

Maurer School of Law: Indiana University Digital Repository @ Maurer Law

Articles by Maurer Faculty

Faculty Scholarship

1978

Social Control in Applied Social Science: A Study of Evaluative Researchers' Conformity to Technical Norms

Ilene Nagel Bernstein
Indiana University School of Law

Follow this and additional works at: <http://www.repository.law.indiana.edu/facpub>

 Part of the [Quantitative, Qualitative, Comparative, and Historical Methodologies Commons](#), and the [Social Psychology and Interaction Commons](#)

Recommended Citation

Bernstein, Ilene Nagel, "Social Control in Applied Social Science: A Study of Evaluative Researchers' Conformity to Technical Norms" (1978). *Articles by Maurer Faculty*. Paper 2079.
<http://www.repository.law.indiana.edu/facpub/2079>

This Article is brought to you for free and open access by the Faculty Scholarship at Digital Repository @ Maurer Law. It has been accepted for inclusion in Articles by Maurer Faculty by an authorized administrator of Digital Repository @ Maurer Law. For more information, please contact wattn@indiana.edu.

Social Control in Applied Social Science: A Study of Evaluative Researchers' Conformity to Technical Norms

ILENE NAGEL BERNSTEIN

Indiana University

This paper is a preliminary exploration of the relationship between social factors, and conformity to a set of prescribed methodological norms in applied social science. Focusing our attention on evaluative research, we seek to estimate how variation in type and nature of research sponsorship, research context, and researcher relationships with sponsor and host affect reported conformity to methodological prescriptions. Analyzing the self-reported responses of 152 evaluative researchers to a mail questionnaire, we find: (a) that conformity to methodological prescriptions is very variable among evaluative researchers; (b) that the social factors here examined seem to affect systematically the degree of conformity; (c) that while no single social factor has a large net effect on conformity, simultaneously occurring values seem more conducive to conformity, i.e., characteristics associated with our "academic model" are correlated with reported higher conformity, whereas characteristics associated with our "entrepreneurial model" are correlated with reported lower conformity. Our findings suggest that, while traditional social control mechanisms increase the likelihood of adherence to methodological prescriptions for those whose work fits or resembles the "academic model," when the model of work moves toward the "entrepreneurial" type, reported conformity decreases with the absence of those traditional mechanisms of social control. Insofar as the "entrepreneurial model" is increasingly becoming the predominant work model in applied social science, our findings suggest that future research should seek to explain variation in conformity among the "entrepreneurs" and to explore the variety of means by which to increase conformity within this model.

INTRODUCTION

Although the application of social science to practical problems of policy and action is still in its early stages, a large body of experience has been accumulated. . . . The experience is there,

Funds for this research were provided by the Russell Sage Foundation. Special thanks are extended to Howard E. Freeman for the central role he played as a collaborator in this research effort. Thanks too are extended to Richard Berk, Peter Burke, Lowell Hargens, and Peter Rossi for their comments on earlier drafts of this manuscript.

Requests for reprints should be sent to Ilene Nagel Bernstein, Department of Sociology, Indiana University, Bloomington, Indiana 47401.

but it has not been systematically reviewed or codified. Consequently, no one knows the present status of applied social science. . . . (Merton, 1949, p. 161).

According to Merton (1949), one neglected responsibility of social science is to study the performance of professional social scientists. As part of that study, he suggests we explore the conditions that limit and make for achievement. In a way, what Merton was calling for nearly three decades ago was the systematic inquiry into the models of social control in applied social science.¹ Of interest is that Cole and Cole (1973, pp. 1-7), in reviewing the sociology of science, note that "Interest in the social organization of science and the ways in which this organization affects scientific development has been slow in coming." This is especially true for the sociology of social science and even more so for the sociology of applied social science.

Kornhauser (1962), Hagstrom (1965), Marcson (1966), Storer (1966), Collins (1968), Boalt (1969), and Blume (1974) have all addressed the question of social control in science. Perhaps the most definitive thesis is Hagstrom's (1965) study of conformity to and deviation from norms in basic science and the social factors affecting it. Using the concepts of socialization, exchange, and social control, Hagstrom (1965, p. 52) explains the institution of science as follows:

The thesis is that social control in science is exercised in an exchange system, a system wherein gifts of information are exchanged for recognition from scientific colleagues. Because scientists desire recognition, they conform to the goals and norms of the scientific community. Such control reinforces and complements the socialization process in science. . . .

Importantly, for our purposes, Hagstrom finds evidence to suggest that there is substantial variation in the degree to which scientists conform to norms, and that certain social factors significantly affect the degree of conformity. However, Hagstrom (1965, p. 294) concludes his thesis by stating that, to the extent that some groups of scientists may be more strongly influenced than others by nonscientists, future research should reexamine his thesis. One might argue that Hagstrom's call for reexamination results partly from the fact that Hagstrom believes, as do Bernal (1939), Polanyi (1951), and Kuhn (1962), that, as science moves from an occupation of curious and ingenious minds supported by wealthy patrons to an industry supported largely by the state, there will be an accompany-

¹ According to Hagstrom (1965, p. 1), the study of social control in science is a study that seeks to determine those factors conducive to conformity to or deviation from scientific norms.

ing decrease in the influence of scientists on science and a concomitant increase in the influence persons external to the scientific community have on scientific work. Furthermore, Polanyi (1951), Kuhn (1962), and Price (1968) have posited that the increased power the State has over scientific work will result in a decreased motivation of scientists to excel. Their argument is predicated on the assumption that freedom to select and define one's own research problem is a central norm of science and the primary incentive to do quality work (see also Roe, 1951; Kornhauser, 1962).

The pursuit of science can be organized therefore, in no other manner than by granting complete independence to all mature scientists. . . . The function of public authorities is not to plan research, but only to provide opportunities for its pursuit. . . . To do less is to neglect the progress of science; to do more is to cultivate mediocrity and waste public money (Polanyi, 1951, p. 90).

Blume (1974) argues that the question is no longer one of debate, i.e., to argue as Polanyi (1951), Price (1968) and others did over *whether* science should be influenced by external forces; instead, the task now is to begin to estimate the effects of these external forces.

Most generally, the research reported herein is addressed to a preliminary exploration of the question first raised by Merton (1949), i.e., how do social factors limit and make for achievement in applied social science. Focusing on the most rapidly expanding form of applied social science research, i.e., evaluative research, we examine one aspect of achievement: conformity to a set of methodological norms, and how a variety of social factors affect such conformity. More specifically and related to Blume's (1974) question, we concern ourselves with the variation on these social factors, e.g., whether the research is funded as a result of a judgment by a peer review panel, such as in the case of grants, or as a result of a judgment by a review not relying on peer scientists, such as in the case of contracts, and how the grant/contract difference correlates with conformity. The study of evaluative researchers is particularly suited to a preliminary exploration of these questions since there is considerable variation on the variety of social factors here considered.

