SUBJECT REACTIVITY IN DIRECT OBSERVATIONAL ASSESSMENT:
A REVIEW AND CRITICAL ANALYSIS

Francis C. Harris
Western Psychiatric Institute and Clinic
University of Pittsburgh School of Medicine
and
Benjamin B. Lahey
University of Georgia

ABSTRACT. Subject reactivity presents a threat to the internal and external validity of observational research. The effect of several factors on subject reactivity were discussed. They include differential subject sets, simple reactivity to observation, participant observation, observation by coders having a history of interaction with subjects, differential rationales given to subjects, and condition-related reactivity. The relevant work was reviewed and suggestions for minimizing the effects of these potential sources of bias were made.

Behavioral scientists long have been concerned with the effects of their measurement procedures on the objects of study (Rosenthal, 1966). The relatively recent proliferation of studies in which human observers (or mechanical data collection devices) are introduced into naturalistic settings (Kelly, 1977) has been accompanied by a line of methodological research and discussion focusing on: (a) statistical considerations in recording system bias such as the assessment of interobserver agreement (e.g., Harris & Lahey, 1978; Hartmann, 1977) and time sampling methods (e.g., Farkas & Tharp, 1980; Powell, Martindale, & Kulp, 1975; Powell & Rockinson, 1978; Repp, Roberts, Slack, Repp, & Berkler, 1976; Sanson-Fisher, Poole, & Dunn, 1980; Sanson-Fisher, Poole, Small, & Fleming, 1979). (b) recording system biases which might cause inaccurate coding behavior (Johnson & Bolstad, 1973; Kazdin, 1977; Kent & Foster, 1977; Lipinski & Nelson, 1974; Harris & Lahey, 1978).

Requests for reprints may be sent to Francis C. Harris, Ph.D., Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA 15213.

Appreciation is expressed to Rex L. Forehand, Joan L. Jackson, Jack O. Jenkins, and Gene H. Brody for critical readings of an earlier version of the manuscript.

523
Note 1), and (c) subject reactivity (Johnson & Bolstad, 1973; Kent & Foster, 1977; Lipinski & Nelson, 1974). The purpose of the present work is to analyze the subject reactivity literature and make suggestions which might enable observational researchers to minimize the influence of subject reactivity. Deliberately excluded have been studies of phenomena commonly referred to as experimenter bias (Azrin, Holz, Ulrich, & Goldiamond, 1961; Rosenthal, 1966), examiner bias (Masling, 1965; 1959), and reactivity to self-monitoring (Nelson, Lipinski, & Black, 1976).

In the present review six different types of subject reactivity studies are discussed. They are (1) studies which have shown that reactivity can occur under certain circumstances, (2) studies investigating simple effects of observation with no attempt made to analyze reactive effects of different components of the observation procedures, (3) studies analyzing the reactive effects of observers having histories of interaction with subjects, (4) studies of reactivity to participant observers who normally are present in the observational setting (e.g., teachers and parents), (5) studies of reactivity to personal characteristics of observers, and (6) studies of reactivity produced by the interactive effects of observation and intervention.

Four basic paradigms have been used in these reactivity investigations (Baum, Forehand, & Zegiob, 1979). In the first paradigm parents or teachers have been instructed to make their children look “good” on certain occasions and not on other occasions. Differences in behavior between conditions are assumed to show that mediators have the capacity to alter their own and their children’s behavior in the presence of an observer (Green, Forehand, & McMahon, 1979: Johnson & Lobitz, 1974; Lobitz & Johnson, 1975; Weinrott & Jones, 1977). In the second paradigm, different levels of observer conspicuousness are compared. For example, the presence and absence of a classroom observer might be varied systematically while data are collected. during observer presence and absence, by other observers stationed behind one-way glass in an adjacent observation room. In the third paradigm, subjects are informed that they are being observed on some occasions and not being observed on other occasions when, in fact, they are observed under both conditions. In the fourth paradigm an observer is introduced and subject behavior is looked at over time, with changes assumed to reflect the habituation of reactivity effects.

