

Evaluation as Meaningful Social Research

Franz-Xaver Kaufmann
Klaus-Peter Strohmeier
Universität Bielefeld

PROGRAM EVALUATION AS “APPLIED BASIC RESEARCH”

Scientific program evaluation research must be seen in the focus of two different, but not necessarily contradictory, perspectives: On one hand, in as much as evaluation attempts to discover causal empirical relationships, it falls under the standards of a logic of (causal) scientific inquiry. Accordingly, strong emphasis must be placed on the complexity, completeness, and causal closure of a system of theoretical hypotheses from which the direction and the nature of a causal process under study can be derived.

It is more or less accepted, according to the logic of scientific inquiry, that an empirical statement about something (as a cause) having affected something else (as an effect) can only be accepted insofar as it confirms an underlying theory (Stegmüller, 1960; Popper, 1964; Opp, 1976; Simon, 1970; Luhmann, 1970). Treating an empirically confirmed relationship between variables as causal will be justified only by the underlying theoretical propositions, not just by the kind of research operations applied.

On the other hand, program evaluation research is applied social research, which is intended for application outside the scientific community.¹ In that perspective it differs from general or basic social research. Its underlying theoretical presuppositions are never “grand” social theory, but

theories of the short- or middle-range type. And the "causes" and "effects" are not just "empirical realizations of generally conceptualized social phenomena" (Weber, 1964), but political or policy-induced variables (Kaufmann et al. 1978). As applied social research, program evaluation is not only committed to high technical standards; but also, to be incorporated into political processes, its results have to be precise and communicable. These two perspectives, both the more practical and the methodological, must be kept in mind in examining the applicability and adequacy of different research designs for the handling of a specific evaluation research problem. Neglecting either of them will lead to methodologically and practically irrelevant research results.

This chapter discusses the relationship between theory and method in program evaluation research; later it will describe two standards, "object adequacy" and "problem adequacy" of a research design. These might help to ensure that program evaluation research be theoretically relevant and practically (politically) useful social research.

PRINCIPLES OF EXPERIMENTAL EVALUATION

The expanding literature about methodological issues in evaluation research documents the predominant type of theory about how and why social programs come to effect as the almost classical "impact model" of the following type (Hellstern and Wollmann, 1977):

a PROGRAM $\xrightarrow{\text{initiates}}$ a CAUSAL PROCESS $\xrightarrow{\text{that produces}}$ EFFECTS.

Factually, this impact model expresses a very simple causal hypothesis, assuming that the program and nothing else will bring about the desired effects, and that, controlling for intervening factors, a direct causal relationship between the program and the dependent (target) variable will remain. Distinguished authors, such as Suchman (1972), Weiss (1972), Campbell (see Campbell and Stanley, 1966), and Scheuch (Scheuch and Rüschemeyer, 1965) have assigned the experimental research technique the highest degree of validity and precision. The simple impact model usually is the grounded theory that underlies the "real experiment" as a technical device in program evaluation research.

"Real" experiments allegedly come nearest to the exact technical devices of the experimental natural (or physical) sciences (Suchman, 1972; Scheuch and Rüschemeyer 1965). There are firm reasons, which, at least, question that methodological position which has so far been prominent among evaluation researchers (Lazarsfeld, 1965). Before expressing the

problems of a strictly experimental technical approach, we want to summarize the basic principles of experimental program evaluation.

In the natural sciences, experimental research can be characterized as the creation of conditions that allow initiating a succession of manipulating changes in the object under study in an ideally innumerable number of cases, practically undisturbed by irritations from external factors. Consequently, applying experimental techniques assumes that it is possible to reproduce theoretically closed systems for practical research purposes.

The "real" experiment in evaluation research appears at the first glance weaker and, on the whole, not similar to the allegedly ideal natural science experiment. Actually, it is completely different.

- Randomly selected units (persons) from a program's target population are assigned to a control group and to at least one experimental or "treatment" group. Random assignment (or other devices, such as "matching") allows the exclusion of systematic bias from the composition of groups, ensuring that potentially intervening variables will either be randomly distributed or can otherwise be regarded as constant.
- Control and experimental groups are then submitted to a first measurement (before-treatment) in which their values on the target variable which are expected to be influenced by the program are measured.
- The experimental group(s) is (are) submitted to a treatment (or, in the case of several groups, to different treatments), whereas the control group gets no treatment.
- At the end of the program (and sometimes in between) additional measurements are taken to identify changes in the dependent (or target) variable. If significant changes from earlier measurements can be observed and if significant differences between groups can be identified, the program and the different modes of treatment under study can be said to have brought about the observed changes (Suchman, 1972; Jones and Borgatta, 1972; Breedlove, 1972).

"OBJECT ADEQUACY" OF EMPIRICAL DESIGNS

Differences between the evaluative experimental technique and the allegedly ideal experimental model of the natural sciences are obvious and self-evident. There is no point in discussing how the social science use of experimental designs can be improved in order to reach natural science standards of validity and precision; this would be a vain effort and, moreover, according to the object of empirical social research, an inadequate attempt. The "social world" as the object of social scientific inquiry is different from the concerns of the physical sciences, and inadequate research standards would not help us (Lazarsfeld, 1965). It is this fundamental gap between the object of research and a growing "complex of exactness" among sociological re-

searchers that made König (1968) call empirical social scientists to a higher consciousness of the "object adequacy" ("*Gegenstandsadäquanzt*") of their research operations.

This argument has also been stressed by interactionist methodologists who, however, equate "object adequate" with "nonquantitative" (Blumer, 1973; Wilson, 1973; Matthes and Schütze, 1973). Tentatively, one could well use these two positions—the one with the "complex of exactness" (König, 1968) on one side and the anti-quantitative entirely "qualitative" approach on the other—to group both general social science and evaluation research into two discrete categories. Nonetheless, neither of them can factually claim to be more "object adequate" than the other. "Object adequacy" is not achieved simply by choosing between two paradigmatic branches of empirical social research which are too often regarded as mutually exclusive. A presupposition that social research be either "qualitative" or "quantitative"—either "valid" and "exact" or "invalid" and "weak"—erects false barriers that hinder the progress of scientific research (and its practical uses are part of that progress).

