Order Number 9029971

Performance feedback and a token economy to increase telemarketing sales skills in a thrift institution

Pratt, Alison, Ph.D.

City University of New York, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

PERFORMANCE FEEDBACK AND A TOKEN ECONOMY TO INCREASE TELEMARKETING SALES SKILLS IN A THRIFT INSTITUTION

by

ALISON PRATT

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1990

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

4/25/90
Date

Nancy S. Hemmes
Dr. Nancy S. Hemmes
Chair of Examining Committee

4/25/90
Date

Herbert D. Saltzstein
Dr. Herbert Saltzstein
Executive Officer

Dr. Bruce Brown

Dr. Claire Poulson

Dr. Richard O'Brien

Dr. Walter Reichman
Supervisory Committee

The City University of New York

Abstract

PERFORMANCE FEEDBACK AND A TOKEN ECONOMY TO INCREASE

TELEMARKETING SALES SKILLS IN A THRIFT INSTITUTION

by

Alison Pratt

Advisor: Professor Nancy Hemmes

Performance feedback has become a widely used technique of behavior change in applied industrial and community settings. The current study used informational feedback and, later, a token economy, in an attempt to increase the sales skill ratios of three telemarketers in an inbound telemarketing department of a thrift institution.

A multiple baseline design across four sales skills (using the customer's name, asking for the sale, cross-selling, and stating product benefits) evaluated the efficacy of informational feedback. Telemarketers tape recorded their phone calls daily (without the customer being heard) from which data on sales skill usage was obtained. Performance feedback consisted of information regarding the percentage of skill usage from two to three days of calls from the previous week. Performance feedback was delivered for a total of 29 weeks. At week 20, a token economy for prizes from a gift catalog was added to the feedback intervention.

All sales skills data were analyzed with interrupted time series analysis. Results indicated that feedback reliably increased the use of sales skills in three out of nine cases. Reinforcement caused reliable increases in three out of six cases. However, many of the intervention effects were not maintained over time.

The thrift implemented a very successful product promotion from weeks 7 through 22, which acted as a confound for both independent variables. This campaign was itself responsible for significant increases in sales skill use for most skills.

Positive correlations between the use of sales skills and sales percentages were found. However, it is unclear whether the use of sales skills led to a sale or whether obtaining a sale led to the use of more skills.

The results of this study are consistent with the notion that feedback, in and of itself, is not a principle of behavior. Rather, information about performance is an environmental stimulus which may acquire antecedent, reinforcing, or punishing control depending on the learning history of the individual.

Acknowledgements

This project could not have been completed without the assistance of many individuals. I would first and foremost like to thank the thrift institution that allowed me to conduct this study with their telemarketing department. Many thanks go to the telemarketers, M., T., and D., who put up with the seemingly endless chore of taping their calls, and who sometimes said hello or told me a joke on the recorders. Mr. Parker Sims offered his support and enthusiasm, for which I am very grateful.

Ms. Lindsay Morrell gave generously of her time and resources. Had I not been able to rely on her consistency and good judgment, I would not have had the confidence to implement this study from a distance of nearly 400 miles away. An unexpected bonus was the development of our friendship as we worked together. The study is now completed, but I hope that our friendship will last for many years to come.

I am very grateful to the Learning Processes Subprogram faculty for their support of my unorthodox dissertation idea and of my graduate career in general. Dr. Nancy Hemmes, my dissertation sponsor, has given me countless hours of her time and much sound advice on this project and on others. Dr. Bruce Brown and Dr. Claire Poulson served on my dissertation committee. I would like to thank them for their guidance on all aspects of this study, and in particular, for their verbal reinforcers that kept me going in spite of the many hassles and delays. In addition, I would like to thank Dr. Richard O'Brien of Hofstra University and Dr. Walter Reichman of Baruch College/CUNY who served as readers of the dissertation, and Dr. Philip Ramsey who served as a statistical consultant.