MANIPULATION OF SETS TO LOOK GOOD AND BAD

Johnson and Lobitz (1974) conducted six observations in the homes of 12 4- to 6-year-old nonbehavior-problem children. Parents were instructed to make their children look “good” during some observations and “bad” during other observations. Using a modified version of the Family Interaction Coding System (Patterson, Ray, Shaw, & Cobb, Note 1), observers found significantly more child deviant behavior on “bad” days than on “good” days. In addition, they recorded significantly more parental commands and negative responses to children on “bad” days. These results occurred in 11 of the 12 families observed. There was no difference in children’s noncompliance ratio (compliances ÷ total commands) between “good” and “bad” days, suggesting parents produced some deviant behavior (i.e., noncompliance) simply by increasing commands. Since noncompliance had been shown to account for at least one-third of child deviant behavior, a separate analysis of deviant behavior, with noncompliance subtracted, was conducted. Results of this analysis paralleled those of total deviancy, suggesting parents also produced other types of deviant behavior. In response to a postexperimental questionnaire, parents re-
ported they were more negative, restrictive, disapproving, and ignoring on “bad” days and more positive, attentive, and approving on “good” days. These results suggest that parents of normal children have the capacity to alter their own and their children’s behavior in the presence of an observer. This capacity is presumed relevant because parents, for a variety of reasons, might want to present their children favorably or unfavorably during observation (Johnson & Lobitz, 1974).

In a subsequent investigation (Lobitz & Johnson, 1975), 12 families of non-deviant children and an equal number of families of behavior problem children were observed in their homes according to the previously described paradigm (Johnson & Lobitz, 1974), with the addition of a “normal” condition in which parents were instructed not to attempt to make their children look good or bad. Ten of 12 deviant families and 9 of 12 nondeviant families produced more deviant behavior on “bad” days than “normal” days, accompanied by increases in parental commands and negative responses and decreases in positive consequences on “bad” days. No overall differences in child deviant behavior between “good” and “normal” days were noted, with 7 of the 12 nonproblem families and only 4 problem families showing differences. Questionnaire responses indicated parents felt they were better at producing good behavior than bad behavior. These questionnaire responses were discrepant with the observational data.

The results of both the above studies, taken together, suggest that family interaction data might reflect parents’ attempts to make the family look bad, but may reflect attempts to look good to a lesser degree (Reid, 1978). Thus, under baseline conditions parents who might want to present their families as needing treatment may produce inflated estimates of deviance; or posttreatment families, who might want to show improvements, could engage in behaviors to hold down negative behaviors.

Green, Forehand, and McMahon (1979) also showed that parents of normal and deviant children can modify the behavior of their children, in accordance with experimenter instructions, in the presence of an observer. Twenty mothers and their 3.9 to 8.3-year-old children participated. One-half of the mother-child pairs included previously referred deviant children. While being observed from behind one-way glass each mother, in a three-part session, (a) spent 10 minutes trying to make her child look obedient, (b) spent 10 minutes trying to make her child look disobedient, and (c) issued 20 specific commands to her child. Results showed mothers of deviant and normal children could manipulate obedience and disobedience. They appeared to do so by manipulating antecedents and consequences. They increased vague and inhibitory commands and criticism during the “look noncompliant” phase and used general and contingent rewards more often during “look compliant” sessions. In contrast to the Johnson and Lobitz (1974) study, parents produced noncompliance without varying their overall rate of commands.

Weinrott and Jones (1977) showed teachers of socially withdrawn first to third-grade children were able to manipulate the occurrence and nonoccurrence of the problem behaviors under differential instructional sets, while teachers of disruptive children were not able to manipulate their students’ disruptive behavior. Twenty pairs of teachers and disruptive children and 20 pairs of teachers and withdrawn children were observed in their respective classrooms, with other children present and participating in ongoing classes, for 6.5 weeks. During the 2.5-week baseline period, observations were made without any specific instructions to the teachers. In the next condition, the teachers were asked to do whatever they could to improve
the withdrawn or disruptive behavior. In the final 2-week condition, teachers were asked to stop making special efforts to improve the target children's problem behaviors. Results showed a significant increase in the social behavior of the withdrawn children during the instruction condition, accompanied by an increase in teacher responses to them. No such improvement was seen in the disruptive children, even though their teachers increased the frequency of their responses to them during the instruction condition. The authors recommended attempting to minimize the influence of differential demand conditions by: (a) including a high demand condition in intervention studies, (b) employing a correction factor to partial out variance accounted for by demand characteristics, or (c) including an attention-placebo control group.