In order, we examine (1) if within this group of funded applied research studies, there is substantial systematic variation in conformity to a set of prescribed technical norms; (2) if so, can the set of social factors here considered account for a significant amount of that variation in conformity; (3) what is the gross and net effect of each of the social factors on conformity; (4) how do variations on the social factors specifically relate to conformity; and (5) based on the results, what are the implications for

understanding social control in applied social science research. By so doing, we hope to lay the groundwork for further research on the sociology of social research, especially the kind of social research that is clearly subject to external influences.

DATA AND METHODS

Sample

The sample used for the analyses that follow consists of 152 evaluative researchers who received grant or contract support in the amount of \$10,000 or more, directly from an agency of the federal government, in fiscal year 1970, for the purpose of measuring the process (implementation) and impact (effect) of a large-scale social action program in one of the areas of health, education, welfare, manpower, income security, public safety, and/or housing.² The data come from these researchers' responses to a mail questionnaire.

The study of evaluative researchers is strategic for addressing the questions raised because the proliferation of program evaluation studies has been accompanied by an increase in diversity of styles of organizing research. Moreover, by limiting our analyses to one kind of applied social science research, we can somewhat reduce the amount of variation in conformity that might alternatively be explained by a relationship between certain types of research problems and adherence to a set of methodological prescriptions. As stated above, our sample includes only evaluators whose research tasks included the measurement of program implementation *and* program impact; and, this definition of the research task was determined both by our screening of proposal abstracts and by confirmation by the evaluators that they agreed with this task definition. By limiting our sample to evaluators, all of whom received their initial funding in the same 1-year time frame, we add a measure of control for changes in the definition of appropriate methodologies that might result

² The data on the 152 evaluative researchers are drawn from a larger sample of evaluative researchers surveyed as research studies by Bernstein and Freeman (1975). Bernstein and Freeman surveyed the total population of evaluation studies first funded by a federal agency in fiscal 1970 ($N = 382$). Of those 382, they report that 83% of the 382 responded to their mail questionnaire, netting a sample size of 318. Of the 318 questionnaires returned, 82 included responses indicating that, despite extensive prescreening, the study named was not an evaluative study. Bernstein and Freeman comment upon these 82 cases, but include in their analysis only the remaining 236 cases. They further report that their available data on respondents and nonrespondents showed no differences between the two groups. Of the 236 studies included in their review, 152 were *comprehensive evaluations*, i.e., evaluative studies that included *both* procedures to measure process (program implementation) and impact (program effectiveness).

Importantly, the analyses and discussion reported here differ significantly from those in Bernstein and Freeman (1975) insofar as here we examine data for a subsample and focus our attention on a subset of factors selected to address in greater depth the theoretical questions posed in our introduction.

with the passage of time. Finally, our exclusion of similar studies that were unfunded, funded by some agency other than a federal funding agency, funded indirectly, or funded with a research budget (independent of the program budget) of less than \$10,000 was motivated by our concern for circumscribing a sample of research studies that were more similar than dissimilar to one another, without losing the diversity in social factors here examined, the effect of which we wished to track.

Our study is limited by the fact that we could not make use of an experimental design, i.e., we could not persuade the government to agree to assign randomly the evaluative studies to a variety of research organization conditions. As such, our analyses can only make statements about corelationships rather than causal relationships. Moreover, in the absence of a true experiment, we cannot control for self-selection effects. While we can make some assumptions about their strength as rival hypotheses, future research will have to probe more directly into the degree to which self-selection alters, modifies, or refutes our findings. These limitations notwithstanding, we proceed with our presentation, keeping in mind that our research is exploratory; nonetheless, it is the first effort to examine systematically one aspect of achievement in applied social science research.

Measuring Conformity

According to Merton (1942), Barber (1952), Kuhn (1962), Hagstrom (1965), Storer (1966), and Blume (1974), there exists in science a dominant set of institutional norms that proscribe and describe appropriate moral and technical behavior. While there has been some debate as to the degree to which the moral norms actually guide scientists' attitudes and behavior (Mitroff, 1974), that there are a set of agreed-upon prescriptions for technical procedures seems not to be challenged.³

Kuhn (1962, pp. 5-6) argues that normal science rests upon the assumption that the scientific community shares definitions of appropriate problems for study as well as *legitimate methods for conducting the various studies*. In the process of articulating these definitions, a set of standard procedures for conducting research is set forth (1962, pp. 27, 76). According to Hagstrom (1965), Storer (1966), and Blume (1974), scientists con-

³ This is not to deny that debates over approaches or orientation in science do not occur. Rather, it is to accept a Kuhnian (1962, pp. 37, 76) notion of science that argues that, *at any particular point in time, there is general consensus about appropriate methodological procedures for researching specific types of scientific problems*. Since we limit ourselves to persons conducting comprehensive program evaluations, each seeking to identify the degree to which program implementation causes changes on the desired outcome measures, it seems appropriate to assume that there is general consensus about prescribed methodologies: and, apart from the few dissensions reviewed in Footnote 4, a review of the evaluation literature would seem to support such an assumption of consensus.

form to these norms because it allows them to continue participation in the exchange system whereby competent responses and recognition are exchanged for efforts that evidence adherence to methodological norms.

One way of measuring conformity to technical norms is to select a set of norms, to examine adherence to each norm individually, and then to construct a composite index and measure the degree of adherence to it. In order to determine which norms should be examined and what they specifically prescribe as methodologically appropriate, it is necessary to examine the methodological literature that sets forth guidelines for conducting research.

A systematic review of the evaluative research literature reveals that six basic methodological issues are repeatedly addressed: (1) the method used to select the sample population; (2) the type of data analytic techniques used, e.g., multivariate statistical analysis; (3) the general nature of the data analytic process, e.g., quantitative; (4) the type of research design, e.g., experimental design; (5) the validity of the measurement procedures; and (6) the representativeness of the sample to the appropriate universe.

Generally, it is asserted that *the sample population to be observed* should be selected as a probability sample (see for example, Suchman, 1967, pp. 102–103; Anderson, 1975; Bernstein, 1975). By randomly selecting the group(s) to be observed, certain rival hypotheses, e.g., selection factors, maturation, become less of a threat to problems of internal validity. Second, it is asserted that *multivariate statistical procedures should be used* (see for example Wholey, Scanlon, Duffy, Fukumoto, and Vogt, 1970, Chap. 6; Alwin and Sullivan, 1975; Eber, 1975). Insofar as evaluative research implies a need for causal analysis, and insofar as evaluative researchers can rarely rely on true experimental designs to rule out threats to internal validity, multivariate statistical techniques are prescribed as a substitute in the absence of random assignment of persons to control and experimental groups. Third, it is argued that evaluative research should be predominantly quantitative in terms of the nature of the data analytic processes (see for example, Rossi, 1972, pp. 3, 16, 46). This present preference for quantification may in part be a reaction against the qualitative evaluations that dominated until the mid-1960s. As Bernstein and Sheldon (1977) note, qualitative methods were often used when the fear of a negative evaluation was present, since softer measures could be reinterpreted more easily than hard empirical data to cover up for program failure. Moreover, the expressed preference for quantitative methods over qualitative, or even over part qualitative and part quantitative, stems in large part from the fact that: (a) the "ideal" model for evaluative research is the experimental model (Riecken and Boruch, 1974); (b) implementing such a design, or an approximation of it, e.g., quasi-experimental, requires quantitative data analysis; and (c) quantitative analyses are far more

readily suited to the making of causal statements. Experimental designs focus on the problem of attribution, and attribution implies causality.