Ms. Barbara Adams, Ms. Deborah Preiss, Ms. Abrah Mintz, Mr. Chris Tavella, and Ms. Suzanne Morrissey served as independent scorers of the telemarketing tapes. This critical task has to have been one of the least rewarding jobs on the face of the earth. Bless their good-natured souls.

Mr. Martin Sliwinsky volunteered his expertise, time, and computer on the statistical analyses, with an enthusiasm that still amazes me. He saved me countless hours and countless headaches, and I greatly appreciate his input.

A dissertation is the culmination of years of work, sacrifice, and a way of life. I would also like to thank some friends who, in the past nine years, have provided the emotional support and laughter that kept it all in perspective: David and Ellen Scarisbrick, Gail Kuslansky, Paul Carro, Debbie Davis-Lang, Martin Gipson, John St. Clair, Sarah Raskin, Rainy Gehles, Mary Elaine Moore, Jeff Cadey, Karleen McSherry, Howard Schneider, Eric Ries, and Patti Lou Watkins.

Heartfelt thanks go to Stephen Hetzel, who kept all the fires lit.

This work is dedicated with love and gratitude to my parents, Herbert and Mary Pratt. I am proud to be your daughter.

Table of Contents

List of Tables

List of Figures

Page

Feedback is a procedure that has been widely discussed and investigated in the psychological literature. While authors describe and debate its definition and mechanism (Ammons, 1956; Annett, 1969; Bilodeau, 1966; Duncan & Bruwelheide, 1985-86; Ford, 1984; Herold & Greller, 1977; Ilgen, Hobson, & Dugoni, 1981; Komaki, 1981; Larson, 1984; Locke, 1967,1980; Locke, Cartledge, & Koeppel, 1968; O'Brien, 1988; Peterson, 1982; Powers, 1973; Salmoni, Schmidt, & Walter, 1984) both laboratory and applied researchers report using feedback to investigate such diverse issues as conditional discrimination performance in pigeons (Peterson, Wheeler, & Trapold, 1980), motor skill learning in college students (Bennett & Simmons, 1984), and sanitary food preparation practices among cafeteria workers (Geller, Eason, Philips, & Pierson, 1980). Indeed, the investigations of feedback in the applied fields and knowledge of results in motor skills learning have been so extensive that several literature reviews have been published (Balcazar, Hopkins, & Suarez, 1985-86; Kopelman, 1982; Nadler, 1979; Prue & Fairbank, 1981; Salmoni et al, 1984) and the term "feedback," though inconsistently used and poorly defined, is commonly accepted in the behavioral literature. The term "feedback" is used in the applied literature, while the term "knowledge of results" is used in the literature on motor skills learning. The terms are synonymous (Duncan & Bruwelheide, 1985-86).

Feedback is defined as a stimulus that, at least on the first occasion, follows task performance and describes some aspect of that performance. This text will discuss the functional properties of feedback and the operational definition of goal-setting in an attempt to discover appropriate definitions or mechanisms of both. Next, by reviewing

the studies of informational feedback alone (without additional treatments), the function of feedback as a reinforcer or punisher can be isolated and examined. Next, because most applied researchers have combined feedback with a variety of other treatments, those studies will be reviewed last. Although the examination of these studies is not likely to reveal the nature of feedback because of confounding treatments, their review reveals the type of research typically conducted in the applied field and can suggest which combination of treatments will result in robust effects.

Functional Properties of Feedback

A "functional property" of a stimulus is the way in which the stimulus operates on behavior; i.e., as a reinforcer, punisher, or discriminative stimulus. Theories of behavior have guided the development of each of these concepts, each of which has a well-specified definition. While applied studies show efficacy of various interventions labeled "feedback," the lack of an agreed-upon theory of feedback makes its boundaries unclear. The term originally came from the field of cybernetics to mean the "... control of a machine on the basis of its actual performance rather than its expected performance" (Weiner, 1954, p. 24) and described a mechanism to maintain steady-state performance. Today, the term is more analogous to the concept of shaping, in that feedback is expected to produce a behavioral change (generally, an "improvement"). It may also be compared to the concept of chaining, in that the feedback stimulus is both the reinforcer for prior behavior and the antecedent stimulus for subsequent behavior (Huber, 1985).