Two other potential solutions to this type of bias have been proposed by Johnson and Lobitz (1974). The first would involve procedures to make observation procedures less obtrusive, including the use of audio or video equipment in lieu of physically present observers (e.g., Bernal, Gibson, Williams, & Pesses, 1971; Johnson & Bolstad, 1975; Johnson, Christensen, & Bellamy, 1976; Kent, O'Leary, Dietz, & Diament, 1979; Moos, 1968; Nordquist, 1971; Purcell & Brady, 1965; Soskin & John, 1963) or procedures to minimize the obtrusiveness of observers (e.g., Brechner, Linder, Meyerson, & Hays, 1974; Burton, 1971; Gerwitz, 1952). In general, ethical considerations might require subjects to be informed they might be observed at any time and that, in the case of audio or video recording, they could turn off the device at any time and otherwise be permitted to censor the recorded material prior to experimenters or coders monitoring the recordings (Johnson & Lobitz, 1974). The second approach would be to collect target behavior data via alternative methods, such as questionnaires or interviews, with convergence between observational and alternative methods strengthening conclusions based upon observational data (Johnson & Lobitz, 1974).

It should be noted that both of the suggestions offered by Johnson and Lobitz (1974) also would appear to be very useful in evaluating and minimizing each type of subject reactivity discussed in the following sections of this review.

PRIOR SUBJECT-OBSERVER INTERACTIONS

There is some evidence (Mash & Hedley, 1975) that the quality of prior interactions between subjects and observers might affect behavior during subsequent observation sessions. A group of children performed a simple motor task: dropping marbles into a box. Some of the children then were exposed to an adult who interacted negatively with them for a short time. Other children spent the same amount of time in the presence of a positively interacting adult. A third group spent the same amount of time without observation by an adult observer and stayed with the experimenter, who did not interact with them. The fourth group spent the same time period in the presence of the adult observer who did not interact with them. Those who experienced positive adult interactions and those who experienced neutral interactions significantly increased their performance rate during the posttest when the adult observer was present. The children who had interacted with the negative adult showed a significantly lower overall performance rate during the observer-present posttest.

Surratt, Ulrich, and Hawkins (1969) also demonstrated possible reactive effects of observers who had previously interacted with subjects. A fifth-grade boy was
taught to operate a console that controlled lights on the desks of four first-grade subjects. He turned on the lights whenever he observed that the child assigned to a particular desk was sitting and working. Tokens representing the number of minutes the lights had been on were exchanged later for admission to a class party. During a follow-up assessment in which the operator's presence and absence was varied while all sessions were videotaped, the boy was shown to serve as a discriminative stimulus for working even though he no longer controlled the lights. Certainly, this type of prior interaction is different from the type we might expect in other observational situations.

The results of these two investigations indicate the importance of using observers who have no history of interaction with their subjects. A possible exception to this rule might be the use as data collectors of people normally present in the observational situation during times when data are not being collected. These people might include parents, teachers, or institutional staff members. Such "participant observation" is discussed below.

**SUBJECT AND OBSERVER CHARACTERISTICS**

It has been suggested that subjects' relative susceptibility to reactivity might be correlated with personality variables, such as different scales of the MMPI (Johnson & Bolstad, 1973; Moos, 1968). It also has been suggested that younger children might be less reactive to observation than older children. In a study by Martin, Gelfand, and Hartmann (1975), described in detail below under "Simple Reactivity," there was significantly more imitative aggressive behavior emitted in the presence of same-sex peer observers than adult same-sex or peer-opposite-sex observers. It must be noted that the observers in this study did not function as observer-coders in the conventional sense.

Zegiob and Forehand (1976) demonstrated that middle and lower class mothers were equally reactive to laboratory observations in terms of their interactive behaviors with their children.

Despite the call for research to determine which observer and subject characteristics might maximize and minimize reactivity (Johnson & Bolstad, 1973), there is a general lack of such data.

**PARTICIPANT OBSERVATION***

Direct observational data collected by people normally present in the environment is appealing to behavioral assessors. In comparison to nonparticipant observers, they are more likely to be present for the occurrence of low frequency behaviors and are less costly, permitting more frequent observations (Hay, Nelson, & Hay, 1977). Also, the possibility of participant observers minimizing subject reactivity exists because they would be less obtrusive than nonparticipant ("stranger") observers (Kubany & Sloggett, 1973).