It is true that the concerns of social scientific inquiry, the "social world" (Blumer, 1973), differ from those of the physical sciences; therefore, we need not bother as social scientists to reach factually inadequate standards of exactness. However, this cannot mean that access to any phenomenon in the social world can be gained only by working on it qualitatively—for example, by means of direct observation, contextual analysis, or narrative interviewing (for the latter see Schütze, 1976). Moreover, we would extend the criterion that research operations be "object adequate" also to the problem of deciding whether to use a qualitative or a quantitative approach (or both) in a given research situation.

In a very broad sense, all social scientific research devices, "weak" or "exact," "qualitative" or "quantitative," can be called "object adequate" insofar as they allow a researcher to deal with a phenomenon within his or her social world of research objects. The specific character of the object under study additionally imposes advantages and restraints that make one technical procedure of research appear more adequate than another. A direct experiment as a special research technique will be adequate only when the causal process investigated is as simple as the "impact model" as a basic causal theory suggests. Other, more complex social processes may well demand the use of technically nonexperimental, even nonquantitative, research designs or a mixture of techniques with varying degrees of rigor. Program evaluation research can be regarded as a specific and fruitful challenge to social science researchers. Its specific concerns do not easily fit into the (mutually conceded) "claims" of either qualitatively or quantitatively orientated researchers, but, instead, are far less conventional.

An evaluation researcher asked to identify the implementation process of a specific program is well advised to use qualitative research devices, such as documentary analysis or interviewing of key persons involved in the implementation. When evaluating the impact of a program, the approach depends entirely upon the type of program studied and the theoretically assumed relationships of variables and their complexity (the impact model). In this case, the researcher might use an experimental or quasi-experimental design to identify program effects (Hellstern and Wollmann, 1977). Only recently, evaluation researchers have conceded that qualitative approaches are adequate and informative procedures for empirical research (Campbell, 1974). Hellstern and Wollman (1977) also emphasize the use of applying multiple research techniques of both qualitative and quantitative approaches to a given research problem.

It is still an indicator for a widespread, but nonetheless erroneous, methodological orientation that, where real experiments are factually impossible, alternative research designs are labeled "quasi"-designs. However, the value of a research technique, in our view, does not lie in its similarity to an allegedly ideal experimental technique. König (1968) has made the point that all analytic scientific inquiry is experimental, in that it follows a basic method of "experimentation." The principle is to keep factors constant while varying others to see the results. We call this the principle of "controlled variation" (Smelser, 1967). All sorts of analytical research techniques, including real experiments and even a *Gedankenexperiment* in the Weberian sense, are under certain conditions distinguished technical realizations of this basic methodological principle.

The selection of the empirical approach will always depend upon the character of the research object and the theory that is to be examined empirically.

"PROBLEM ADEQUACY"

Our understanding of "object adequacy," developed in the preceding section, comprises methodological standards that are applicable to all empirical social research. In the special case of program evaluation, however, it would suggest that there is no such thing as an absolute ranking of designs, going "downwards" from the direct experimental to other, perceptibly weaker, devices in terms of validity, "hardness" of data, and results. Instead, "experimental" in the sense of "controlled variation" is a basic criterion distinguishing analytical scientific research from other (for example, descriptive), fact-collecting types of research. The various research techniques existing in the social sciences are all technical realizations of this basic experimental method—each of them is adequate, provided it is applied to a specific type of objects. Object-adequate research designs alone will not,

however, make an evaluation "problem-adequate" as to the political problem of concern.

Donald Campbell (1974) considers "evaluating the outcomes of deliberately induced political innovation" as the main objective of program evaluation research. This definition may hold for the United States; however, it is problematic when applied to Germany. Evaluation research as a specific mode of interaction between the social sciences and the political system (Kaufmann, 1980) has been imported into the Federal Republic from the United States. Consequently, the methodological and theoretical affiliations that most West German evaluation researchers adhere to have also been imported. Nevertheless, there remain crucial differences in the structure of the political systems and their modes of initiating, implementing, and improving policy. Our political system is by no means experimental, neither in the sense that political innovation would play an important role in state activities nor in that it would be implemented in a way that followed the logic of scientific experimentation (Kaufmann and Schneider, 1975). Experimental reforms as well as established policies in the Federal Republic will be implemented, delivered, and administered grounded on legal regulation. Our political process is basically regulated by law and only secondarily by attainment of goals. The administration through which a policy program is implemented commits itself to a principle of "legality" (*Rechtmäßigkeit*); that is, the compatibility of an individual administrative action with general legally codified norms. This particular structure and function of the German political process, (policy being nonexperimental and policy regulated by general law), pose particular problems to evaluation research that we will discuss in this section.

The primacy of "legality" (*Rechtmäßigkeit*) in German policy can be most consequential for empirical research on the impacts of policy. Legal regulation determines which target population under specified conditions and in specific modes use a program (whether established or innovative) and receive associated benefits. No one who fits those legally defined conditions can be excluded from using a program, for example, by being assigned to a control group. Conversely, nobody who has the right to use a program can be urged to submit himself to a treatment.

In a situation in which a potential client submits himself to a "treatment" in a social program, the way in which (and the conditions under which) he does so necessarily become a subject for evaluation. Unlike a physical scientist, who creates an experimental setting and can observe a successive relationship between changes in specific variables, the evaluator of a social program is essentially unable to assume causal closure with an empirical research design. Human beings (or social units in general), can "allow" themselves to be effected by a treatment. One can say that the disposition or

motivation of clients to use a program is an essential prerequisite for it to come to effect. It is here more appropriate to consider "causal" relationships not in a deterministic sense but as contingent.

That basic motivation of a client to participate in a program, however, is not just a matter of entirely voluntary individual decision. It has systematic and objective bases in the client's life situation and personality. An evaluative research design will then be problem adequate when it allows reconstruction of conditions for those meaningful social processes of self-selection among potential clients and determines their influence in a program.

With established programs in the Federal Republic a purely experimental approach is not possible. Even with reforms where one might employ experimental evaluation techniques, one would be unable to identify the complex sets of influences in the social settings of the program and their impact on effectiveness. The experimental design does, in fact, allow an adequate empirical reconstruction of the simple impact model mentioned earlier. However, it is the impact model that provides an oversimplified view of reality, especially when nonexperimental, legally regulated policy is evaluated.

(It is true that results of experimental research are easily communicable. However, their information value for political purposes is relatively limited. In fact, direct experiments would be able to demonstrate only how programs work without the influence of the social reality in which they operate. The interest of policy makers in program evaluation goes further (at least, it ought to) than just wishing to know if a program works or not; normally, it also includes potential ranges of program effects, the conditions under which programs are used, and the utilization by the target population.