Fourth, with respect specifically to the type of research design, Rossi (1972, p. 16) argues, "[The] main methodological model [for evaluative research] is the controlled experiment and its derivatives." Similarly, Riecken *et al.* (1974), Boruch (1974), and others have argued that the experimental *or* quasi-experimental design model is the most desirable and appropriate one as long as evaluations seek to demonstrate causality (see for example, Campbell, 1971; Cain and Hollister, 1972; Stanley, 1972).⁴ Fifth, the importance of reliable and valid *measurement procedures* is underscored by Hyman and Wright (1967, p. 743), Suchman (1967, p. 110), and Nunnally (1975). Last, it is said that the *sample(s) observed* should be a reasonable sample of the universe to which one wants to generalize (see Suchman, 1967; Bernstein, Bohrnstedt, and Borgatta, 1975; and Cook and Campbell, 1975). Suchman (1967, p. 103) states: "If the results [of the evaluation] are to be meaningful, the group one uses in the evaluative project should be representative of the target group for the full scale operating program."⁵

Analysis Procedures

In accordance with the literature review referred to above, a six-item index of methodological prescriptions was constructed to serve as the dependent variable, i.e., the measure of conformity to basic technical norms. To determine the degree of conformity, the evaluative researchers' responses to questions ascertaining *what methodological procedures were used* were coded for each of six items. The six items include: method used for sample selection, type of data analytic techniques used, nature of data analyses procedures, type of research design, validity of measurement procedures, and representativeness of the sample.⁶ Importantly,

⁴ Weiss and Rein (1970) and Edwards, Guttentag, and Snapper (1975) argue that alternative approaches, e.g., a decision theoretic approach, may be more useful to evaluation given the political context in which it is ordinarily conducted. However, even the critics of the experimental model do not reject it so much as they see it as too difficult to implement. The fact that Boruch (1974) has mustered substantial evidence to the contrary, i.e., he has delimited an extensive list of experiments successfully mounted in evaluative research, coupled with the absence of evidence for general acceptance of the decision theoretic approach, leads us to conclude that, at this point in time, the general consensus still supports the use of the experimental model.

⁵ Williams and Evans (1969, pp. 118-132) and Wholey *et al.* (1970, p. 98) similarly argue that the only meaningful evaluations are those executed in program settings representative of settings likely to be used for similar programs and on target populations representative of likely future target populations.

⁶ For example, the questionnaire asked "Does the research plan to measure impact or change include an experimental or quasi-experimental design?" and "Instead of or in addition to an experimental design, are any of the following approaches utilized? Check all that apply: Longitudinal study without control or comparison groups; cross-sectional study without control or comparison groups: . . ."

however, each item was first analyzed separately by means of tabular analysis to determine the appropriateness of using a composite index. Since the pattern of relationships between each of the independent variables and each of the dependent variables separately was consistent across all six dependent variables, it seemed appropriate to create one composite index (the construction of the index is explained in detail below). Furthermore, we computed an α measure, according to Bohrnstedt (1969), to determine the reliability of the constructed composite index. The α for our six-item index is .69.

In order to assign scores to each questionnaire response for each of the six items, we referred back to the literature review and rank ordered the responses for each item, giving the highest score to the procedure most highly recommended and the lowest score to the procedure least recommended. For example, on the item "research design," a "3" was assigned to the use of an experimental *OR* quasi-experimental design *with both* randomization and control groups; a "2" was assigned to the use of an experimental *OR* quasi-experimental design *with either* randomization or control groups; a "1" was assigned to the use of a longitudinal or cross-sectional design *without* a control or comparison group; and, a "0" was assigned to the use of descriptive testimony or narratives as the research design.⁷

To construct the index, each of the six items, i.e., methodological procedures, was weighted by $1/m-1$ where m was the number of response categories. In this way, each of the six items has equal weight in the composite scale. The total score was equal to the sum of the products of the response code multiplied by the weight for each item. The total scores range from 0 to 6, 6 representing the highest degree of conformity.⁸ The mean score on the index is 4.0; the standard deviation is 1.4.

To estimate the effects of the social factors on the index of conformity, dummy variable regression procedures are used.⁹ To estimate net effects,

⁷ An example of a descriptive design is the case of an evaluator who reported measuring program impact by asking the program staff to describe the program effects at several points during the intervention. An example of a narrative design is the case of an evaluator who reported that he observed the program in action and used his own verbal description as evidence of program effect.

⁸ Using self-reported responses may cause concern about the reliability of such responses. However, given that the questionnaire asks "What methodological procedures were followed," and given it is common knowledge that certain procedures are preferred, we suggest that, if any bias exists, it is in the direction of inflating the degree of conformity. Furthermore, respondents did know that access to materials that would verify or question their responses could be made available to us. Thus, we suggest that the likelihood of false responses was low.

⁹ While some may question the appropriateness of regression analysis with these kind of data, the arguments by Cohen (1968) and Kerlinger and Pedhazur (1973) are quite convincing. Moreover, as indicated earlier, the analyses were run using tabular formats too, i.e., each of the independent variables was run against the dependent variable "conformity,"

i.e., the amount of unique variance, we deleted, in turn, each set of dummy variables used to represent the particular independent variable from the full equation. Thus, the net effect is the difference between the R^2 when all the dummy variables are entered into the equation and the R^2 when the set of dummies for the particular independent variable in question was deleted.

Independent Variables

In determining a set of independent variables to consider, our intent was to include social factors that tap some aspects of the external influence, the kind of which Polanyi (1951), Hagstrom (1965), and Price (1968) posited to be likely to affect conformity to scientific norms. Accordingly, we examine the effect of each of the following on our measure of conformity to methodological norms: (1) the federal agency that sponsored the evaluator's research effort, e.g., NIMH, LEAA; (2) the nature of the research award given the evaluator, e.g., contract, grant; (3) the amount of research funds allocated for the evaluator's research; (4) the number of years for which funding was approved; (5) the type of organization with which the evaluative researcher is affiliated, e.g., profit, nonprofit, university; (6) the major audience (as defined by the evaluator) to whom research findings are addressed, apart from the research sponsor; (7) the formal organizational arrangement between the organization administering the social action program being evaluated and the organization with which the evaluator is affiliated, e.g., the evaluator is affiliated with the same organization as that administering the action program, i.e., the evaluator is an inside evaluator; (8) the informal working relationship between the evaluation staff administering the action program vis-à-vis research decisions, e.g., decisions concerning research design are made jointly between the two staffs; and (9) the informal working relationship between the staff of the federal agency sponsoring the evaluation and the evaluation staff vis-à-vis research decisions.¹⁰

once using conformity as a seven-field category (0-6) and once collapsing the categories into three groups arbitrarily called high, medium, and low conformity. As expected, the results of the tabular analyses using nonparametric tests are identical to those reported for the regression analyses.