Hay, Nelson, and Hay (1977) investigated the reactive effects of observation by

*The term "participant observation" also has been used to describe observations by observer-coders stationed in the observation setting solely to collect data as well as data collection by individuals normally present in the environment of interest (Wiggins, 1973). In the present work only the latter circumstance is considered participant observation.
teachers on the behavior of the teachers themselves and their elementary school children. Four students from each of eight classrooms and their teachers were observed for one hour per day for ten schooldays across two phases. Two of the children in each class were chosen on the basis of their teacher's reports of their frequent off-task behavior. The other two were selected on the basis of previous direct observations of high rates of off-task behavior. In the first phase observers coded student behavior as disruptive, appropriate, or passive and teacher behavior as either prompts, criticisms, or praises. Prior to the second phase the teacher was instructed in the procedure for coding child behavior. During continuing observations by independent coders, the teachers recorded the behavior of two of the four students participating in the experiment. Analysis of the independent coders' recordings of teacher behavior showed the frequency of prompts to the students who were observed by the teacher was greater during the teacher-recording phase than the baseline phase, while prompts to the children not observed systematically did not differ between the two phases. There were no between or within teacher differences in praise or criticism. Students who were observed by the teacher showed greater changes in appropriate behavior from baseline to teacher observation conditions than the control children. Significant interactions of phase and referral status for passive and disruptive behavior were observed such that for referred students teacher-observation affected disruptive behavior and for nonreferred students teacher-observation affected passive behavior.

A subsequent investigation (Hay, Nelson, & Hay 1980), in which individual subject data were presented, also suggests that the behavior of some participant observers and their subjects is affected by the observation process. The study was conducted in the classrooms of two first-grade and two second-grade teachers. The behavior of the teachers and four behavior problem children in each class was coded by independent observers for a total of 35 days across baseline and intervention conditions. In each class a multiple baseline design across students was employed. During the baseline phase, standard classroom procedures were followed as the coders recorded each students' appropriate and inappropriate verbalizations as well as teachers' positive, negative, and instructional behaviors toward each student. During the intervention two of the teachers counted students' appropriate verbalizations and the other two counted their students' inappropriate verbalization. No significant effects were found in three of the classes. However, statistically reliable effects were found in one class. One teacher's frequencies of positive and instructional behaviors were greater during the intervention, in which she recorded students' appropriate verbalizations, than during the baseline condition. In addition, students' appropriate verbalizations were more frequent during the intervention. Comparisons of the independent observers' and teachers' recordings revealed that: 1) although there were strong correlations, $r = .77 - .83$, the teachers recorded significantly fewer occurrences of each target behavior, and 2) the absolute magnitude of the difference increased as the behavior frequency increased, $r = .50 - .62$. Unfortunately, no description of the way in which the teachers or independent coders were trained was provided. Thus, differences in recorded behavior may have been a function of different training methods. The fact that reactivity effects attributable to the participant observation procedure were seen only in one of the four classrooms obviously weakens the generality of the findings. However, these results, in combination with the group results in the earlier
study (Hay, et al., 1977), certainly do cast some doubt on the degree to which the participant observation procedure is free from reactivity.

Barker and Wright (1955) reported an observational study in which coders were permitted to interact with the family members that they were observing in naturalistic situations. Although the authors claimed their study provided evidence of only limited reactivity, Johnson and Bolstad (1973) pointed out that the results, which showed 20% of the observed children's interactions were with the observers, actually indicate a great deal of reactivity.

These results, along with data showing therapeutic effects of teacher recording of problem behavior (Forehand, 1973) and reactive effects of self-monitoring (Nelson, Lipinski, & Black, 1976) suggest that caution should be exercised when using participant observers. In fact, unless data are forthcoming to show that participant observation does not produce reactivity in a given situation (directly or indirectly) by affecting changes in participant observers' behavior toward subjects, participant observation should not be depended upon to provide accurate data.

Given the previously described potential advantages of participant observation, future research efforts should be directed at evaluating observer training strategies designed to teach participant observers to minimize the reactivity they produce.

**Rationale for Observation**

Johnson and Bolstad (1973) suggested that giving subjects a thorough rationale for the collection of observational data might reduce guardedness and anxiety and, hence, reactivity. Although this hypothesis still awaits empirical validation, it has some face validity and is consistent with standard ethical practices for the use of human subjects (American Psychological Association, 1973).

**Simple Reactivity to Observation**

For the purpose of the present review, simple reactivity studies are those which investigate the influence of observer presence and make no attempt to determine the effect of specific components of the observation procedure. Simple reactivity to observation has been investigated in the context of several settings, subjects, and target behaviors. They include counselor and client verbal behavior in an outpatient setting, institutional staff and resident verbal behavior, teacher and student behavior in normal and special classrooms, and family members' behavior in laboratory and naturalistic settings.

