on Prison Reform and Effect" (with P. H. Rossi), in Measuring Social Judgements: The Factorial Survey Approach, Steven Nock (ed.), Beverly Hills: Sage Publications, 1980.

15. "A Further Test of the Stability of Punishment Hypothesis" (with D. Rauma and S. Messinger), in Methodological Developments in Criminological Research, J. Hagan (ed.), Beverly Hills: Sage Publications, 1982.

16. "Where Angels Fear to Tread and Why," in Standards for Program Evaluation, P. H. Rossi (ed.). San Francisco: Jossey Bass, 1982.

17. "Training in Quantitative Methods," in Applied Sociology, H. E. Freeman, Russell R. Dynes, P. H. Rossi, and W. F. Whyte (eds.). San Francisco: Jossey Bass, 1983.

18. "The Criminological Diagnosis of Collective Behavior: An Argument for Disaggregate Labeling" (with L. Mazeoti), in Criminological Diagnosis: An International Perspective, Wolfgang and Ferracuti (eds.), Boston: Lexington Press, 1983, (written in 1974).

19. "Mutual Combat and Other Family Violence Myths" (with S. F. Berk, D. R. Loseke and D. Rauma), in The Dark Side of Families, Finkelhorn et al. (eds.), Beverly Hills: Sage Publications, 1983.

20. "Introduction and Overview" (with Marilynn Brewer), in Beyond Nine to Five: Sexual Harassment on the Job, M. B. Brewer and R. A. Berk (eds.), special issue of The Journal of Social Issues, Vol. 34, No. 4, 1982.

21. "Supply Side Sociology of the Family: The Challenge of the New Home Economics" (with S. F. Berk), in the Annual Review of Sociology, Vol. 9, R. H. Turner (ed.), 1983.

22. "The Scope of Evaluation Activities in the United States" (with Peter H. Rossi), in Evaluating the Welfare State: Social and Political Perspectives, S. B. Spiro and E. Yuchtman Yar (eds.), New York: John Wiley and Sons, 1983.

23. "An Introduction to Applications of the General Nonlinear Model," in Handbook of Survey Research, P. H. Rossi, J. Wright and A. Anderson (eds.), New York: Academic Press, 1983.

24. "Some Consequences of Convenience Samples in Crime and Justice Research," in Proceedings of the American Statistical Association Workshop on Crime and Justice Statistics, A. Gelfand (ed.), Washington D.C.: National Institute of Justice, Bureau of Justice Statistics, 1983.

25. "The Randomization of Arrest" (with Lawrence Sherman), in Randomization and Field Experimentation, R. F. Boruch and W. Wothke (eds.), San Francisco: Jossey-Bass, 1985.

26. "Data Collection Strategies in the Minneapolis Domestic Assault Experiment," in Collection Evaluation Data: Problems and Solutions, H. Freeman, P. H. Rossi and L. Burstein (eds.), Beverly Hills: Sage Publications, 1985.

27. "Program Evaluation," in the Final Report of the Attorney General's Commission on the Prevention of Drug and Alcohol Abuse, Sacramento: Office of the Attorney General, 1986.

28. "A Conceptual Framework for Measuring Norms" (with Peter H. Rossi), in the Social Fabric, James F. Short, Jr. (ed.), Beverly Hills: Sage Publications, 1987.

29. "Causal Inference as a Prediction Problem," in Classification and Prediction in Criminal Justice Decisions, D. Gottfredson and M. Tonry (eds.), Chicago: University of Chicago Press, 1987.

30. "The Economics of Household Production: A Critique," in the New Palgrave: A Dictionary of Economic Theory and Doctrine, J. Eatwell, M. M. Milgate and P. Newman (eds.), Macmillan Press, 1987.

31. "How Applied Sociology Can Save Basic Sociology," in The Future of Sociology, E. Borgatta and K. Cook (eds.), Newbury Park, CA: Sage Publications, 1988.

32. "Causal Inference for Sociological Data," in The Handbook of Sociology, N. Smelser (ed.), Newbury Park, CA: Sage Publications, 1988.

33. "An Empirical Analysis of Police Responses to Incidents of Wife Battery" (with S. F. Berk and F. J. Newton), in Coping with Family Violence: New Research, D. Finkelhorn, G. Hotaling and M. Straus (eds.), Newbury Park, CA: Sage Publications, 1988.

34. "The Social Consequences of AIDS: Some Preliminary Observations," in The Social Impact of AIDS in the U.S., R. A. Berk (ed.), Cambridge, MA: Abt Books, 1989.

35. "A Primer on Robust Regression," in Modern Methods of Data Analysis, S. Long and J. Fox (eds.), Newbury Park, CA: Sage Publications, 1990.

36. "What Your Mother Never Told You About Randomized Field Experiments," in Community Based Care of People with AIDS: Developing a Research Agenda. AHCPR Conference Proceedings, Washington D.C.: U.S. Department of Health and Human Services, 1990.