¹⁰ It may seem that certain critical variables, although not representative of the kind of external influence about which Polanyi was concerned, are left out in our analyses. For example, such variables might relate to characteristics about the evaluative researcher. We should note here that, in prior analyses, we found that the highest degrees of the evaluators as well as their years of research experience had no significant effects on conformity. In fact, it was their lack of effect that partly stimulated our interest in homing in on the factors considered here. Since our research is exploratory, we suggest that future analyses gather more in-depth data on both the characteristics of the researcher as well as the social factors considered here so as to allow for a meaningful comparison of their relative effects on conformity.

Results

First, in order to determine whether or not there is deviation from the index of technical norms, we examined the distribution of scores. Given that the norms were described in the literature as basic to any well-done evaluation study, ideally, one would expect the majority of scores to be near perfect. Our analyses indicate that 11% ($N = 16$) of the 152 evaluative researchers used all six prescribed methodological procedures, i.e., attained a score of 6 on our index. Clearly, deviation from technical norms is high in evaluative research. If one recalls that our data are based on self-reported responses, this finding becomes more significant insofar as if the self-reported nature of the data collection is biased, the bias is almost certainly in the direction of producing an inflated rather than deflated estimate of conformity.

The finding of substantial deviation from technical norms is consonant with the conclusions of those who have previously examined scientific research (Hagstrom, 1965) and policy research, e.g., Whyte and Hamilton (1964), Wilkins (1969), Wholey *et al.* (1970), and Stromsdorfer (1972). Wholey *et al.* (1970), for example, in reviewing evaluative research funded by four federal agencies, concluded that few significant program evaluations have been undertaken, and most of those carried out were poorly conceived, poorly executed, and poorly disseminated.

Second, in order to determine the degree to which the variation in conformity found could be explained by the social factors delimited earlier, we regressed the index of conformity on the nine factors. The R^2 is .301, significant at the .001 level.¹¹ Thus, we conclude that the social factors examined here do have a statistically significant effect on conformity.

Third, to estimate the effect of each of the nine factors individually on conformity, we regressed the index of conformity on the set of dummy variables used to represent each factor. Table 1 presents the gross and net effects for each variable. Seven of the nine factors examined have statistically significant gross effects on the index of conformity: the amount of funds and the working relationship vis-à-vis research decisions between the evaluators and the sponsoring agency staff are the two exceptions. The net effects, however, are all very small.¹² This is probably due to the fact that no single factor has a large unique net effect on conformity. Rather, the variance is shared, and, as will be discussed later, patterns of

¹¹ Given the nature of our particular sample, it is not clear that significance tests are relevant. However, until the significance test controversy is resolved, we will opt to continue to present them.

¹² While we present the net effects to provide an estimate of the unique variance explained, it should be kept in mind that variables can be very important and can have nonsignificant net effects because their contribution to the explained variance is shared.

TABLE 1
Gross and Net Effects for Each of the Social
Factors on the Index of Conformity

Variable	Gross effects	Net effects
Sponsoring agency	.111***	.008
Nature of award	.165***	.032*
Amount of funds	.009	.009
Number of years for award	.150***	.055*
Type of organization with which researcher is affiliated	.077**	.011
Major audience for findings	.082***	.002
Formal organizational arrangement between education and action program organizations	.102***	.002
Working relationship between evaluation and action staffs re research decisions	.060**	.000
Working relationship between evaluation and sponsoring agency staff re research decisions	.035	.005

* Significant at .05.

** Significant at .01.

*** Significant at .001.

values occurring simultaneously seem to be more or less conducive to the kind of conformity being considered.

To examine the way in which the specific dummy variable categories, e.g., grant, affect the conformity index, we present the unstandardized regression coefficients (Table 2). Given our finding that only two of the independent variables have significant net effects, we are not surprised to find that few of the regression coefficients are statistically significant. While the fact that they are not statistically significant means only that the category which the dummy variable represents does not significantly differ from the left out category, we cannot reject the null hypothesis of no difference. Based on Table 2, we conclude that evaluators funded under contract auspices are less likely to conform to methodological prescriptions, as are evaluators whose time frame falls between 12 and 18 months.

The fact that the total variance explained is .301 and that seven of the nine independent variables have statistically significant gross effects, yet few variables have significant net effects (and few dummy variables have significant regression coefficients) leads us to reiterate our earlier conclusion that the variance explained is largely shared rather than unique. As such, we turn our attention toward a search for patterns of simultaneously occurring values that might have some predictive utility and theoretical meaning.

TABLE 2
Unstandardized Regression Coefficients, Standard Errors, *t* Values, and
Frequencies for Each Category of Each Independent Variable

Variables	Unstandardized regression coefficient (SE)	<i>t</i> Value	Frequency
Sponsoring agency ^a			
NIH/NIMH	+ .026 (.460)	.056	30% (45)
Housing and Urban Development, Agriculture, Labor, Office of Economic Opportunity	— ^b		24% (37)
Office of Education or HEW other than NIH, NIMH, SRS	— .034 (.408)	.084	9% (14)
Department of Justice	— .159 (.558)	.286	22% (34)
Social Rehabilitation Services	— .360 (.437)	.825	15% (22)
			100%
Nature of award			
Grant	—		59% (89)
Contract	— .986 (.403)	2.445	41% (63)
			100%
Amount of funds			
\$150,000 +	—		29% (42)
\$10,000–49,999	— .143 (.347)	.411	25% (36)
\$100,000–149,999	— .342 (.349)	.980	21% (31)
\$50,000–99,999	— .355 (.329)	1.079	25% (36)
			100%
Duration of award			
3 years +	—		29% (43)
Less than 1 year	— .051 (.442)	.116	11% (17)
1.5–2 years	— .275 (.361)	.762	19% (28)
2–3 years	— .349 (.314)	1.113	22% (33)
1–1.5 years	— 1.007 (.342)	2.949	20% (30)
			100%
Type of organization			
Educational institution or affiliate	+ .461 (.325)	1.420	32% (49)
Profit research corporation	+ .247 (.367)	.673	28% (42)
Nonprofit research corporation	+ .236 (.359)	.660	21% (31)

(Continued)