Initially we remarked that causal relationships identified through scientific investigation are confirmations of a theory from which the nature of the causal process and the causal character of the observed relationships will become plausible. In evaluation research, this theory describes a social or political problem addressed by a program. In this sense, problem adequacy is not only a property of the empirical research operations, in as much as they allow identification of the influences of the social settings in which programs normally will work. In fact, problem adequacy also characterizes the problem-solving or policy-improving capacity of an evaluation study.

Object-adequate and problem-adequate evaluation starts from a theoretical "impact model" that is sufficiently differentiated and informative to conceptualize cause-effect relationships within a heterogeneous social field. Finally, it will conclude with a theory, empirically confirmed, that informs both the politician and the administrator how their social programs work and which effects they have under specified circumstances. Insofar as meaning-

ful evaluation research is committed to general standards of methodology, it will fulfill the professional norms of quality, such as a sufficient amount and quality of theoretical conceptualizations, as well as object- and problem-adequate research procedures. The more comprehensive the underlying impact model and the more object-adequate and problem-adequate the empirical investigations, the less is the probability that evaluation results, as empirically grounded theories, will vanish in a drawer (Hellstern and Wollmann, 1978; Blankertz, 1976) somewhere in a government agency and be used only for the purpose of political legitimization. The latter, however, is most likely to happen when evaluation limits itself to merely delivering data, no matter how exact, instead of theories or models of reality. Evaluation research as applied basic research performs a scientific reconstruction of the conditions through which and the modes in which social programs operate.

Initially we said everyone involved in a program will do his personal subjective evaluation and, in terms of practical theories or definitions of the situation, will have a conception about how and why "his" program works. Improving policy through evaluation serves to correct those everyday conceptions of reality where they need to be corrected. Evidently, in that respect, social scientific evaluation will only then gain a policy-improving function or become "professionally effective" (Breedlove 1972) when it is able to confront those practical, everyday interpretations of reality with an empirically grounded scientific reconstruction of the way in which social programs work.

EVALUATING ESTABLISHED PROGRAMS

An eventual policy-improving function of an evaluation study will depend upon the quality of its grounded theory or "impact model"; that is, the way in which it conceptualizes relationships in the social field in which policy programs operate. We will exemplify the need for a complex impact model and for "object-" and "problem-" adequate research strategies by introducing the basic theoretical propositions and a few selected results taken from an evaluation study we conducted for the Federal Ministry of Youth, Family and Health Affairs (Bundesministerium für Jugend, Familie und Gesundheit) between 1974 and 1978 (Kaufmann et al., 1978, 1980).

The policy question appeared very simple at first glance: "What is the impact of the State's established policy programs² for families with young children of preschool-age upon the socialization process in the family?" The programs or policies considered comprise the following heterogeneous activities:

- (a) financial benefits, such as
 - monetary family allowances (*Kindergeld*)
 - monetary housing allowances (*Wohngeld*)
 - monetary subsidies that grant a minimum living to the very poor (*Sozialhilfe*)
- (b) social services, such as
 - nursery schools (*Kindergarten*)
 - adult education for parents (*Elternbildung*)
 - advisory services (*Erziehungsberatung*)
 - health services (*Mütterberatung*).

Literal equivalents in the United States are difficult, but the common factor is that all policies potentially influence the life situation of children. They are all established programs, some of them for decades. For an evaluation researcher the policy question appears unanswerable in its original formulation. Instead, we had to translate the simple political problem into a workable scientific research problem:

(a) What do we have to evaluate? Rather than an innovation, or something that has not existed before, we must evaluate policies that to a great extent have shaped the everyday world of the family in our country. Achinger (1958) stated that the social world, the reality in which we all live, has broadly been effected and is continually influenced by social policy. Everyone knows what a “Kindergarten” is, although they may have never attended one.

(b) How do we measure the “impact” of social policy as a part of social reality upon the socialization process? The “programs” we studied have no explicit and operative “program goals” (which a less rigorous evaluation might adopt as dependent or target variables (for educational programs, see Blankertz, 1976).

As social scientists we asked ourselves, “What effects can be expected if the programs were to improve the life situation and developmental conditions of young children in the family?” In accordance with recent trends in socialization theory, and sociologically elaborating Bronfenbrenner’s “Ecology of Child Development” (1974), which emphasizes the predominant influence of children’s enduring environment upon developmental processes, we have identified parents’ communicative and regulative activities (their “performance” in their social role as parents) as the key variable of the socialization process (see, for details, Kaufmann et al., 1979, 1980; Herlth and Strohmeier, 1980). Parental activities decide the quality of a child’s home environment, and the modes in which a child gains access to the outside social and physical world are also a result of parents’ regulative activities. Thus, the dependent variables in our study were not questionable

measures for the problem at hand, as children's IQs for example, but specific modes of parental competence and behavior.

(c) How can the (causal) links between the policy programs and parental performance be conceptualized in an impact model? Primarily, it cannot be taken for granted that there are any causal links. Simple impact models only allow the conclusion that a program has or has not worked. Our interest, however, was directed at the "causal processes" that explain the observed "effects" or "non-effects." We have developed a theoretical framework to represent the "impact model" that, taking into account the actual research literature, conceptualizes the "environment dependency of the family socialization process."

The basic theoretical proposition is that families need "social resources," such as knowledge on socialization (see Lüscher, 1977), goods, money, housing, and social services (medical care and day care), to bring up children in qualitatively sufficient ways (according to a child's social chances). Some of those "social resources" are "internal resources"; that is, families "have" them by virtue of parents' education or occupation, the structure of the family, and the personality of its members. Others, however, are "external resources," such as goods which families have to gain from the "socioecological context" of their environment. External resources are transferred into the family via temporary interaction of its members within other social systems, such as the employment sector, social networks, and social agencies and social services. Thus, the social policies being evaluated and the agencies providing the services must be treated as environmental systems which offer special kinds of external resources—money, advice, child care, parent education, and so on. As such, the use or nonuse of specific services by a family represents a distinctive familial property—"selectivity" of participation—in specific social systems in the socioecological context. This theoretical assumption also considers the special characteristic of "programs" (they are conditional and they are not more than just offers) that can be but do not have to be accepted.