37. "Current Perspectives on Environmental Problems and Environmental Policies" in Values in Conflict: Policy Interactions in the Pacific Basin, J. P. Montgomery (ed.), CSIA Discussion Paper 91-98, Kennedy School of Government, Harvard University, September 1991.

38. "Time Series Analysis," in Encyclopedia of Sociology, E. F. Borgatta (editor-in-chief), Macmillan Publishing Co., Inc., 1991.

39. "Research Priorities for the Human Dimensions of Global Environmental Change," in Global Environmental Change, NATO ASI Series, Vol I, F. Anderson, J. Eddy and R. Corell (eds.), Newbury Park, CA: Sage Publications, 1992.

40. "Thinking More Clearly About Hate-Motivated Crimes" (with E. Boyd and K. Hamner), in Hate Crimes: Confronting Violence Against Lesbians and Gay Men, G. M. Herek and K. T. Berrill (eds.), Newbury Park, CA: Sage Publications, 1992.

41. "Studying the Human Dimensions of Global Climatic Change," in Global Climate Change and California: Potential Impacts and Responses, N. Dowling and J. Knox (eds.), Davis, CA: University of California Press, 1992.

42. "Arresting Ex-offenders in Spouse Abuse Incidents: What the Scientific Evidence Shows," in Understanding Family Violence, Richard Gelles and Donileen Loseke (eds.), Newbury Park, CA: Sage Publications, 1993.

43. "Hate Crime" in M.S. Hamm (ed.), Hate Crime: International Perspectives on Causes and Control, New York: Anderson Publishing, 1993.

44. "Uncertainty in the Construction and Interpretation of Mesoscale Models of Physical and Biological Processes," Peter Coffman and Gene Likens (eds.), New York: Chapman and Hall, 1994.

45. "Sexual Activity as Told in Surveys," (with Paul R. Abramson and Paul Okami) in P.R. Abramson and S. Pinkerton, Sexual Nature, Sexual Culture, Chicago: University of Chicago Press, 1995.

46. "Statistical Assumptions as Empirical Commitments," (with David A. Freedman), in T. Blomberg and S. Cohen (eds.), Law, Punishment, and Social Control: Essays in Honor

of Sheldon Messinger, Part V/ pp. 245-248; Biographical sketches of contributors page 305. Aldine de Gruyter, November 1995.

47. "Statistical Inference for Sociological Data," in M. Alexander et al., (eds.), Refashioning Sociology: Responses to a New World Order, pp. 215-220, TASA Conference Proceedings, Brisbane, 1996.

48. "Computer Simulations as Experiments: Using Program Evaluation Tools to Assess the Validity of Interventions in Virtual Worlds," (with Jason Bond, Rong Lu, Richard Turco, and Robert Weiss), in L. Bickman (ed.), Donald Campbell's Legacy (Volume II): Contributions to Research Design, Sage Publications, 2000.

49. "Managing the Demand for Water by Subsidizing the Purchase of Water-Efficient Household Technology," (with Daniel Sebaumer) in Sustainable Development: Its Ideology and Practice, Dal-Yeon Jeong (ed.), UNESCO/Korea, Forthcoming.

Exhibit B

Report on "A Survey of the Status of Equality in Public Education in California --- A Survey of a Cross-Section of Public School Teachers"

Richard Berk

Department of Statistics

UCLA

I. Introduction

I was asked by the attorneys for the defendants to review for scientific and statistical credibility the study conducted by Louis Harris Organization, "A Survey of the Status of Equality in Public Education in California --- A Survey of a Cross-Section of Public School Teachers." I was also asked to review claims made by plaintiffs' experts based on the Harris study. To undertake this task, I examined the following materials.

1. "A Survey of the Status of Equality in Public Education in California --- A Survey of a Cross-Section of Public School Teachers" (March, 2002)
2. The "Final Version" of the questionnaire on which the Harris study was based, dated January 2002.
3. The data from the Harris study.
4. The First Amended Claim for Injunctive and Declaratory Relief.
5. Plaintiffs' Liability Disclosure Statement.
6. Expert Report from Robert Corley
7. Expert Report from Glen I Earthman

8. Expert Report from Ross Mitchell
9. Expert Report from Michelle Fine
10. "Access to Textbooks, Instructional Materials, Equipment and Technology: Inadequacy and Inequality in California's Public Schools," Jeannie Oakes.
11. "Education Inadequacy, Inequality, and Failed State Policy: A Synthesis of Expert Reports Prepared for Williams v. State of California, Jeannie Oakes.
12. Untitled document from Jeannie Oakes.
13. *Survey Costs and Survey Errors*, Robert M. Groves, John Wiley, 1989.
14. Sampling (second edition), Steven K. Thompson, John Wiley, 2002.