**Client and Counselor Behavior**

Roberts and Renzaglia (1965) used covertly made recordings of counseling sessions to demonstrate that clients solicited from undergraduate classes and practicum student counselors talked differently depending on whether or not they had reason to believe the session was being recorded. Eight counselor-client pairs participated in a series of counseling sessions. During some sessions recording equipment was in full view of the participants and in the others they were told no recording was being made. In fact, all sessions were recorded. Clients made significantly more positive self-statements in the presence of the recording equipment, while coun-
seior made significantly fewer client-centered statements when the equipment was visible.

**Institutional Staff and Resident Behavior**

Hagen, Craighead, and Paul (1975) studied the effects of observation in an institutional setting where direct observations of staff had gone on for a considerable period of time. Staff verbalizations to residents were monitored continuously via a microphone covertly placed in one section of a unit. The presence and absence of an observer in that section of the unit was varied systematically during times when a staff member was assigned to make programmatic interactions with a group of residents. No difference in staff’s “correct” verbalizations was found between observer-present and absent conditions. The authors speculated that no difference was found because habituation to observers had occurred, observations were scheduled frequently and unpredictably, and the staff members had been trained thoroughly to use a “correct” style of interaction.

Purcell and Brady (1965) reported that adolescents in a day treatment center who were outfitted with radio transmitters for one hour per day reduced their comments about being observed to zero within two days. The elimination of such comments obviously provides limited information on the reactivity of other behaviors.

**Laboratory Observations of Children and Parents**

Martin, Gelfand, and Hartmann (1971) investigated the effects of observation on 50 male and 50 female preschool children. Immediately after exposure to an adult model hitting a doll, individual children were taken to a room containing an identical doll under one of the following conditions: (a) no observer, (b) adult male observer, (c) adult female observer, (d) male preschool student observer, or (e) female preschool student observer. All data were collected by coders stationed behind one-way glass. Results showed no difference between the no-observer condition and any of the other conditions. These results should be interpreted cautiously, however, since the in-room observers did not actually function as data collectors. They sat in the room with the subjects but did not attend to them in any way. The adult observer read a book and the child observer colored and drew pictures.

White (1977) investigated the effects of observation on the activity levels of families in a laboratory situation. During two hours of observation presence and absence was varied, while other observers coded from behind one-way glass. During observer presence 50% less moving about the room was recorded. The author pointed out that these results might not hold over longer observation periods or observations in homes.

Three laboratory studies have looked at reactivity in parent-child interactions (Zegiob, Arnold, & Forehand, 1975; Zegiob & Forehand, 1978; Zegiob, Forehand, & Resick, 1979). In the first study (Zegiob et al., 1975) 12 mothers and their nursery school children, six boys and six girls, were observed via one-way glass during informed and uninformed conditions of 15 and 11 minutes, respectively. Three maternal behaviors, positive verbalizations, interactive playing, and structuring...
(questions and commands) occurred significantly more often during informed observations. Recordings of three other behaviors, negative verbal, out of contact, and helping, did not differ between conditions.

One component of another investigation (Zegioh & Forehand, 1978) also compared parent and child behavior under informed and uninformed observation conditions. Fourteen middle class and 14 lower class mothers and their preschool age children were observed for 16 minutes during informed conditions and 15 minutes during uninformed conditions. Results generally confirmed those of the previous study (Zegiob et al., 1975). For both groups, mothers' interactive playing, questions and commands, positive verbal behavior, and helping behavior were significantly greater during informed observations, while out of contact was significantly less during informed conditions, and negative verbal behavior occurred hardly at all during either condition. In addition, children in both groups engaged in more solitary play during the informed condition.

Zegiob, Forehand, and Resick (1979) examined habituation of presumed reactivity effects in three mother-child pairs. The three pairs were observed for seven, eight, and nine 30-minute sessions. One pair was observed approximately once each week and a half for 10 weeks, while another was observed three times per day for three days, with one week separating observation days. The third pair was observed three times per week, with sessions occurring every other day for 2.5 weeks. All three mothers' behavior showed quite similar habituation effects even though each one experienced different densities of observation sessions. Mothers' positive verbal, helping, and interactive playing behaviors decreased over sessions, with an immediate increase when a different observer was introduced, followed by rehabituation. Mothers' negative verbal behavior did not show consistent habituation effects. The authors suggested that the use of the same observer across sessions would facilitate habituation. While these results provide some support for the use of laboratory observational data following an habituation period, caution should be exercised in generalizing these results due to the small number of subjects used.