A theoretical impact model such as ours has two main functions. It can explain under which conditions people actually use programs designed for them and the extent of the utilization. On the other hand, it can identify factors influencing the dependent variables apart from the program(s) evaluated. Furthermore, it clarifies additional conditions under which programs work and in what ways self-selection of clients explains eventual program effects measured. This type of impact model performs the analytical reconstruction of a concrete social field in which policies realistically work. It contains variables on different analytical levels. We have thus decided to use a multi-level and multiple-research-methods design. The relevant variables needed were measured on an individual (family), an aggregate (socioecolo-

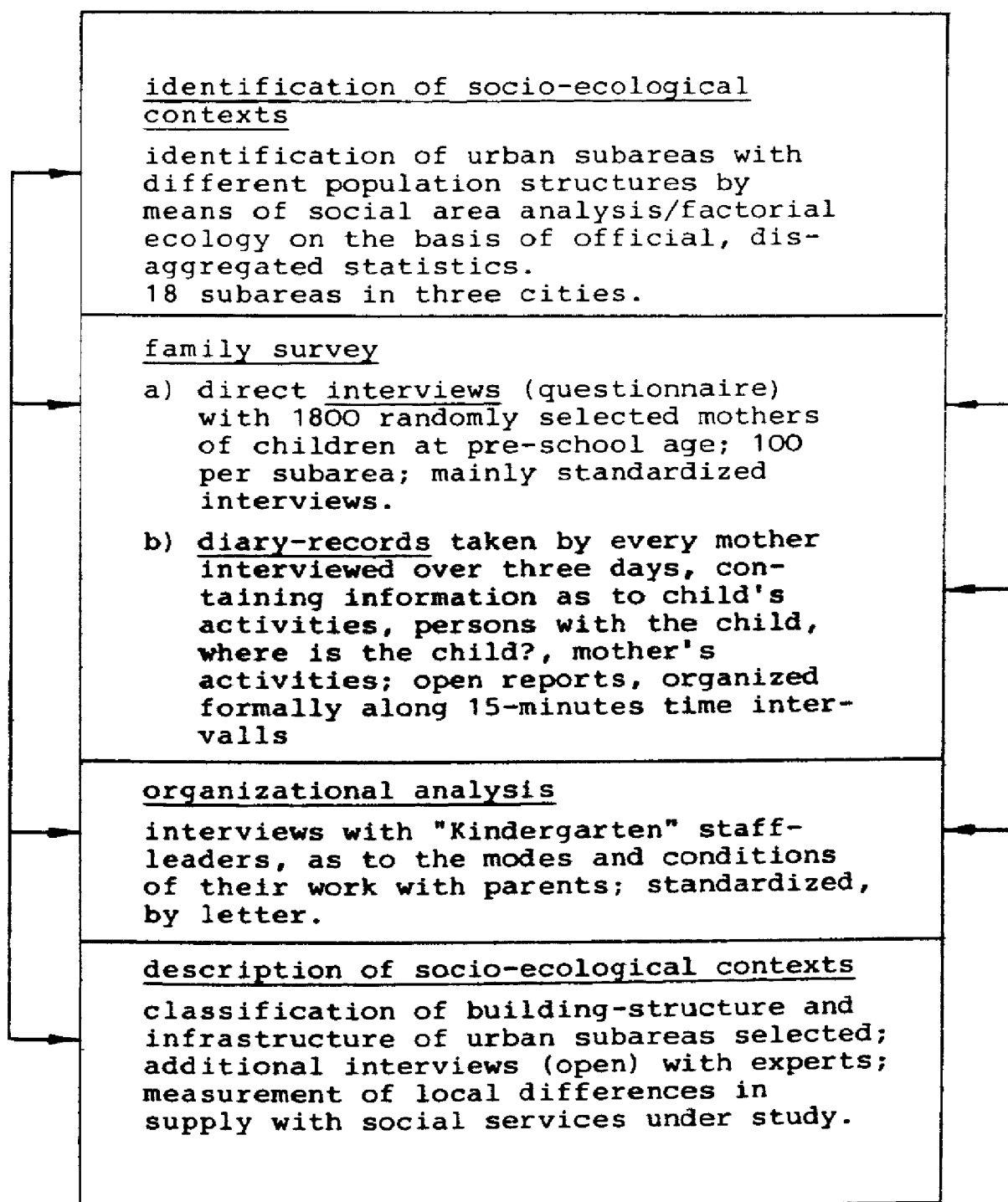


FIGURE 8.1 Analytical Levels of Research and Research Instruments

gical context), and an organizational level—the kindergarten.

The scheme presented in Figure 8.1 compiles the different levels of research and the research techniques, qualitative and quantitative, applied. “Bows” denominate research procedures with cross-validating function; that is, variables were redundantly measured in different procedures.

Finally, we would like to demonstrate the influence variables in the structure and life situation of families have upon clients selectively submitting themselves to different modes of political treatment, and the ways in which the same variables determine the impact of a treatment. We use data about the use and impact of nursery-schools (kindergarten). The nature of our dependent variable, competence and performance of parents in their role as parents, suggests that we concentrate upon what a kindergarten does with parents, how far it affects the "parent-child-system" (Bronfenbrenner, 1974). Depending upon the degree to which parents are involved in the activities of the kindergarten (that they can factually participate is assured by legal provisions), we can distinguish four different groups that receive different "treatments."

- B = Parents who do not and will not send their children to a kindergarten. This group will serve as a control group with no treatment.
- A1 = Parents whose children attend a kindergarten but who themselves do not participate in parents' activities (formal assemblies, pedagogic discussion, activities with parents and children, and the like).
- A2 = Parents who send their children to a kindergarten and irregularly take part in parents' activities.
- A3 = The group with the most intense treatment: the child attends a kindergarten and the parents regularly participate in kindergarten activities.

Actually, we can expect differences in the composition of groups B, A1, A2, and A3. Figure 8.2 demonstrates these compositional differences in a profile of subgroups' mean scores on life situation and family structure indicators. The only similarity of the control group to the treatment groups (A1 to A3) is the age of their preschool children. On all other variables it has a markedly different profile. Nonuser families have the lowest socioeconomic status; a low percentage of mothers are working; families live under the worst economic and housing conditions and have relatively many children. On the other hand, a growing intensity of "treatment" is accompanied by increasing values on the status indicators and in the financial and housing situation. In other words, we see that it is mainly the social status of families (as an indicator of specific types of life situations; (see Strohmeier and Herlth, 1979) through which the different modes of participation can be explained.