II. Purpose of the Harris Study

The purpose of the Harris study is clear. At the top of the first page of the Executive Summary we are told that the goal of the study is to "find out what the public school system is providing" to its students. These students attend public schools in California, and in the first summary table we find projections to all "California public school students." Just below the table we are told that table's figures are "results of a sample of overall state-wide results." The text that follows in the Executive Summary and in the body of report further underscores that the findings are meant to apply to all students and all schools in California at the time the study was conducted. In more formal language, the population would seem to be all students in all schools in California in 2002. Then, the report is essentially a needs assessment of what the Harris study calls "key ingredients essential for children to learn ...".

III. Achieving the Harris Study Goals

A useful way to begin an evaluation of the statistical and scientific credibility of a study is to describe what an ideal study might entail. The ideal serves as an important benchmark. In this case, an ideal study to achieve the goals of the Harris report would be to collect data from the full set of schools in California. With respect to the schools being studied, there would be no sampling, and no risk of sampling error; both systematic and random sampling error would be eliminated.

Then, to each one would send "auditors," who had the expertise to document the "key ingredients" of concern. Auditors could actually examine the plumbing in the bathrooms, the performance of the air conditioners, the availability of textbooks, and other features of the educational infrastructure. They could also examine school records to document teacher turnover, the credentials and experience of teachers, and reports of rodents, leaks, broken pipes and the like. While no system of measurement could ever be perfect, one would have access to primary data with two powerful assets: proximity to the features of interest and transparency of the measurement process. For example, to see if the air conditioning works, one could try to turn it on. This measurement is proximate because it addresses directly what is of interest, and it is transparent because it comports with common sense. A similar logic would apply a wide variety of the "key ingredients" that the Harris report discusses.

Finally, a study of all California schools conducted by knowledgeable auditors would have the additional asset of being easily understood. Because there would be no sampling

(and therefore no random sampling error), there would be no need to talk about confidence intervals and statistical tests. Because the measurement process would be proximate and transparent, the tabulations reported could, at least in principle, be taken at face value. A reader of the report could easily understand a statement like "on the day auditors visited each school, X% of the toilets did not flush," or that "Y% of the students did not have textbooks."

It is important to stress that for the ideal study one requires both on-site, systematic auditing and data collection from the population of schools in California. Auditing a convenience sample of schools on a hit and miss basis will not do. Thus, Robert Corley's observations do not qualify.

IV. What the Harris Study Actually Did

It is clear that the Harris study is a very long way from this ideal. First, the data do not come from auditors who systematically assessed each school's educational "ingredients," but from teachers asked to provide broad, subjective generalizations about various features of their school's infrastructure and learning environment. At the very least, proximity and transparency are sacrificed. Second, the manner in which these data are analyzed and reported is highly misleading and contradicts sound statistical practice. The study's conclusions do not follow from the data. Third, the data come from a subset of teachers in California for a subset schools. Thus, there is substantial sampling error to deal with. And the manner in which the sampling was done leads to both random

sampling error and systematic sampling error. Then, efforts after the fact to remove the bias produced by the systematic sampling error are inadequate. Now the details.

Measurement of Teacher, School and Classroom Attributes

Surveys commonly try to elicit three kinds of information.

1. Surveys can elicit respondents' opinions and/or attitudes. There are no facts external to the responses that are of interest. A public opinion poll taken to determine, for instance, whether respondents favor or oppose affirmative action is an illustration. A given respondent can be for affirmative action, against affirmative action, or perhaps be unsure one way or the other. There is no right or wrong answer.
2. Surveys can elicit from respondents facts about themselves. For example, it is common in marketing surveys to ask about respondents' "demographics." The assumption is that there are correct and incorrect answers and that respondents are willing and able to provide the correct ones. Household income is one illustration.
3. Surveys can elicit from respondents facts about other people (e.g., their bosses), collections of people (e.g., their families), or institutions of various kinds (e.g., their places of work). As such, respondents are being used as *informants*. The purpose of the survey is to use respondents to collect facts not otherwise more easily and/or accurately obtained. Once again, the assumption is that there are correct and incorrect answers and that respondents are willing and able to provide

the correct ones. One implication is that if one has access to several informants, each knowledgeable about the same fact, there should be virtual unanimity in the information elicited. For example, if one informant says it is raining and another says it is not, and if there is no other information, the fact of the matter remains unknown.

The key questions in the Harris survey were an effort to elicit facts. Sometimes, the facts were about the respondent (e.g., "What subjects do you teach?"). Sometimes the facts were about the respondent's classroom (e.g., "... During the past year was your classroom uncomfortably hot or cold, or not?"). And sometimes the facts were about the school as a whole (e.g., "Are the student bathrooms in your school clean and open for student use throughout the day, or not?").