TABLE 2 (Continued)

Variables	Unstandardized regression coefficient (SE)	t Value	Frequency
Public service or planning agency	—		20% (30) 100%
Major audience ^c			
Cosmopolitan	+.198 (.281)	.705	30% (45)
Local	—		70% (106) 100%
Formal organizational arrangement			
Evaluator "inside"	+.204 (.330)	.618	41% (63)
Evaluator "outside"	—		59% (89) 100%
Working relationship between evaluator and action			
Action staff review decisions about research made by the evaluator	+.039 (.305)	.129	23% (34)
Evaluator and action staff make decisions about research jointly	+.036 (.305)	.117	44% (67)
Evaluator makes decisions about research independent of the action staff	—		33% (50) 100%
Working relationship between evaluation and sponsoring agency			
Sponsor makes research decisions independent of evaluator	+.258 (.502)	.515	5% (8)
Sponsors review decisions about research made by evaluator	+.158 (.272)	.581	38% (58)
Evaluator makes decisions about research independent of sponsor	—		40% (60)
Evaluator and sponsoring staff make decisions about research jointly	-.095 (.345)	.273	17% (26) 100%
Intercept = 4.538.			

^a Agencies were grouped together on the basis of similar operating practices and similar missions.

^b Left-out category.

^c Major audience was defined as "cosmopolitan" if the respondents indicated that they did or intend to maximize communication with other colleagues and researchers and to publish the results in journals and/or monographs. If the respondent indicated otherwise, e.g., communication would be maximized with government officials and results would be disseminated in government reports, the response for audience was classified as "local."

An examination of a correlation matrix for each of the response categories reveals two interesting patterns. The first is one of awards funded by research oriented agencies, e.g., NIMH, in the form of grants, for 3 or more years, to evaluators in educational institutions who define their major audience as cosmopolitan, who work as "insiders," and who make research decisions jointly with the staff of the action program (see Table 3).¹³ The second is one of awards funded by service-oriented agen-

TABLE 3
Correlation Matrix of Response Categories for Academic-Type Group^a

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
\$50-99 on funds (1) ^b	—							
NIH/NIMH sponsored (2)	21**	—						
Grant (3)	11	41***	—					
3+ years (4)	00	28***	26***	—				
Evaluator in educational institution (5)	03	32***	29***	17**	—			
Cosmopolitan audience (6)	11	37***	35***	35***	21**	—		
"Inside" evaluator (7)	11	40***	45***	26***	13*	26***	—	
Evaluator and action staff make joint research decisions (8)	12	28***	33***	21**	18**	16*	60***	—

^a The variable "working relationship between sponsor and evaluator" is not included because there is no clear pattern of association between any particular response on that variable and these other response categories. While the same is true of the variable "budget size," it is presented here as an item of possible interest.

^b Budget is in thousands.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

cies, e.g., SRS, in the form of contracts, for 12 to 18 months, to evaluators in profit research corporations who define their major audience as locals, who work "outside," and who make research decisions independent of the staff administering the action program (see Table 4). For convenience sake, we label the first group the "academic model," and the second, the "entrepreneurial model."

To determine whether a comparison of these two model types provides insight into our understanding of conformity, we computed a predicted

¹³ The amount of funds allocated for the research and the relationship between the sponsoring agency and the evaluation staff regarding research decisions are not included in the discussion since neither variable was found to have a significant gross effect on conformity.

TABLE 4
Correlation Matrix of Response Categories for Entrepreneur-Type Group^a

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
\$150+ funds (1) ^b	—							
Sponsor: HUD, AG, OEO, Labor (2)	35***	—	—					
Contract (3)	28***	60***	—					
1-1.5 years (4)	-14*	-06	15*	—				
Evaluator in profit research corporation (5)	24***	37***	47***	18**	—			
Local audience (6)	11	28***	35***	08	28***	—		
"Outside" evaluator (7)	12	28***	45***	12	36***	26***	—	
Evaluator makes research decisions independent of action staff (8)	10	15*	21**	04	20**	14*	39***	—

^a The variable "working relationship between sponsor and evaluator" is not included here for reasons cited in the note in Table 3.

^b Budget is in thousands.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

mean score for each pattern by adding the regression constant, i.e., 4.54, to the unstandardized regression coefficient for each response category (dummy variable) named above. The predicted mean for the "academic type" is 5.47; the predicted mean for the "entrepreneurial type" is 2.43. This difference suggests that these ideal-type patterns are correlated with conformity, and that the academic-type model is more conducive to the likelihood that there will be adherence to technical norms.

DISCUSSION

In order to explore why it is that evaluations more closely fitting the "academic type" are correlated with reported higher conformity to methodological prescriptions, we organize our discussion around a consideration of each of the categories included in that model type. Importantly, however, one must keep in mind the fact that few of these categories have unique effects on conformity.

Beginning with the *sponsoring agency*, from Table 1 we conclude that the sponsoring agency affects conformity. From Table 3 we note that being sponsored by NIH or NIMH is correlated with the "academic model." The question is, what is it about NIH/NIMH sponsorship that might explain its correlation with conformity to methodological norms. First, at the time of our study, NIH/NIMH was the only agency that made extensive use of an *external* peer review system. Second, by its heavy

reliance on the open grant system, agency decisions to support proposed research were purportedly made on the basis of projected conformity to technical norms and evidence of ability to adhere to technical prescriptions. Third, the heavy reliance on the open grant system may have created a highly competitive system whereby only proposals evidencing merit of support would be funded. In terms of the use of an external peer review system, Polanyi (1951, pp. 65–66) and Price (1968) argue that the conferring of research support must remain in the hands of representatives of the scientific community if professional standards are to be maintained. Their argument is predicated on the assumption that scientists alone are competent to judge the viability of proposed work. In terms of the relationship between competition and conformity, Simmel (1955, p. 60) and Hagstrom (1965, p. 39) note that competition is likely to increase commitment to norms and that the lack of it is likely to increase deviation.

The *nature of the award*, a variable highly correlated with the sponsoring agency, has a substantial gross effect on conformity and one of the only two significant net effects (see Table 1). From Table 2 we note that evaluators funded under contract auspices are less likely to conform to technical norms than are those funded under grant auspices. This finding may be a function of the fact that grants and contracts can be differentiated according to (a) the degree of participation of representatives of the scientific community in decisions about which research to support, (b) the degree of autonomy afforded the researcher by virtue of the type of award,¹⁴ and/or (c) the degree of competition for receipt of the award. Grants tend to be allocated as the result of peer review; they are traditionally awarded to researchers studying scientific problems identified and delimited *first* by the researcher, and competition for them is quite extensive. Contracts, at the time of our study, were generally awarded not after peer review, but rather on the basis of decisions by funding agency staff, usually persons who were not formally credentialed representatives of the scientific community. Furthermore, contracts were usually awarded on an RFP system; thus, the research problem and design for execution were defined and initiated first by the funding officer rather than by the researcher. Last, partly because of the particular demands of contracts, e.g., time specifications, competition for contracts is limited to the very few persons and groups capable and willing to work under the specific terms of the award.¹⁵

References for support of the finding that peer review in a competitive

¹⁴ By autonomy, we mean autonomy to define the research problem, delimit the scope of the problem, and select the methods by which to study the problem.