We have chosen four indicators to demonstrate the effects of those different modes of parents' participation in the kindergarten upon their behavior as parents:

- the number of *children's books* available in the family;
- frequency of *children's playmates* allowed to visit the family home;
- frequency of *intensive mother-child interaction*—that is mother spending more than half an hour in intensive communication (play, reading a story, and so on) with her preschool child; and

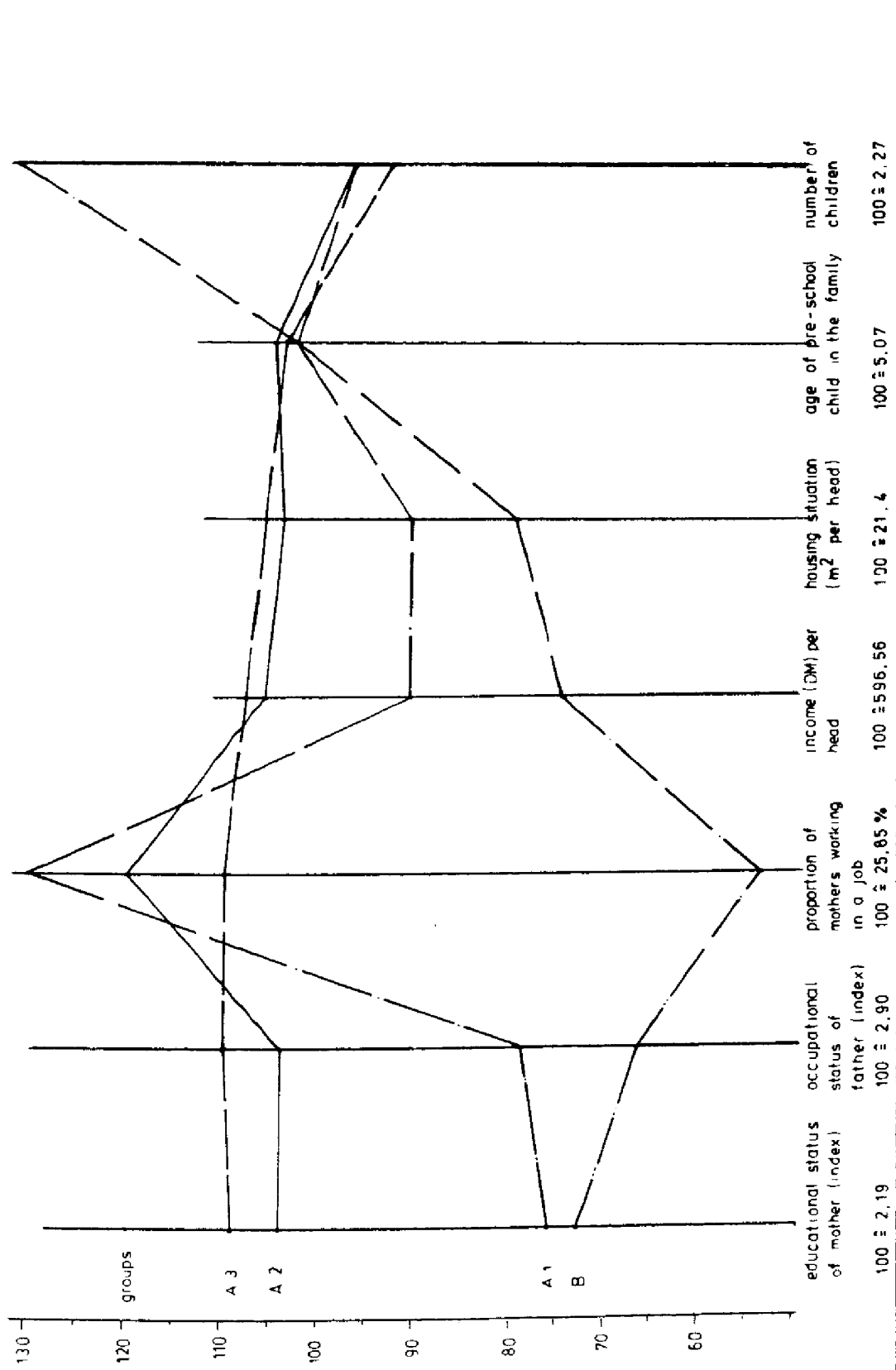


FIGURE 8.2: Parental participation in nursery-schools ("Kindergarten") on the background of families' structural and situational properties