For each of kind of question, one must be convinced that every respondent knows the correct answer and is able to provide it. And in the first place, that depends on asking clear questions able to call forth accurate information. In fact, many of the questions in the Harris study are vaguely worded so that it is not apparent what fact is being sought. Consider first questions about the school as a whole.

What does it mean to ask if a classroom is "uncomfortably" hot or cold during the past year? For whom? When? How often? What is the fact that is being sought? What does it mean to ask "Are the student bathrooms in your school clean and open for student use throughout the day, or not." Who defines clean? Which day? All the time? And what does it mean to ask "Have you seen evidence that cockroaches, rats, and mice have been

a problem in your school over the past year, or not." What does either a "yes" or "no" convey about whether at any time there really was a problem in the school with cockroaches, rats, or mice? Questions such as these at best elicit a teacher's subjective assessment. The relationship between that assessment and the facts is unknown.

But the problem goes deeper. For questions such as these, well-meaning respondents will often disagree. Then what? What can be learned about the facts when the individuals polled say different things?

For the Harris study, it is possible to gain some insight about how common the disagreements were. By happenstance, teachers were selected so that for about 100 schools, there were two or more respondents for each. For questions about the school as a whole, Figures 1 and 2 show the percentage of time there was disagreements between the teachers at the same school. For Figure 1, the questions have five response categories: "Excellent," "Good," "Only fair," "Poor," "Not sure." The analysis was conducted excluding all "Not sure" responses. This lowers the estimated amount of disagreement, but the disagreement reported reflects firmly held positions. For Figure 2, the questions have three response categories (e.g., "Are clean And open," "Are not," and "Not sure"), and again, the "Not sure" response is excluded from the analysis.

In Figure 1, we see that about two-thirds of the time when it is possible to check, teachers disagree about the facts for their school. For example, about 63% of the time there was disagreement between teachers about the "The adequacy of physical facilities in your school." Of course, this is not surprising given that the responses are each teacher's

subjective assessment. In Figure 2, about one-third of the time there is disagreement. For example, about 28% of the time teachers disagree about whether there is evidence of cockroaches, rats, or mice. Because this too is a subjective assessment, disagreements are to be expected.

What is one to make of disagreements? Suppose that in a given school one teacher says that the bathrooms "are clean and open" and another says they "Are not." To simply state that in this particular school 50% of the teachers say the bathrooms are not clean fosters a very misleading inference. First, the sample of two respondents is so small that conventional confidence intervals based on simple random sampling cover the entire range from 0% to 100%. The story for samples fewer than six are all about the same and in this study, no school has more than five respondents.

Second, common sense indicates that if one teacher says one thing and another teacher says the opposite, no credible conclusions can be drawn about the bathrooms. Whatever the facts happen to be, there is no agreement about them. And the problem generalizes. Suppose there are four respondents for a given school, and the split is three saying the bathrooms are clean and open and one saying they are not. It would be misleading to simply say that 25% of the teachers in that school said the bathroom were not clean and open without also saying 1) that under simple random sampling the confidence interval covers most of the range between 0% and 100% and 2) the strong majority of teachers in that school actually said the bathrooms were clean and open.

Three conclusions follow. If samples of appropriate size had been collected from each school included in the study, there would have to be overwhelming evidence of substantial differences on the facts. Second, given the small sample actually collected school-by-school, the results from each school are far too unreliable to be taken seriously. Third, when there is disagreement, it is necessarily very unclear what conclusions to draw about the facts. What does it mean if half the teachers in a school say one thing and half say another? Or what does it mean if a third of the teachers in a school say one thing and two-thirds say another? Any attempt to arrive at the facts from an opinion poll of teachers requires a measurement model linking the facts to the teachers' responses. Not only is there no measurement model in the Harris analysis, the entire problem of disagreements within schools is swept under the rug. This violates sound statistical practice and the scientific requirement of honest reporting.

These difficulties are carried along when conclusions are drawn about all California schools, and new problems result. Suppose there actually was a rule allowing one to make inferences about the facts from a vote of observers. For simplicity, assume that there are two response categories, and we adopt the voting rule that the response with the greatest number of votes (i.e., the mode) is taken as the best estimate of the truth. While this is certainly not the only voting rule possible, it has some justification as the "best bet" if teachers are able to convey the true situation, except for some random noise. More complicated voting rules can be formulated using a different response model if there is good reason to adopt it. Suppose, for example, teachers see the survey as a way to lobby for more resources; that is, teachers are inclined to emphasize problems with the schools.

In this case, the votes implying problems within schools might be discounted substantially.

Now, assume for simplicity there are only three schools in California with three respondents in each. Also assume that the Harris survey shows that in each of the schools, two of the three respondents say the bathrooms in their school are clean and open. By our simple voting rule, the best guess for each school is that the bathrooms are clean and open; that is what the majority says.