¹⁵ An additional difference between grants and contracts that we can assume to be present, but about which we have no evidence or data, is that persons may self-select themselves into those two categories on the basis of factors that may be related to our measure of conformity. For example, it may be that when a person has a clear research problem, with a good research plan, that person might apply for a grant: however, faced with

system is likely to increase conformity have already been cited. Regarding the relationship between autonomy and conformity, Polanyi (1951), Roe (1951), Whyte (1956), Kuhn (1962), Kornhauser (1962), Hagstrom (1965), Freidson (1970), and Zuckerman and Merton (1971) argue that autonomy is a primary incentive for conformity, and the lack of it a stimulus for deviation.

Before turning to the next variable, an additional point needs to be made, i.e., that our data suggest that the presence of a peer review system is correlated with greater conformity. Table 5 presents a mean breakdown of scores on the index of conformity by agency and by nature of award.

TABLE 5
Mean Scores of Evaluation Researchers Receiving Grant or Contract Support from Federal Agencies on the Index of Conformity^a

National Institutes of Health/Mental Health		Social Rehabilitation Services or Law Enforcement Assistance Administration		Office of Education, HUD, Agriculture, Labor, OEO, or HEW other than NIMH/NIH or SRS	
4.62 (45)		4.04 (48)		3.51 (58)	
Grant	Contract	Grant	Contract	Grant	Contract
4.86 (40)	2.70 (5)	4.08 (42)	3.69 (6)	4.79 (7)	3.33 (51)

^a NIH, NIMH, and Office of Education peer reviewed grants but not contracts in FY 1970. SRS, LEAA rarely used peer reviews on either grants or contracts in FY 1970. HUD, Agriculture, Labor, OEO, and HEW (other than SRS, NIH, NIMH) did not award any grants in FY 1970. Note: Overall mean is 4.0; SD is 1.4.

Although the cells are small, the pattern seems clear. In funding agencies in which grants and contracts differed according to the presence or absence of a peer review system, i.e., NIH/NIMH, persons receiving grants fared better on the index than did those receiving contracts. However, for agencies in which the grant and contract mechanisms did not differ, e.g., Social Rehabilitation Services, which at the time of our study made little or no use of peer review, regardless of whether the award was a grant or contract, we find no difference between persons supported under grant versus those supported under contract auspices. In view of the recent debates over peer review in *Science* (Gustafson, 1975), this seems important to note.

The *number of years* allocated for the research significantly and

a "murkier" research problem, and a research plan open to methodological criticism, that person might apply for a contract, hoping that the importance of the "murky" problem will outweigh the liabilities of the methodological inadequacies. Or, it may be that more methodologically sophisticated researchers apply for grants, while those with less methodological skills apply for contracts. In the absence of data, we can only raise these issues as potentially relevant to understanding our own findings as well as important considerations for future research.

uniquely affects conformity (see Table 1). From Table 3 we conclude that evaluative researchers working under minimum-time constraints (3 years or more) are more likely to conform to technical norms. According to Parsons (1951, p. 335) and Kornhauser (1962, pp. 52–53), scientists generally prefer work arrangements that allow for lengthy probes and the exhaustive testing of ideas.

The *type of organization* has a significant gross effect on conformity (see Table 1). From Table 3 we note that being affiliated with educational institutions is a characteristic of the "academic model." Thus, the question again is one of explaining what it is about educational institutions that might contribute to its correlation with conformity to methodological norms. Our contention is that educational institutions are most likely to define goals that are similar, if not identical, to the goals of the scientific community, e.g., publishing theoretically relevant, methodologically competent work. Moreover, in educational institutions, responsibility for work is usually allocated to particular individuals rather than to diffuse groups; and educational institutions are purportedly governed by colleague rather than hierarchical control. According to Parsons (1951), Orth (1959), Kornhauser (1962), and Hagstrom (1965), the university is the institution most likely to embody the norms of science. Furthermore, Hagstrom (1965) asserted that conformity to norms is most likely to occur when work responsibility is allocated specifically (to individuals) rather than diffusely (to groups), since only with specific responsibility can major rewards be obtained. Finally, Whyte (1965), Weinberg (1961), and Kornhauser (1962) suggest that conformity to scientific norms is more likely when the authority structure of the work organization is characterized by colleague rather than hierarchical control, since hierarchical control may lead to a redefinition of goals and the bases for rewards other than those ordinarily embraced by the scientific community, e.g., the goal of getting more research funds may replace the goal of conducting quality research.

From Table 1 we note that the *major audience* has a significant gross effect, and from Table 3 we see that defining one's major audience as cosmopolitan¹⁶ is part of the "academic model." For the group of evaluators (the cosmopolitans) who define other researchers and colleagues as their major audience and who indicate that they have or will publish their findings in refereed journals and/or monographs, it is likely that they conform to technical norms because their reference group and the mode of dissemination they use (or seek to use) require such conformity. However, for the other evaluators (the locals), who define the government and public officials as their major audience, and who publish (or plan to publish) their findings in government publications or mimeo-

¹⁶ For a definition of the difference between "cosmopolitan" and "local," see Footnote *b* on Table 2.

graphed organizational documents, their reference group and selected mode of dissemination do not require conformity to technical norms. As such, they are not motivated to do so or at least not by their reference group.

The *formal organizational arrangement* between the evaluative researcher and the organization administering the action program being evaluated has a significant gross effect on conformity (see Table 1). According to Table 3, having an organizational affiliation the same as the organization administering the action program, i.e., being "inside," is part of the academic model. (It is critical to note here that few "inside evaluators" are the same people as the administrators of the program being evaluated.)¹⁷ The finding that being "inside" is part of the model type more highly correlated with conformity is particularly interesting in view of the fact that a review of evaluative research literature reveals very mixed support for the "inside" position (see Caro, 1971, p. 17). Arguments made in favor of insiders hypothesize that insiders will be better equipped to develop knowledge about and measurements for program process and program effectiveness, better able to do continuing research, have greater access to important data records, and be better able to mediate problems ensuing from practitioner-scientist research relationships. Arguments made in favor of outsiders hypothesize that outsiders will be better able to maintain objectivity, more able to include criteria that question organizational premises and effectiveness, better protected from problems of marginality and status incongruity, and better able to withstand requests to engage in time-consuming nonresearch acts. We contend that the political liabilities of being outside may be so great so as to force evaluators to compromise scientific rigor in order to be able to execute some small part of the intended research. Argyris (1958), Angell (1967), Aronson and Sherwood (1967), and Coleman (1971) have all addressed the litany of complaints of evaluative researchers whose work has been compromised or thwarted by uncooperative social action staffs. Since this finding is somewhat unexpected, we will return to a further exploration of it shortly.