Results of these laboratory investigations of parent-child interactions strongly suggest that the frequency of maternal behaviors, especially positive ones, is inflated during observer present conditions, although there is some evidence that habituation to observation might occur after a several session "warm-up" period. The demand characteristics of the laboratory situation appear to inhibit negative maternal behaviors (Zegiob & Forehand, 1978; Zegiob, et al., 1979), at least during the limited observation periods employed. Further research certainly is necessary to determine which child behaviors are reactive to observation in laboratory situations.

Classroom Observations of Teachers and Students

An early study of simple reactivity effects (Masling & Stern, 1969) investigated habituation to observation in 23 different fourth- and fifth-grade classrooms. Correlations between the three successive observation periods for five student behaviors did not show consistent habituation effects across behaviors. The absence of statistical analysis and relatively brief duration of the study limit the generality, reliability, and usefulness of these data.
In another study (Sattler & Swoope, 1976), teachers' token delivery was monitored on selected occasions over a 4- to 6-week period. The study occurred in the context of a classroom management class participated in by the teachers. During all class periods the children colored squares on a chart whenever they received a token. The number of colored-in squares served as the dependent variable during the observer-absent condition, and observer collected data served as the independent measure during observer presence. Results showed significantly more token delivery during the observer-present condition. Unfortunately, no information regarding the observers was provided to permit the ruling out of their association with the classroom management class as an explanation for the effect. A major flaw in the study involved the use of observer collected data in one condition and the children's recordings in the other. Thus, the obtained results may have been the result of different data collection procedures.

Mercatoris and Craighead (1974) investigated reactivity to classroom observers in an institutional special education language class comprised of six developmentally disabled adult females. A camera was placed in the classroom for two weeks prior to the study. The teacher was told the camera would be in use only when an observer was in the classroom to operate it. During the first phase of the experiment no observer was present, and, in fact, from the teacher's and students' points of view, it was not different from the adaptation period. In the next 10-day condition an observer was present in the classroom. The next two 5-day conditions replicated the first two, providing an ABAB design. Videotapes of all sessions, observer present and absent, were coded by observers who were unaware of which condition was in effect for any given tape. Results showed an increase in the total number of student-teacher interactions during observer present conditions, which did not appear to decrease as a function of habituation to observation. There was no difference in the appropriateness of child behavior between the two conditions. In addition, the number of tokens delivered by the teacher was correlated with the intervention.

A series of three well-controlled studies of reactivity to classroom observers was conducted by Dubey, Kent, O'Leary, Broderick, and O'Leary (1977). In all three studies, the Disruptive Behavior Code (O'Leary, Kaufman, Kass, & Drabman, 1970) was used by coders stationed behind one-way glass to record the behavior of eight severely deviant conduct-problem children. In the first experiment an observer was present and absent in the classroom on alternating days for a 14-day period in which a token economy program for classroom behavior was in effect. No reactivity to observation was shown in any of the categories. In the second study, the children's behavior was coded under baseline and token economy conditions in a relatively unstructured math class, while several categories of teacher behavior were recorded only during the token economy condition. Results showed a significant interaction of phase and observer presence for students' off task behavior and a significant main effect observer presence on teachers' "educational comments" and a composite "overall teacher behavior" category during the token economy condition. The third study was undertaken to attempt to replicate the significant results obtained in the second study. Observing teachers and students in highly structured classes during 10 baseline and 10 token economy sessions, they found no significant main or interaction effects of observation. A subsequent series of observations during unstructured mathematics classes in baseline and token economy conditions showed a significant main effect of observer presence for students time off task. This effect was not found in the first two studies. Across all three studies, 63 tests
(main and interaction) for observer effects were conducted and only three reached statistical significance. This was fewer than the number that would be expected to occur by chance. Moreover, the attempts to replicate the significant effects that were found were unsuccessful. Since only group data were reported, individual subjects might have been reactive differentially. If that were the case, these results would not be generalizable to single subject investigations.