The Harris analysis would produce a different conclusion. Averaging over all teachers *without respect to school*, 33% say that the bathrooms are not clean and open. But it does *not* logically follow that one school in three has bathrooms that are not clean and open. Nor does it logically follow that 33% of the students in those three schools have bathrooms that are not clean an open. These conclusions, derived from averaging over all teachers without respect to their school, represent fundamental and misleading errors, and just the kinds of errors that follow for the way the analysis was conducted in the Harris study. Moreover, a general outcome of these errors is to overestimate problems in the public schools.

Consider as examples three questions from the Harris study. One asked about instructional materials, a second asked about available technology, and a third asked about the bathrooms. Consistent with the Harris analysis strategy, the responses to the first two questions were collapsed so that "Excellent" and "good" were considered

"satisfactory" with "Only fair," and "Poor" considered "unsatisfactory." For Ninety-nine schools there was more than one respondent, for a total of 214 respondents overall.

The answers to these three questions can be analyzed in two ways: 1) as in the Harris study averaging over teachers or b) first characterizing a school my majority vote and then averaging over schools. For the instructional materials question, about 18% of the teachers felt the materials were unsatisfactory while only about 2% of the schools had unsatisfactory materials. For available technology, about 37% of the teachers felt that the availability technology was unsatisfactory while only about 21% of the schools had unsatisfactory available technology. For the school bathrooms, about 16% of the teachers said the bathrooms were unsatisfactory while about 7% of the schools had unsatisfactory bathrooms.[1]

It is clear from these three questions how the incorrect form of analysis used in the Harris study can overestimate problems in the schools. And the overestimates are substantial. Here, the overestimates range from 180% to 900%.

In short, one cannot try to establish facts from an opinion poll without having a rule by which a credible estimate of the facts is determined. And then once that rule is determined, it cannot be ignored when aggregate statistics are computed. In the Harris study, no voting rule was employed. There is no logical way, therefore, to get from teachers' subjective opinions to the truth of the matter. However, using a reasonable

[1] When at the school level there was a tie, that school was treated as being satisfactory for purposes of this analysis. The analysis averaging over teachers was about the same whether or no the "Not sure" responses were included in the denominator. At the level of the school, the majority vote rule was applied including "Not sure" as a response.

voting rule, the Harris analysis can lead to substantial overestimates of the problems reported in schools.

Even when there is no disagreement within schools, the Harris style of analysis can be terribly misleading. Again suppose that there are 3 schools in the State. The survey data by school are as follows.

1. School A: 1 respondent who says the bathrooms are fine.
2. School B: 1 respondent who says that the bathrooms are fine.
3. School C: 3 respondents who say that the bathrooms are not fine.

In 67% (two-thirds) of the schools, the teachers say that the bathrooms are fine. But 60% (three-fifths) of the teachers say the bathrooms are not fine. If one were to take the survey data seriously, the truth is that the majority of schools in the state are not having problems with their bathrooms. Yet, under the Harris reporting system, one is given the impression that the majority of schools are in fact having problems with their bathrooms. The apparent contradiction could be resolved by proper weighting of the teachers, but there is no evidence that the weights used actually did this.

The errors just discussed affect as well teachers' assessments of their own classrooms and teaching materials. There are, for example, several questions about the quality and physical condition of textbooks. Thus, one question is "How would you rate your textbooks on giving students up-to-date information?" The response choices are

"Excellent", "Good", "Only Fair", "Poor", and "Not sure." Might it not have been more direct for an auditor to determine the publication date of the book?

But if one is to rely on teachers' subjective ratings, what can one to make of what the teachers say? At the college level, I have routinely been part of conversations with faculty evaluating introductory statistics textbooks, all of which are adequate. Yet, one faculty member will judge a given book as "poor" while another faculty member will judge that same book as "excellent."

The point is that for these kinds of subjective judgments, disagreements are to be expected. And then, it is not clear what to do with the results. Unfortunately, one cannot tell from the Harris data which teachers are rating the same textbooks, so there seems to be no way to document disagreement. This is a pity because disagreements would almost certainly surface about textbooks and any number of others. And the real question, never addressed, is how the ratings of textbooks given by teachers translate into how well students learn. In short, while there would be significant disagreements between teachers about the facts for their own classrooms, there is no way explore that with the Harris data.

To summarize, the survey questions used in the Harris study are meant to elicit teachers' subjective generalizations about their teaching environment. While this might make some sense for a study of, say, teacher morale, it makes little sense if the goal is to learn about the quality of key educational ingredients. Likewise, while the statistical analysis of those questions might appear to have clear meaning, that meaning disappears upon close inspection.

Sampling

If the goal is to project from the sample of teachers to all schools in California, and by implication, all students in California, one requires a probability sampling of teachers from that population.