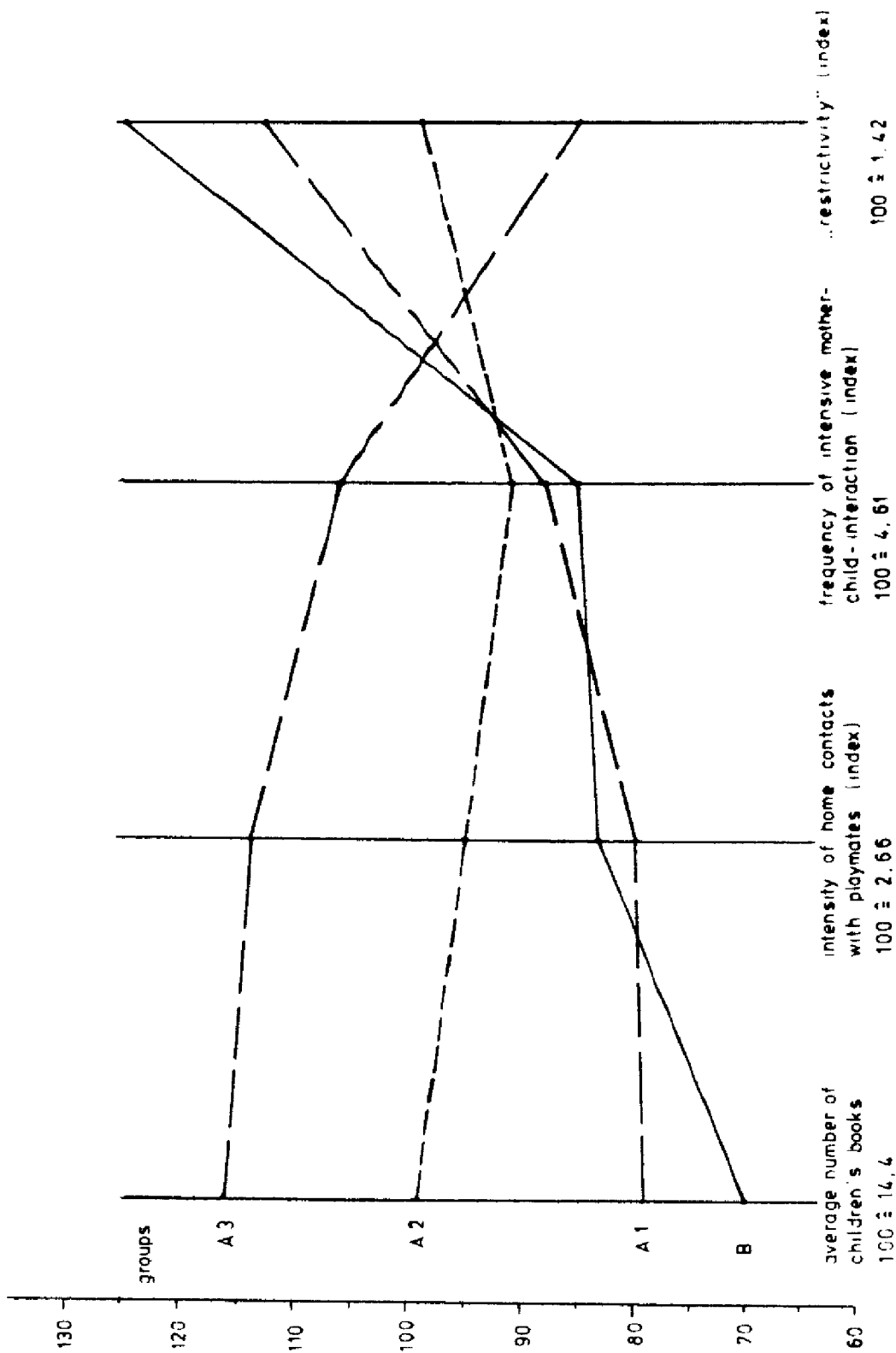


FIGURE 8.3: Average scores on indicators of parental behavior for groups of parents with different degrees of participation in nursery-schools ("Kindergarten")

—parental “*restrictivity*”—that is, parents not allowing their children to play with “dangerous” objects such as a saw or knife.

Figure 8.3 shows profiles across the mean scores on those indicators for the four “treatment categories” introduced above.

There is almost no difference between “control group” B and group A1. They are the most restrictive, and we find them below average on the other dimensions. The scores for group A2 are almost average (100) on every dependent variable, whereas the profile of group A3, the one with the most intensive treatment, is significantly over average.

It is most likely, however, that these results are merely effects of the composition of groups as demonstrated in Figure 8.2. If, however, we select only those families from our sample for which intensive participation in the kindergarten would, according to their structure and life situation, be most untypical (that is, families with social status lower than average, number and age of children more than average, in financial and housing situations being both worse than average) and if we apply the same kind of analysis to these “problem families,” significant differences can be found. Note, however, that in Figure 8.4 the values taken as 100 are the respective means of this subgroup of “problem families” and not, as in Figure 8.3, those of the entire sample.

Figure 8.4 illustrates that among “problem families” only those in which parents participate in a kindergarten’s parental work most intensively (A3) show scores on the dependent variables that indicate (compared with other families in similar life situations) relatively favorable conditions for children. Thus, Figure 8.4 confirms an assumption that even depressed and underprivileged families show an effect of the treatment, provided it is extremely intense. However, it still must be demonstrated how and under which circumstances it will come to these effects.

We have analyzed the kind of causal processes involved here by means of multivariate (path) analysis. Thus, we had to differentiate the assumption that with depressed families compensatory effects of the kindergarten can be expected. They can be, provided certain minimum standards of living are fulfilled. Families with a middle-class background showed no demonstrable effects. Families from the lowest working-class strata, segregated in low-ranking social areas (which, nevertheless, were not marginal, such as homeless families or ethnic strangers) with a very low per capita income and with the lowest standards of accommodation and many children (even if the parents have participated in the kindergarten in a highly intense manner), have hardly benefited from the treatment as far as the quality of parent-child interaction at home is concerned. Only “restrictivity” appeared slightly influenced by the treatment itself and not by only the depressed life situation of the families. The simple reason for that apparently was that, in those families, minimum standards in their life situation were not available to form the