A similar study was conducted by Weinrott, Garrett, and Todd (1978). Observations of six teacher-referred socially aggressive boys were made in two different classes (i.e., highly structured math and less structured language) from behind one-way glass. Observer presence was varied systematically during all four phases (ABAB) of a 23-day evaluation of a token program. Analyses of appropriate and inappropriate classroom behavior categories showed no significant reactivity effects during either condition for the group, as a whole, or any individual subjects. It should be pointed out that the two composite categories in which data were presented were formed by collapsing 17 molecular categories, such that reactivity may have occurred in some of the molecular codes.

A multiple baseline analysis of subject reactivity in four "normal" junior high school classrooms was conducted by Nelson, Kapust, and Dorsey (1978). Coders recorded student out-of-seat, student and teacher looking at the observer (or observation window), student volunteering, student on task, teacher at desk, teacher speaking, teacher attending to students, and student and teacher manipulating objects. No reactivity effects were found for any of the behaviors; but, again, no data on individual subjects were presented.

In general, well-controlled studies have failed to demonstrate consistent reactivity effects in classroom settings. It should be noted, however, that reactivity to classroom observers may occur as a function of other factors covered in this review, such as prior interactions between observers and subjects or particular characteristics of observers. It also should be pointed out that these investigations were conducted in classrooms containing one-way glass or cameras and that teachers and perhaps students were aware, in most cases, of the possibility of covert observation. Thus, these results might not be generalizable to situations where covert observation is less likely.

**Naturalistic Observations of Family Members**

Johnson and Bolstad (1975) investigated simple reactivity in the homes of 12 families of children 4–8 years old. An audiotape recorder in the families' homes was activated for 45 minutes on six consecutive days. An observer was present for three of the observations. Paid coders then used a modified version of the Family Interaction Coding System (Patterson et al., Note 1) to collect data on child deviant behavior, parent negative behavior, and parent commands. Results showed no differences in any of the behaviors during observer-present and absence conditions and no evidence of habituation to observation within or across sessions. These results suggest comparability of data collected by human observers and recording equipment but should not be taken as a demonstration of generalizability to totally unobtrusive data collection situations. As the authors indicated, arranging a completely covert assessment situation would be virtually impossible and unethical in home observations. Thus, data collected with in-home observers and recording equipment may not accurately reflect absolute levels of behavior when overt data collection is not in effect.
In an evaluation of the use of recording equipment for home data collection, Bernal et al. (1971) had a family activate the equipment for 7 minutes four times per day for 6 weeks. An observer was present during one of the sessions per week. A correlation of $r = .89, p < .02$, was obtained between mothers' commands during observer-present and absent conditions, suggesting comparability of data collected by the two methods.

The fact that reactivity has been seen more often in family interactions in laboratory observations than home observations probably is due to the difference in the obtrusiveness levels of the observation procedure during observer-absent conditions. In the laboratory studies, data were collected via relatively covert procedures including hidden recording equipment and one-way glass, while in the home studies observer-absent data were collected via recording equipment which was clearly visible, identifiable, and, in one case, even operated by the family members being observed.

### Summary of Simple Reactivity Literature

In general, simple reactivity effects have been seen in counselor-client and parent-child laboratory interactions and not in on-unit institutional staff verbal behavior, family interactions in home settings, and teacher and student classroom behavior. A possible explanation for this pattern of results might lie in the setting differences. Reactivity effects might be minimized in group settings where several people witness behaviors of individual subjects, such that the introduction of one more person (i.e., the observer) might be relatively unobtrusive and produce no reactivity effects. On the other hand, reactivity may be maximized in generally private settings where the introduction of an observer would be very obtrusive. A study by Bales (1950), cited by Johnson and Bolstad (1973), of three different levels of observer conspicuousness did not show any reactive effects in the behavior of participants in a discussion group. These results support this "public vs. private settings" hypothesis.

The relative absence of reactivity in classrooms and homes also might have been a function of the subjects' awareness of "covert" data collection. The ethical and procedural difficulties of including completely covert data collection in reactivity studies reflects a difficulty inherent in observational methodology in general. Unless generalizability to completely unaware circumstances can be assured, any observational data are subject to reactivity. This assurance of generalizability may be attained by (a) including comparisons of periodically collected covert data with data collected during aware conditions, or (b) reference to a data base comprised of comparisons of overtly and covertly collected data. If such a covertly collected data base did not exist and completely covert data collection were impossible, investigators might refer to a data base comprised of data from different levels of observer conspicuousness and/or vary levels of conspicuousness during an investigation. The absence of demonstrated generalizability to unaware situations argues for the use of alternative data collection methods to support observational data (Johnson & Lobitz, 1974).