Probability sampling is desirable for three reasons. First, it provides a formal and scientifically valid vehicle for projections from a sample to the population from which it was selected. Second, it provides a formal and scientifically valid way to estimate the uncertainty in such projections. Third, probability samples pass the "sniff test" in that the selection is undertaken with procedures that are absolutely neutral with respect to the issues at hand. That is, there are no grounds for suspicion that the sample was chosen in a manner that favors any particular set of facts. Without probability sampling, all three desirable attributes are lost. And they can be lost because of an inappropriate study design or because a design is poorly implemented.

The data come from three samples purchased from Market Data Retrieval Inc. (MDR). The Harris report describes the samples as follows.

"**Calhome**. A random sample of names and home phone numbers of teachers in MDR's database for California public school teachers..."

"**Caltech.** A random sample of names and school phone number of teachers in MDR's database for California public school teachers.

"**Calholo.** "A random sample of names and school number of teachers in MDR's database for California public school teachers residing in the lower income census tracks" [sic].

The sampling procedures described in the Harris study raise a number of important issues. First, how are the MDR lists constructed? Who is on and who is not? Without such information it is impossible to know what biases may have been introduced by the sampling frame, even before sampling was initiated.

Second, cooperation rates, not response rates, are reported. By common statistical practice, the cooperation rate is "the ratio of completed interviews to all contacted cases capable of being interviewed" (Groves (1989: 141). This seems to be consistent with the calculations in the Harris report, which are presumably meant to convey that no significant biases were introduced because only a subset of people contacted completed interviews. From the appendix of the report, we learn that the cooperation rate for the Calhome sample is 61%, the cooperation rate for the Calholo sample is 72%, and the cooperation rate for the Caltech sample is 84%. These figures are taken to be acceptable by the Harris analysts.

But according to conventional standards "the most universally endorsed" calculation of the possible biases from non-response is the response rate Groves (1989: 141), which

adds to the denominator of the cooperation rate the number of no-contacts, and other potential contacts not made. When this definition is applied to Harris survey, the response rate for the Calhome sample seems to be about 18%, the response rate for the Calholo sample seems to be about 25%, and the response rate for the Caltech sample seems to be about 18%. For high quality survey research with a great deal riding on the outcome, these response rates are disappointing. They are a clear signal that serious biases could well be present in the study regardless of the quality of the MDR sampling frames.

Even if there are no problems with the MDR sampling frames, the low response rates also undermine any attempts to construct meaningful confidence intervals and tests. The low response rates likely degrade the probability sample beyond recognition and likely invalidate all statistical inference.

Is there a way to fix the problems with the sample after the fact? The Harris study proceeds as if it can done. Comparisons are made between 1) distributions for some key characteristics of the three samples and 2) distributions for these same characteristics in the population of teachers, as reported in statewide figures apparently available from the California Department of Education. The reported figures from the two sources are much alike, which is taken as evidence that the sample is representative. Nevertheless, before the reported analyses were undertaken, the data were weighted to still better match up to available statewide figures (See the Harris report technical appendix).

However, weighting of this sort will at best achieve comparability only on the variables used in the weighting. According to the Harris report technical appendix, these variables are gender, ethnicity (i.e. Latino v. Anglo), kind of school (i.e. elementary school, middle school, and high schools), and poverty (i.e., the proportion of LEP students in each school, and the proportion of children eligible for free/reduced meals in each school).

Clearly, there are a host of variables for which weighting was not undertaken, and these could severely bias the results: teacher seniority, teacher ethnicity, size of the school, ethnicity of the students beyond Latino v. Anglo, school drop out rates, crime rates in the surrounding neighborhoods, and so on. For example, even after the weighting, we do not know if teachers in smaller schools from rural areas are underrepresented or if teachers near retirement are overrepresented. But perhaps more important than the variables for which in principle weighting could be done, are any number of variables representing a teacher's views on the issues for which weighting is effectively impossible. In particular, when there are low response rates, survey researchers rightly worry that the mix of completed interviews will substantially over represent respondents who have a particular axe to grind. These are the respondents who will be more inclined to agree to be interviewed and then more likely to complete the interview. In the Harris study, one obvious risk is that the completed interview will over represent the teachers who are highly motivated to lobby for more resources, who are strongly critical of the education infrastructure in their classrooms and schools, or who have grievances against their local schools.

Data Analysis

The tabulations reported are relatively straightforward, even with the weighting. But how to combine data from the three samples in those tabulations is not straightforward, nor is how to calculate confidence intervals and tests of statistical significance.