Duncan and Bruwelheide (1985-85) have reviewed the formal definitions of feedback and have concluded that feedback is typically defined as either information,

reinforcement, or some combination of the two. Brethower (1972) defined feedback as "infomation about past performance which is used to guide future performance" (p. A-1). Connellan (1978) defined it as "information (communications) about present or past performance that is communicated to the individual employee, presumably in a way that will influence future performance" (p. 106). These definitions assume that information alone is sufficient to alter behavior. Many other researchers appear to believe that feedback alone is insufficient to change behavior because they combine it with external contingencies such as approval (e.g., Brown, Willis, & Reid, 1981), peer pressure via public posting (e.g., Fellner & Sulzer-Azaroff, 1984a), bonus payments (e.g., Harris, Bushell, Sherman, & Kane, 1975), and "thank you" letters (e.g., Runnion, Johnson, & McWhorter, 1978), to name a few.

Stoerzinger, Johnston, Pisor, and Monroe (1978) found that feedback given to salvage operation workers to increase production was initially effective for several weeks, after which performance returned to baseline levels. An initial impact of feedback on behavior that subsides after a number of presentations may be explained by a conditioned reinforcement and extinction model. In this explanation, feedback is a nonreinforcing stimulus presented following performance of a task. An individual who initially shows a feedback effect (the target behavior increases) may have a learning history in which feedback (information about performance) had been paired with either positive or negative reinforcers. Feedback represents conditioned reinforcement that acts to increase performance. However, if feedback for the targeted behavior is not paired with more potent reinforcers, its effectiveness will diminish. In other words,

feedback as a conditioned reinforcer undergoes extinction. For example, a rat can be trained to press a lever in the presence of a light that is paired with a food reinforcer. The light itself becomes a conditioned reinforcer and controls responding. In a new context, the light can initially function as a reinforcer, but if it is not paired with reinforcement, it will undergo extinction and cease to control responding. An individual with a history of feedback paired with consequences can be predicted to be reinforced, at least initially, by positive feedback when it is presented in a new context. If the individual learns, over repeated feedback presentations, that there are no consequences associated with feedback, then levels of behavior that can initially be maintained with feedback will decrease (extinction).

Feedback may also be viewed as functioning as an antecedent stimulus. Although the information contained in the feedback pertains to prior performance, the time of the delivery of feedback will be prior to the next response. In this case, effective feedback may function as a discriminative stimulus for correct responding. A discriminative stimulus that is not reliably correlated with reinforcement eventually ceases to control behavior. An individual's behavior may be initially controlled by a discriminative function if he or she has a learning history in which information has led to reinforcement; however, feedback will eventually lose its discriminative function if it is not differentially paired with reinforcers.

Studies in applied settings have indicated that feedback is more effective when delivered as soon as possible after the response (Balcazar et al, 1985-86). This finding indicates that a conditioned reinforcement model may be appropriate. However, the

content of feedback may also influence its functional role as either an antecedent or a consequence. Tosti & Jackson (1981) have attempted to separate these two functons by delivering feedback of different content at different times. They define two feedback procedures. "Summative feedback" refers to positive reinforcement (e.g., praise) following performance. Any information that might be perceived as punishing, negative, or a request for change is eliminated. "Formative feedback" refers to instructions or information for improvement, may contain positive and negative evaluations, and is given just prior to the next response. To date, there are no published reports comparing formative and summative feedback; however, there is ongoing research in this area (Brewer & Gipson, 1989).

The use of techniques without theory testing led Peterson (1982) to argue that few researchers have provided a behavior analysis of their results in terms of basic principles. He states that feedback is not a new or different principle of behavior; in fact, the term is professional slang. Feedback (as information about performance) can serve any number of functions (e.g., reinforcer, punisher, discriminative stimulus) depending on the participant's learning history. He urges behavior analysts to use direct procedural descriptions of their independent variables (e.g., "performance graphs" instead of "feedback") and to provide behavior analyses. These steps will determine which variables produce maximum results without invoking new theories or intervening variables.