### CONDITION-RELATED REACTIVITY

In addition to investigating simple reactivity to observation in the classroom, two previously described studies (Dubey et al., 1977; Weinrott et al., 1978) addressed a crucial question that had not received attention previously. It will be recalled that
observer presence was varied systematically while data were collected by other coders stationed behind one-way glass during baseline and token economy conditions, and that there was no main or interaction effect of observation. A significant condition X presence/absence interaction in such a design would pose a threat to the internal validity of the commonly used paradigm in which target behaviors are recorded during baseline and intervention conditions. That is, the observation procedure might interact with the intervention to produce results that easily could be mistaken for a treatment effect. This especially might be likely when subjects are kept unaware of the specific behaviors being recorded in one condition (e.g., baseline) and are made aware of them in other conditions (e.g., intervention). This differential awareness may be the result of specific information provided by a mediator or may develop as subjects come to learn just what is expected of them. In fact, subjects may develop new skills as a result of the intervention but only demonstrate them in the presence of an observer. The absence of condition-related reactivity in the special classroom studies certainly should be encouraging to education researchers, but it should alert other observational researchers to the possibility of such a confound in their own research.

Harris and Lahey (Note 2) suggested that condition-related reactivity might be a factor in non-classroom situations. In a multiple baseline across subjects analogue intervention study of preschool children's peer praising behavior, an observer's presence and absence was varied systematically during baseline and intervention sessions while each entire session was audiotaped covertly for later coding. Four pairs of children separately participated in a series of 12 ten-minute drawing sessions in which they were free to talk to each other. A film, modeling and describing peer praising, was shown to each pair of children following their final baseline session. Results showed that observer presence had no effect on praising during the baseline condition, but had significant effects during the intervention condition for the group as a whole and two of the four pairs. That is, what otherwise could have been explained as a treatment effect really was an effect attributable to condition-related reactivity. The authors suggested that several differences between their study and the earlier classroom investigations (Dubev et al., 1977; Weinrott et al., 1978) might have been responsible for the divergent results. These differences included the age and behavior problem status of the children, size and makeup of child groups, presence of a teacher who consequated recorded behaviors in the classroom studies, and target behaviors and settings (i.e., appropriate and inappropriate classroom behaviors in a relatively public setting as opposed to interpersonal behaviors in a relatively private setting. In any event, condition-related reactivity presents a serious problem to observational researchers and should be the object of future inquiry. Until such research can establish the circumstances in which condition-related reactivity is likely to occur, its effects may be minimized and/or evaluated by: (1) using the least obtrusive data collection method possible, (2) comparing observational data collected under different levels of observer conspicuousness, and (3) collecting target behavior information via means other than observational data in all experimental conditions.

CONCLUSION

Since subject reactivity constitutes a threat to external and internal validity (Campbell & Stanley, 1963), it certainly will continue to be a matter of concern to observational researchers. Investigators would be wise to become familiar with the
reactivity literature \textit{vis à vis} his or her subjects, observers, research settings, and target behaviors and take whatever precautions are dictated by that literature. If an absence of reactivity effects has not been demonstrated with a given combination of subjects, observers, procedures, settings, and target behaviors, it is incumbent upon individual investigators to convince research consumers that subject reactivity was minimized in their particular study. To this end, it is recommended that researchers (a) use procedures that minimize the obtrusiveness of the observation process such as electronic monitoring equipment (e.g., Bernal et al., 1971; Johnson & Bolstad, 1973; Johnson et al., 1976; Kent et al., 1979; Moos, 1968; Nordquist, 1971; Purcell & Brady, 1965; Soskin & John, 1963) and/or equipment such as one-way glass and portable observation booths (e.g., Brechner et al., 1974; Burton, 1971; Gerwitz, 1952) which minimize contact between observers and subjects, (b) periodically collect observational data through less conspicuous levels of observation than are used routinely during an investigation, and (c) include independent assessments of target behaviors using methods of assessment other than observational data (Johnson & Lobitz, 1974). Divergence between conspicuousness levels or between independent methods would alert investigators to an assessment problem.

As Harris and Lahey (Note 2) suggested, unless it has been well-documented that reactivity is not a factor in a given situation, observational data may be taken as a demonstration that a particular behavior is in a subject's repertoire, but not that it is performed in the absence of observation.

**REFERENCE NOTES**


**REFERENCES**


学霸图书馆
www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，
提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具