Even if one ignores the serious problem with low response rates, the study *by design* over samples for teachers living in low-income census tracts. Recall that the Calholo sample is designed to reach teachers residing in low-income census tracts. When analyses are undertaken with the full set of data, all of the observations must be properly weighted to represent the true mix of teachers. In effect, the responses of teachers living in low-income census tracts need to be given relatively less weight than the responses of teachers living elsewhere. But how much less? The poverty measures used in the weighting are the proportion of LEP students and the proportion of students eligible for free or reduce price meals. For the weighting to work, these two variables must be very closely related to the income levels in census tracts where teachers reside. A key assumption is that teachers, by and large, live very close to where they work. In fact, one learns from the Harris report that in previous research using a similar design, a substantial fraction of the teachers living in low-income census tracts did not teach in schools serving low-income students (Technical Appendix). Moreover, by over sampling teachers living in low-income census tracts, the study is also over sampling for variety of other attributes common among such teachers. These might include being in a single earner household, having a lower income (because of less seniority) and various ideological predispositions. In short, the weights used to compensate for the overrepresentation of teachers living in low-income census tracts are almost certainly not getting the job done.

The result is bias in the findings reported; the views of teachers living in low-income census tracts are likely to be overrepresented.

More subtle are the potential biases introduced by combining the Calhome samples and Caltech samples without explicit weights to accurately represent the population of teachers in California. One would need to know the probability that each teacher in California would appear on either (or both) of the MDR lists. Without this information, proper weights cannot be constructed. And without proper weights for these two samples, generalizations from the data to all teachers in the State risk serious biases

As noted above, at least as problematic are the confidence intervals (and margins are error) reported in the Harris study. These depend on proper calculations of the standard errors (Thompson, 2002). The Harris report does not explain how the standard errors were computed, but it is likely that they are substantially wrong.

1. Given the disagreements between teachers in the same school, there is within school variance that must be taken into account. Alternatively, the within school variance may be viewed as random measurement error. In either case, this additional source of uncertainty was ignored and the reported margins of error are too small. There is less precision in the estimates than is reported.
2. When analyses are undertaken for subsets of respondents (e.g., when the "Not sure" response is dropped from the analysis), the sample size for that analysis becomes a random variable and another source of uncertainty. Conventional

expressions for the standard errors, based on a fixed sample size, are incorrect. Again, false precision results.

3. The low response rate leads to substantial violation of the assumption of random sampling. As a result, the computed standard errors, and everything that uses them, are essentially nonsense.

V. Comparisons between Different Kinds of Schools

The Harris report makes a great deal out of comparisons between schools characterized by students' income levels and ethnicities, operationalized as an "Index of Risk." Estimates are provided, for example, about how much worse off across California the students in the highest risk schools are compared to students in the lowest risk schools. While one can certainly quarrel with the way "risk" is measured, the litany of problems already discussed make the risk index moot. If it is impossible to place any credence in the estimates for all schools, it is ridiculous to make something out of comparisons across different kinds of schools.

VI. Reliance upon the Harris Study by the Plaintiffs' Experts

It is clear that several of the Plaintiffs' key experts rely on the Harris study. In particular, Professor Oakes in her synthesis of the Plaintiffs' expert reports frequently and at length cites the Harris findings (e.g., on page 19, 21, 22, 22, 24, and 25). Professor Oakes relies on these findings uncritically when in fact the study is badly flawed. Moreover, it is clear from Professor Oakes's deposition that while she had little grasp of the technical aspects

of the Harris report, she makes pronouncements from it nevertheless (See, for example, pages 751-779).

VII. Conclusions

There are surely schools in California whose educational infrastructure is unsatisfactory. And just as surely, many of those schools are in low-income areas. But, the Harris study adds nothing credible to these observations. Indeed, the systematic provision of misinformation is a disservice. If the goal is to estimate the number of California students statewide who are being served inadequately by the public schools, the study fails. If the goal is to single out particular features of schools across the State that are in particular need of improvement, the study fails. And if the goal is to estimate with any real precision disparities in the distribution of education resources, the study fails.

There remain several issues about the Harris study that are still being explored, in part because some depositions from the Plaintiffs' experts have yet to be taken. When this additional information becomes available, the views expressed above may change.