Feedback, like any other environmental stimulus, may acquire a discriminative function (set the occasion for responding), a reinforcing function (increase behavior), a punishing function (decrease behavior), or no function at all (no effect on behavior). It

may acquire a discriminative function if it precedes the behavior and is reliably paired with reinforcement following a response of a specified topography. It may acquire a reinforcing or punishing function if it follows behavior and becomes paired with either positive or negative consequences. The functional properties of feedback, like those of any other environmental stimulus, will depend on the timing of the presentation (before or after the task) and its effect on behavior (increase, decrease, or no effect). In many cases, feedback may acquire antecedent and consequence functions simultaneously, and these functions may not be separable.

Operational Definition of Goal-Setting

In contrast to feedback, goal-setting is a distinct, though related, procedure. Nevertheless, feedback and goal-setting have become so intertwined in the literature that it is practically impossible to discuss one without invoking the other. Goals have been described procedurally as "stimuli that precede behavior. When the antecedent goal reliably accompanies a reinforced response it acquires 'discriminative control,' increasing the probability that it will cue the individual to repeat the behavior" (Fellner & Sulzer-Azaroff, 1984b, p. 34). Therefore goals, like feedback, can acquire a range of functions for the individual.

These principles can be illustrated with a study by Becker (1978), who compared the effects of goal-setting to those of goal-setting plus feedback on residential electricity consumption. The study was conducted in a planned community of approximately 3,000 homes in central New Jersey. Of a pool of 175 randomly selected families, 100 agreed to participate.

Participants were assigned to one of four treatment conditions in this 2 x 2 factorial design. The first independent variable, goal-setting, had two levels, "easy" and "hard." "Easy" goals were operationalized as a daily conservation goal of 2%, based on data collected the preceding year in the same community. "Hard" goals were operationalized as a daily conservation goal of 20%. The second independent variable was the presence or absence of feedback. Participants were informed of their goals during a visit from a research assistant, during which feedback was also explained to those families who were to receive it. Feedback was supplied by the researchers who placed a plastic pocket on the outside of a window of the participant's home. Feedback cards could be dropped into the pocket from the outside so that they could be read from inside the house. The cards showed a graph from the daily meter readings of "percent conserved" or "percent wasted" which could be added to over time. Depending on the goal group, a dark line at either 2% or 20% was drawn across the graph. Daily meter reading and feedback continued for approximately one month. The results showed that the 20%-goal plus feedback group was the only group that used significantly fewer kilowatt-hours (KWH) than the control group. Becker concluded that "motivational effects" typically attributed to either goal-setting or feedback are really due to their joint effect. Since the effects of feedback alone (without any goal-setting) were not tested, however, that claim is questionable. That is, from these data, goal-setting may be a necessary but not sufficient condition of behavior change, but from this particular design, it is unknown whether feedback alone would have been sufficient. It is clear that goal-setting without feedback (which presumably acted as the reinforcer) did not change behavior. In addition, a 2%

conservation goal may not have been a potent antecedent to electricity conservation because conservation behaviors associated with a mere 2% reduction in electricity usage may be unclear. The 20% goal may have been a more potent antecedent because virtually all electricity usage must be curtailed to achieve it.

For Locke and his colleagues (Locke, Shaw, Saari, & Latham, 1981; Latham & Baldes, 1975; Locke et al, 1968; Locke, Feren, McCaleb, Shaw, & Denny, 1980) goal-setting is a conscious intention that regulates action. Behavior depends partially on the individual's emotions; for example, whether the goal is accepted by the individual. In Locke's view, setting the goal impels the individual to feel motivated to act, and feedback is useful only in helping to set goals.