¹⁷ Bernstein and Freeman (1975) report that, of the 236 evaluation studies they reviewed, 38% were conducted by evaluators working within the same organization as that administering the social action program: 42% by evaluators working in partly or completely different organizations: 12% by evaluators working in different organizations where one organization was a subcontractor of the other: and 8% by evaluators who, along with the action staff administrators, were both separate subcontractors of a third organization. Of the 38% where the evaluation staff and action staff were part of the same organization, in only 25% of the 90 cases was the evaluation staff the same person(s) as the staff administering the action program. Thus, only 9% were simultaneously program administrators and evaluative researchers. An example of an "inside evaluator" is a psychiatrist on the staff of Harvard who acts as the evaluator for a mental health program run out of the Harvard Medical Center. While the evaluator and the program administrator are both members of the same organization, their immediate work groups are quite distinct.

The *working relationship between the evaluation staff and the action program staff vis-à-vis research decisions* has a significant gross effect (see Table 1). From Table 3 we note that joint decisions made by the two staffs are correlated with the academic model. Freeman and Sherwood (1965), Coleman (1971), and Kellam (1973) state that the politically sensitive nature of conducting evaluative research makes it essential for the two groups to work jointly in making research decisions. Extensive consultation with the action program staff can provide access to essential data about program process and program goals, as well as serve as a vehicle for establishing a cooperative relationship between the two staffs. They suggest, too, that this cooperative working arrangement is more crucial when the evaluator is an "outside" evaluator.

Because these last two variables are intercorrelated,¹⁸ and because the literature is so varied in terms of what is likely to correlate with conformity, we present Table 6 as an exploratory consideration of what might occur if we were to consider simultaneously the effects of these two variables on conformity.¹⁹ While the "inside" evaluative researcher fared better on conformity than the "outsider," when inside, it seems less

TABLE 6
Mean Scores of Evaluation Researchers in Varied Formal Organizational Arrangements with Varied Informal Working Arrangements on the Index of Conformity^a

Inside		Outside	
Evaluator and action staff in same organization		Evaluator and action staff in different organizations	
4.55 (62)		3.63 (89)	
Research decisions made inter-dependently with action staff	Research decisions reviewed by or made independent of action staff	Research decisions made inter-dependently with action staff	Research decisions reviewed by or made independent of action staff
4.55 (48)	4.56 (14)	4.01 (19)	3.52 (70)

^a Note: Overall mean is 4.0; SD is 1.4.

¹⁸ The correlation between being inside and working interdependently is .56. The correlation between being outside and working interdependently is -.60. The correlation between being outside and working independent of the action staff is .39. These three correlations are all significant at the .001 level.

¹⁹ We say exploratory consideration because the small number of cases does not allow for a meaningful examination of this possible interaction effect. In spite of this, we did do a difference of means test: the differences were not statistically significant.

important that the two make research decisions interdependently. However, when the evaluator was "outside," it seemed that an interdependent (joint) working relationship vis-à-vis research decisions increased the likelihood of greater conformity. Both these findings suggest that what might be important is that the two groups have *either* some formal or informal arrangement so that cooperation can be facilitated and mechanisms for gaining access to the necessary information can be provided. Summarily, while the data don't show a statistically significant interaction effect, such an effect might be demonstrable if one had a larger sample of cases.

CONCLUSION

In terms of Merton's (1949) original question about the factors that make for achievement in applied social science, our data lead us to conclude that, (a) When the representatives of the scientific community are active participants in the decision-making process to award support to persons whose research proposal shows evidence of promise and competence, and (b) when the award is made to persons presently members of educational institutions *or* the like, where there is likely support for scientific work that evidences methodological competence, and (c) when the researchers in accordance with the reward structure of the organization with which they are affiliated, define their major audience as including other members of the scientific community, and (d) when the researcher works cooperatively with the host agency in executing the research, there will be a greater likelihood of achievement in applied social science. Our assumption is that a prerequisite for achievement is adherence to methodological norms.

In terms of the Hagstrom question about external influences and their effect on applied scientists, our data suggest that not only are external factors affecting scientists' behavior, at least the degree to which they adhere to technical norms, but also, our comparison of the academic versus entrepreneurial types suggests that as the input of nonscientists increases, so does the likelihood of their influence having a negative impact on conformity to technical norms increase.

In terms of the implications for social policy, our exploratory findings raise two immediate questions. First, given that the input of nonscientists is increasing, what are the conditions under which one can translate that input into an effect positively correlated with adherence to methodological norms. Second, given that the entrepreneurial model is likely to dominate in applied social science research, what mechanisms of social control can be introduced to increase conformity to methodological norms within the entrepreneurial model. Both these questions make obvious the fact that further research is needed before we can begin to delimit a model for structuring the organization of applied research conducive to the production of quality work.

REFERENCES

- Alwin, D., and Sullivan, M. (1975), "Issues of design and analysis in evaluation research," *Sociological Methods and Research* 4, 77-100.
- Anderson, A. B. (1975), "Policy experiments: Selected analytic issues," *Sociological Methods and Research* 4, 13-30.
- Angell, R. C. (1967), "The ethical problems of applied sociology," *in The Uses of Sociology* (P. Lazarsfeld, W. Sewell, and H. Wilensky, Eds.), pp. 725-740, Basic Books, New York.
- Argyris, C. (1958), "Creating effective research relationships in organizations," *Human Organizations* 17, 34-40.
- Aronson, S., and Sherwood, C. (1967), "Researcher versus practitioner," *Social Work* 12, 89-96.
- Barber, B. (1952), *Science and the Social Order*, Free Press, Glencoe, Illinois.
- Bernal, J. D. (1939), *The Social Function of Science*, Routledge and Kegan Paul, London.
- Bernstein, I. N. (1975), *Validity Issues in Evaluative Research*, Sage Publications, Beverly Hills, California.
- Bernstein, I. N., Bohrnstedt, G. W., and Borgatta, E. F. (1975), "External validity and evaluation research: A codification of problems," *Sociological Methods and Research* 4, 101-128.
- Bernstein, I. N., and Freeman, H. E. (1975), *Academic and Entrepreneurial Research: The Consequences of Diversity in Federal Evaluation Studies*, Russell Sage Foundation, New York.
- Bernstein, I. N., and Sheldon, E. B. (1977), "Methods of evaluative research," *in Social Science Methodology* (Robert Smith, Ed.), Free Press, New York.
- Blume, S. (1974), *Toward a Political Sociology of Science*, Free Press, New York.
- Boalt, G. (1969), *The Sociology of Research*, Southern Illinois University Press, Carbondale, Illinois.
- Bohrnstedt, G. W. (1969), "A quick method for determining the reliability and validity of multiple item scales," *American Sociological Review* 34, 542-548.
- Boruch, R. F. (1974), "Bibliography: Illustrative randomized experiments for planning and evaluating social programs," *Evaluation* 2, 83-87.
- Cain, G. and Hollister, R. (1972), "Methodology of evaluating social action programs," *in Evaluating Social Programs, Theory, Practice and Politics* (P. Rossi and W. Williams, Eds.), pp. 109-137, Seminar Press, New York.
- Campbell, D. T. (1971), "Reforms as experiments," *Readings in Evaluative Research* (F. Caro, Ed.), pp. 233-261, Russell Sage Foundation, New York.
- Caro, F. (Ed.) (1971), *Readings in Evaluative Research*, Russell Sage Foundation, New York.
- Cohen, J. (1968), "Multiple regression as a general data analytic system," *Psychological Bulletin* 70, 426-443.
- Cole, J., and Cole, S. (1973), *Social Stratification in Science*, University of Chicago Press, Chicago.
- Coleman, J. S. (1971), "Problems of conceptualization and measurement in studying policy impact," *Social Science Research Council Conference*, Virgin Islands.
- Collins, R. (1968), "Competition and social control in science: An essay in theory construction," *Sociology of Education* 41, 123-140.
- Cook, T. D., and Campbell, D. T. (1975), "The design and conduct of quasi-experiments and true experiments in field settings," *in Handbook of Industrial and Organizational Research* (M. D. Dunnette, Ed.), pp. 223-236, Rand McNally, Chicago.
- Eber, H. (1975), "Multivariate methodologies for evaluation research," *in Handbook of Evaluation Research* (E. Struening and M. Guttentag, Eds.), pp. 553-570, Sage Publications, Beverly Hills, California.