Percentage of Time Respondents Disagreed about Conditions at the Same School

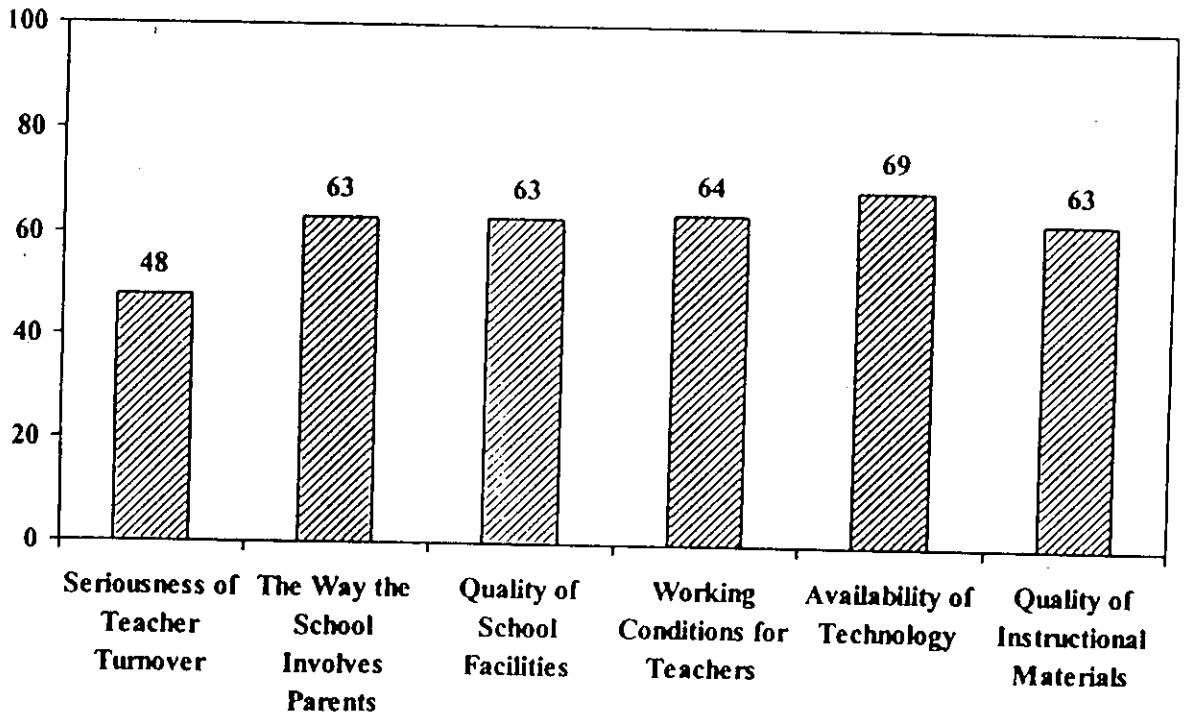

Figure 1

Percentage of Time Respondents Disagreed about Conditions at the Same School

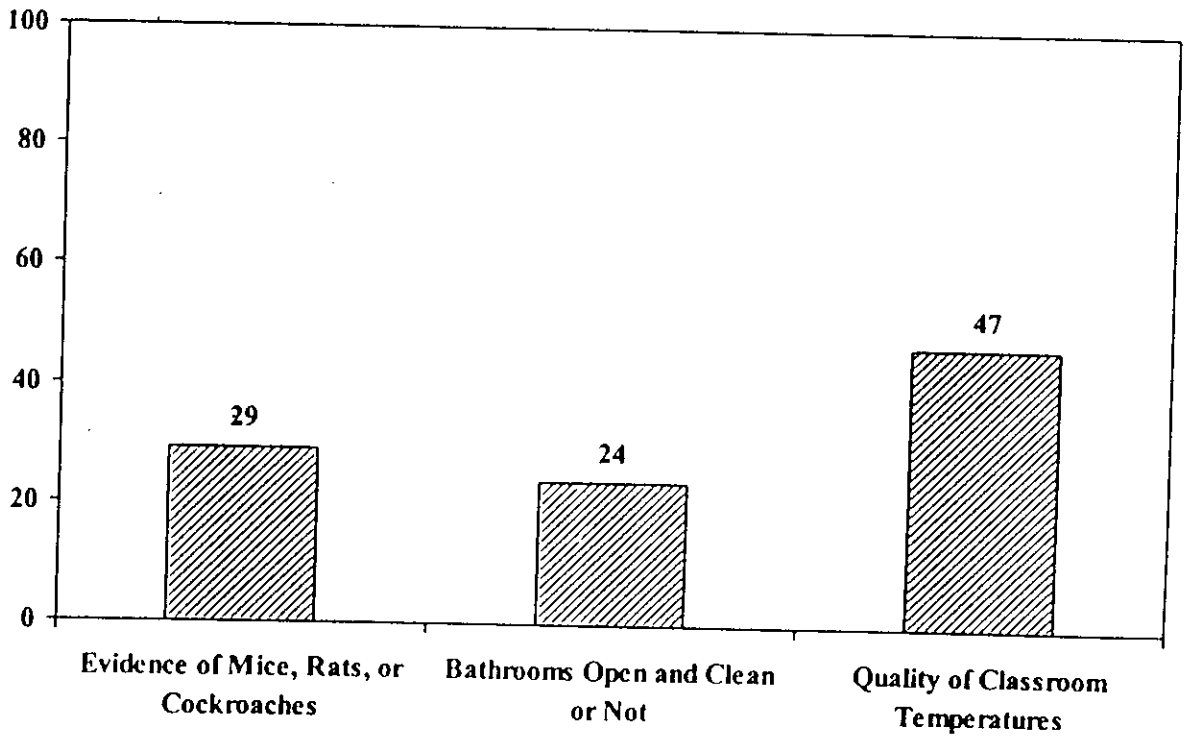

Figure 2