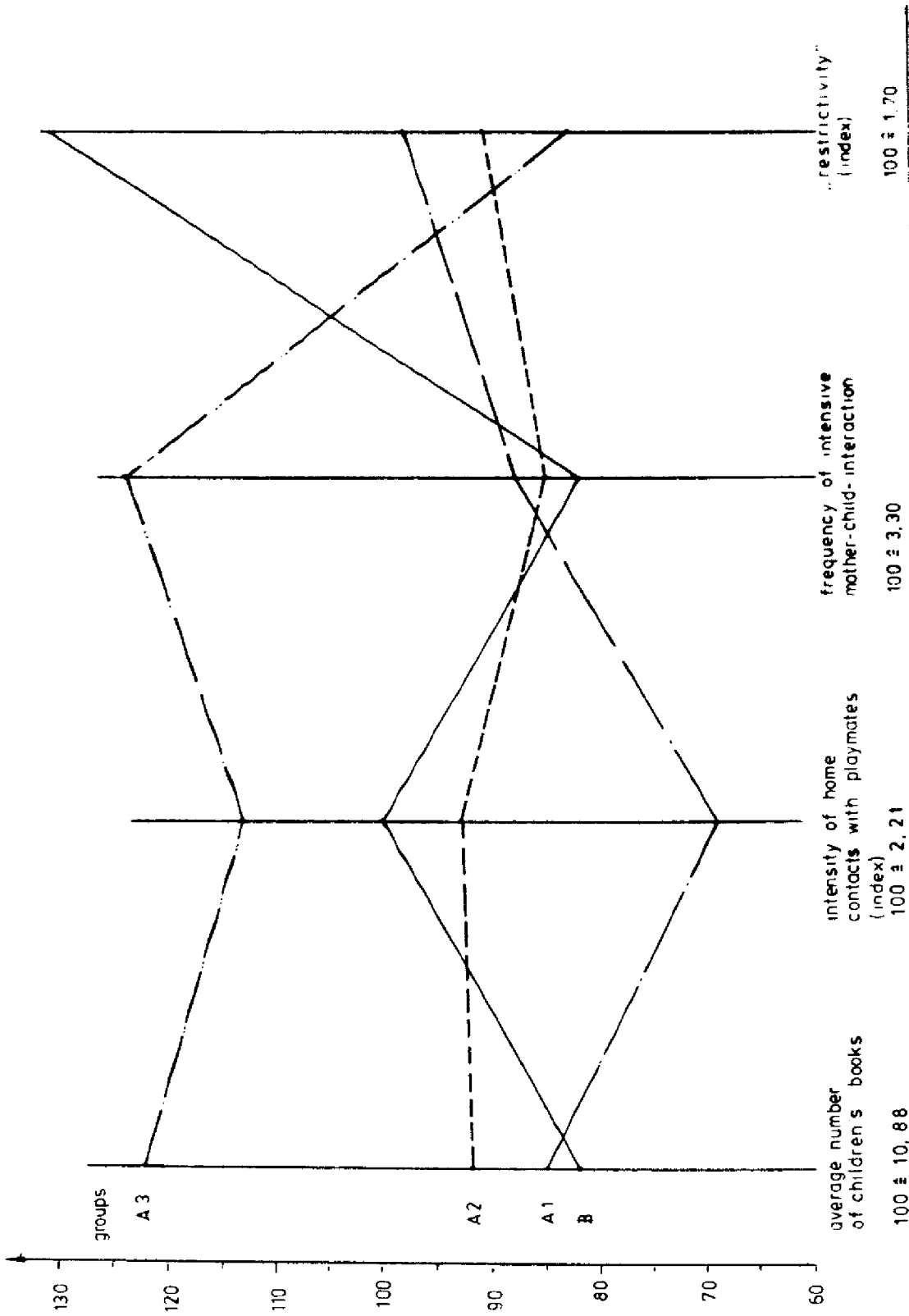


FIGURE 8.4: Average scores on indicators of parental behavior for groups of parents with different degrees of participation in nursery-schools ("Kindergarten"), only: "problem families"

basis upon which parental participation as a political treatment can bring about the desired compensatory effects. Consequently, working-class families which at least fitted those minimal standards appeared broadly effectuated by the program.

SUMMARY

Our study comprises a critical discussion of "traditional" approaches in social scientific evaluation research which more or less follow the experimental model of the natural sciences. The evaluation of political intervention, however, demands research designs that do *not* presuppose causal closure of processes studied. The scientific value and the practical use (the policy-improving function) of an evaluation do not depend upon a research design's formal exactness but, instead, result from its "object" and "problem adequacy." Depending upon the object of research, we suggest deliberate application of qualitative or quantitative research techniques; certain research problems may demand a combination of both.

Different from social policy in the United States, political intervention in the Federal Republic is generally implemented on the basis of permanent laws and not in the form of revisable "programs." To be "problem adequate" evaluation research in Germany demands empirical designs which have to be more complex than those merely applicable to the evaluation of limited political "programs." Possible research strategies are demonstrated using the example of our own research project on the effects of social policy upon the socialization process. Rather than simply measuring how far program-defined goals may have been attained, we suggest a complex and theoretically confirmed "impact model" enabling us to analyze the modes in which political intervention hypothetically comes to effect and considering subjective and objective factors that explain the utilization of programs (or social services) by their target population. This comprehensive approach consequently leads to more than just the identification of global "casual" relationships. In fact, it allows us to demonstrate how far the effects of political intervention are influenced by the social context in which policies normally operate.

NOTES

1. Ideally, the results of applying evaluation research in the political process perform a policy-improving function (Kaufmann, 1977, 1980). In fact, under specific conditions of interaction between science and politics, evaluation research can also have a "killer function" as well as merely a "legitimizing" function. "Applied" social research, consequently, is not a distinct concept. When the application and applicability of research results are discussed in this chapter, we concentrate upon the first of the above-mentioned types of application—improving or rationalizing policy.

2. We hesitate to use the term "program" because it does not characterize the legally regulated (and not explicitly goal-oriented) type of established policies of concern. The proper term, "*Sozialleistungen*," however, cannot be adequately translated as "social services." We prefer to use the term "program" or "policy(ies)" synonymously in the meaning of monetary or personal aids and services delivered to families by established state agencies.

REFERENCES

- ACHINGER, H. (1958) *Sozialpolitik als Gesellschaftspolitik. Von der Arbeiterfrage zum Wohlfahrtsstaat*. Reinbek bei Hamburg: Rowohlt.
- BLANKERTZ, H. (1976) Was heißt "Erfolg" oder "Scheitern" von Bildungsreformen. *GFPF-Materialien*, 8:29-46.
- BLUMER, H. (1973) "Der methodologische Standort des symbolischen Interaktionismus." pp. 80-146 in *Arbeitsgruppe Bielefelder Soziologen (eds.), Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit (vol. 1)*. Reinbek bei Hamburg: Rowohlt.
- BREEDLOVE, J. L. (1972) "Theory development as a task for the evaluator." pp. 55-70 in E. J. Mullen et al. (eds.), *Evaluation of Social Intervention*. San Francisco: Jossey-Bass.
- BRONFENBRENNER, U. (1974) *Wie wirksam ist kompensatorische Erziehung?* Stuttgart: Klett.
- CAMPBELL, D. T. (1974) "Qualitative knowing in action research." Presented at the meeting of the Society for the Psychological Study of Social Issues, New Orleans, September 1.
- _____ and J. STANLEY (1966) "Experimental and quasi-experimental designs for research on teaching." pp. 171-246 in N. L. Gage (ed.), *Handbook of Research on Teaching*. Chicago: Rand McNally.
- _____ (1977) "Methodische Vorstudie für die Analyse der städtebaulichen und stadtstrukturellen Wirkungen ausgewählter Sanierungsmaßnahmen nach dem Städtebauförderungsgesetz (final report). Unpublished manuscript. Berlin.
- HELLSTERN, G. M. and H. WOLLMANN (1978) *Sanierungsmaßnahmen. Städtebauliche und stadtstrukturelle Wirkungen*. Bonn-Bad Godesberg: Schriftenreihe "Stadtentwicklung" des Bundesministers für Raumordnung, Bauwesen and Städtebau.
- HERLTH, A. and K. P. STROHMEIER, (1980) "Sozialpolitik und der Alltag von Kindern." In L. Vaskovics (ed.), *Der Mensch als soziales und personales Wesen (vol. 6). Sozialökologische Einflußfaktoren familialer Sozialisation*. Stuttgart: Klett.
- JONES, W. C. and E. F. BORGATTA, (1972) "Methodology of evaluation." pp. 39-54 in E. J. Mullen et al. (eds.), *Evaluation of Social Intervention*. San Francisco: Jossey-Bass.
- JUHOS, B. (1960) "Welche begrifflichen Formen stehen der empirischen Beschreibung zur Verfügung?." pp. 101-158 in E. Topitsch (ed.), *Probleme der Wissenschaftstheorie*. Festschrift für Victor Kraft. Vienna: Springer.
- KAUFMANN, F.-X. (1980) "Social policy and social services: some problems of policy formation, program implementation and impact evaluation." In D. Grunow and F. Hegner (eds.), *Responsiveness of Social Service Organizations to Clients' Needs*. New York: Praeger.
- _____ (1977) "Sozialpolitisches Erkenntnisinteresse und Soziologie. Ein Beitrag zur Pragmatik der Sozialwissenschaften." pp. 35-75 in C. von Ferber and F.-X. Kaufmann (eds.), *Soziologie und Sozialpolitik. Sonderheft 19 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Opladen: Westdeutscher Verlag.
- _____ and SCHNEIDER, S. (1975) "Modelleinrichtungen—ein Instrument für experimentelle Reformverfahren in der Sozialpolitik?" *Neue Praxis* 3: 206-218.
- KAUFMANN, F.-X., A. HERLTH, and K. P. STROHMEIER (1980) *Sozialpolitik und familiäre Sozialisation—zur Wirkungsweise öffentlicher Sozialleistungen*. Stuttgart: Kohlhammer.