- Edwards, W., Guttentag, M. and Snapper, K. (1975), "A decision-theoretic approach to evaluation research," in *Handbook of Evaluation Research* (E. Struening and M. Guttentag, Eds.), pp. 139-181, Sage Publications, Beverly Hills, California.
- Freeman, H. E., and Sherwood, C. C. (1965), "Research in large scale intervention programs," *The Journal of Social Issues* 21, 11-28.
- Freidson, E. (1970), *Profession of Medicine: A Study of the Sociology of Applied Knowledge*, Dodd, Mead, New York.
- Gustafson, T. (1975), "The controversy over peer review," *Science* 190, 1060-1066.
- Hagstrom, W. O. (1965), *The Scientific Community*, Basic Books, New York.
- Hyman, H., and Wright, C. (1967), "Evaluating social action programs," in *The Uses of Sociology* (P. Lazarsfeld, W. Sewall, and H. Wilensky, Eds.), pp. 741-783, Basic Books, New York.
- Kellam, S. (1973), *Discussant Presentation*, American Sociological Association Seminar, New Orleans.
- Kerlinger, F. N., and Pedhazur, E. J. (1973), *Multiple Regression in Behavioral Research*, Holt, Rinehart and Winston, New York.
- Kornhauser, W. (1962), *Scientists in Industry: Conflict and Accommodation*, University of California Press, Berkeley, California.
- Kuhn, T. (1962), *The Structure of Scientific Revolution*, University of Chicago Press, Chicago.
- Marcson, S. (1966), *Scientists in Government: Some Organizational Determinants of Manpower Utilization in a Government Laboratory*, Rutgers University Press, New Brunswick, New Jersey.
- Merton, R. K. (1942), "Science and technology in a democratic order," *Journal of Law and Political Sociology* 1, 115-126.
- Merton, R. K. (1949), "The role of applied social science in the formation of policy," *Philosophy of Science* 16, 161-181.
- Mitroff, I. (1974), "Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists," *American Sociological Review* 39, 579-595.
- National Science Foundation (1967), *Basic Research, Applied Research, and Development in Industry*, p. 101, NSF 67-12, Washington, D.C.
- Nunnally, J. (1975), "The study of change in evaluation research: Principles concerning measurement, experimental design, and analysis," in *Handbook of Evaluation Research* (E. Struening and M. Guttentag, Eds.), pp. 101-137, Sage Publications, Beverly Hills, California.
- Orth, C. D. (1959), "The optimum climate for industrial research," *Harvard Business Review* 37, 55-64.
- Parsons, T. (1951), "The institutionalization of scientific investigation," *The Social System*, Free Press, Glencoe, Illinois.
- Polanyi, M. (1951), *The Logic of Liberty*, Routledge and Kegan Paul, London.
- Price, D. K. (1968), *Presidential Address to the American Association for the Advancement of Science*.
- Riecken, H., and Boruch, R. (1974), *Social Experiments: A Method for Planning and Evaluating Social Programs*, Seminar Press, New York.
- Roe, A. (1951), "A psychological study of eminent biologists," *Genetic Psychology Monographs* 43, 121-239.
- Rossi, P. (1972), "Testing for success and failure in social action," in *Evaluating Social Programs* (P. Rossi and W. Williams, Eds.), pp. 11-49, Seminar Press, New York.
- Simmel, G. (1955), *Conflict and the Web of Group Affiliations*, Free Press, New York.
- Stanley, J. C. (1972), "Controlled field experiments as a model for evaluation," in *Evaluating Social Programs* (P. Rossi and W. Williams, Eds.), pp. 67-72, Seminar Press, New York.

- Storer, N. (1966), *The Social System of Science*, Holt, Rinehart and Winston, New York.
- Stromsdorfer, E. (1972), *Review and Synthesis of Cost-Effectiveness Studies of Vocational and Technical Education*. ERIC Clearinghouse on Vocational and Technical Education, Columbus, Ohio.
- Suchman, E. (1967), *Evaluative Research*, Russell Sage Foundation, New York.
- Weinberg, A. (1961), "Impact of large scale science on the United States," *Science* **134**, 161-164.
- Weiss, R. S., and Rein, M. (1970), "The evaluation of broad-aim programs: Experimental design, its difficulties, and an alternative," *Administrative Science Quarterly* **15**, 97-109.
- Wholey, J., Scanlon, J., Duffy, H., Fukumoto, J., and Vogt, L. (1970), *Federal Evaluation Policy*, The Urban Institute, Washington, D.C.
- Whyte, W. H. (1956), *The Organization Man*, Doubleday, Garden City, New York.
- Whyte, W., and Hamilton, E. (1964), *Action Research for Management*, Dorsey, Homewood, Illinois.
- Wilkins, L. T. (1969), *Evaluation of Penal Measures*, Random House, New York.
- Williams, W., and Evans, J. (1969), "The politics of evaluation: The case of head start," *Annals of the American Academy of Political and Social Science* **385**, 118-132.
- Zuckerman, H. and Merton, R. K. (1971), "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System." *Minerva* **9**, 66-100.