- KAUFMANN, F.-X., A. HERLTH, H.-J. SCHÜTZE, and K. P. STROHMEIER (1979) "Wirkungen öffentlicher Sozialleistungen auf den familialen Sozialisationsprozeß." pp. 181–211 in K. Lüscher (ed.), *Sozialpolitik für das Kind*. Stuttgart: Klett.
- _____. (1978) "Sozialpolitik und familiale Sozialisation. Schlußbericht des Forschungsprojekts "Wirkungen öffentlicher Sozialleistungen auf den Sozialisationsprozeß." Unpublished.
- KÖNIG, H. (1968) "Beobachtung und Experiment in der Sozialforschung." pp. 17–47 in R. König (ed.), *Beobachtung und Experiment in der Sozialforschung. Praktische Sozialforschung* (vol. 2). Cologne: Kiepenheuer und Witsch.
- LAZARSELD, P. F. (1965) "Wissenschaftslogik und empirische Sozialforschung." In E. Topitsch (ed.), *Logik der Sozialwissenschaften*. Cologne: Kiepenheuer und Witsch.
- LUHMANN, N. (1970) *Soziologische Aufklärung—Aufsätze zur Theorie sozialer Systeme*. Opladen: Westdeutscher Verlag.
- LÜSCHER, K. [ed.] (1979) *Sozialpolitik für das Kind*. Stuttgart: Klett.
- _____. (1977) "Sozialpolitik für das Kind." pp. 591–628 in C. von Ferber and F.-X. Kaufmann (eds.), *Soziologie und Sozialpolitik. Sonderheft 19 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Opladen: Westdeutscher Verlag.
- MATTHES, J. and F. SCHÜTZE (1973) Introduction: "Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit." pp. 11–53 in *Arbeitsgruppe Bielefelder Soziologen* (eds.), *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit* (vol. 1) Reinbek bei Hamburg: Rowohlt.
- OPP, K. D. (1976) *Methodologie der Sozialwissenschaften*. Reinbek bei Hamburg: Rowohlt.
- POPPER, K. R. (1964) "Naturgesetze und theoretische Systeme." pp. 87–103 in H. Albert (ed.), *Theorie und Realität. Ausgewählte Aufsätze zur Wissenschaftslehre der Sozialwissenschaften*. Tübingen: Mohr.
- SCHEUCH, E. K. and D. RÜSCHEMEYER (1965) "Soziologie und Statistik—über den Einfluß der modernen Wissenschaftslehre auf ihr gegenseitiges Verhältnis." pp. 345–363 in E. Topitsch (ed.), *Logik der Sozialwissenschaften*. Cologne: Kiepenheuer und Witsch.
- SCHÜTZE, F. (1976) "Zur Hervorlockung und Analyse thematisch relevanter Geschichten im Rahmen soziologischer Feldforschung—dargestellt an einem Projekt zur Erforschung von kommunalen Machtstrukturen." pp. 159–260 in *Arbeitsgruppe Bielefelder Soziologen* (eds.), *Kommunikative Sozialforschung*. Munich: Fink.
- SIMON, J. L. (1970) "The concept of causality in economics." *Kyklos* 23: 226–255.
- SMELSER, N. J. (1967) "Sociology and the other social sciences." In P. F. Lazarsfeld, W. H. Sewell, and H. L. Wilensky (eds.), *The Uses of Sociology*. London: Basic Books.
- STEGMÜLLER, W. (1960) "Das Problem der Kausalität." pp. 170–190 in E. Topitsch (ed.), *Probleme der Wissenschaftstheorie. Festschrift für Victor Kraft*. Vienna: Springer.
- STROHMEIER, K. P. and A. HERLTH (1979) "Sozialräumliche Bedingungen familialer Sozialisation." In H. Walter (ed.), *Region und Sozialisation* (vol. 2). Stuttgart: Frommann-Holzboog.
- SUCHMAN, E. A. (1972) "Action for what? A critique of evaluative research." pp. 52–84 in C. H. Weiss (ed.), *Evaluating Action Programs. Readings in Social Action and Education*. Boston: Allyn & Bacon.
- WEBER, M. (1964) *Wirtschaft und Gesellschaft*. J. Winkelmann, (ed.). Cologne: Kiepenheuer und Witsch.
- WEISS, C. H. [ed.] (1972) *Evaluating Action Programs. Readings in Social Action and Education*. Boston: Allyn & Bacon.
- WILSON, T. P. (1973) "Theorien der Interaktion und Modelle soziologischer Erklärung." pp. 54–79 in *Arbeitsgruppe Bielefelder Soziologen* (eds.), *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit* (vol. 1). Reinbek bei Hamburg: Rowohlt.