Locke (1967) conducted a 2 x 2 factorial study with two levels of feedback (presence or absence) and two levels of goal-setting (specific hard goals versus "do your best" goals) while participants worked on an addition task. His results yielded no effect for feedback, but a significant effect for goals, with hard specific goals being superior. Locke surmised from these results that incentives (feedback, money, praise) were effective only to the degree to which they affect the participants' goals, which can be self-generated or assigned. Goal-setting and goal evaluation are mediating variables between the performance and the incentive (feedback, reward.) The form of the incentive can be potentially important because it might affect the level at which the participant sets his or her goal (Locke et al, 1968). Nevertheless, in Locke's view, it is the goal-setting (and specifically, the setting of hard goals) and not the feedback, that is the primary predictor of future behavior.

Huber (1985) has pointed out that Locke's theories, while intuitively appealing, are impossible to test because cognitive intentions and goal-setting cannot be measured directly. Goals must be inferred from self-report by either (a) assigning goals and asking partcipants whether they accepted them; (b) giving participants a choice of goals from which to choose; or (c) asking participants to set their own goals. The notion that any of these processes mediate the relation between environmental events and behavior is speculative since none are directly available to measure or test. The behavioral model also explains why goals do not necessarily affect behavior. If responses have been met with reinforcement, goals will acquire a discriminative function. Goals that are too hard will be rejected because the required behavior is too difficult to perform and no reward will follow. Goals that are nonspecific (e.g., "do your best") do not consistently meet with reinforcement because the individual's best may not be the level that is required for reinforcement. In addition, nonspecific goals do not provide specific instructions, which if followed, could be met with reinforcement. Therefore, while it is empirically correct to say that specific goals are better than nonspecific goals (Locke, 1967; Locke et al. 1968; Latham & Yukl, 1976; Latham & Baldes, 1975), the best explanation for that statement comes from environmental contigencies that can be measured and manipulated.

A second, critical notion in Locke's theory is that people evaluate a goal and decide whether to accept it. People can, and do, change their minds and appear to be acting like free agents. The behavioral model states that the variables controlling the changes in behavior can be identified in the environment, and therefore, if enough information is known about those variables and the person's learning history, that person's behavior can

be predicted without reference to the mind. This parsimonious explanation does not resort to unobservable, unmeasurable variables that are, by definition, outside the realm of science. As O'Brien (1988) pointed out, the "free will" argument implies that people are unpredictable, and therefore, no empirical relationships can be demonstrated. An "explanation" that does not allow for prediction is theoretically barren and retreats to whimsy (O'Brien, 1988).

Informational Feedback

Feedback and goal-setting have thus been defined as environmental stimuli which may acquire consequential and antecedent control over behavior. It may now be asked whether feedback that is strictly information about performance can acquire those functions. Informational feedback will be defined as information on task performance without additional treatments. Informational feedback may be either process feedback, which is information about the task itself (e.g., the number of times a sales skill is used in a sales presentation), or outcome feedback, which is information about the task result (e.g., whether a sale resulted from a sales presentation). The relative efficacy of process versus outcome feedback has received some attention (e.g., Johnson & Frederiksen, 1983); however, most research uses one or the other type without a direct comparison of the two.

The next three sections will describe three experimental conditions in which informational feedback without other contingencies was implemented. In the first section, informational feedback comprises the data that participants collect on some aspect of their own behavior, known as self-monitoring. In this case, the goal was to collect steady

baseline indices of performance. In contrast with most uses of informational feedback in which change is expected and desirable, behavioral change as a result of self-monitoring data collection is undesirable. The next section will describe cases in which informational feedback was ineffective in changing behavior. The last section describes cases of individual informational feedback that were effective as reinforcers and as punishers.

Self-Monitoring

Self-monitoring is a form of informational feedback because participants collect information about some aspect of their own behavior (e.g., number of cigarettes smoked), which fulfills the definition on page 10 of informational feedback. The literature on self-monitoring has provided evidence that feedback not paired with reinforcement will cause the target behavior to extinguish. In some cases, but not all, self-monitoring has been shown to produce changes in the target behavior without an additional intervention, a phenomenon known as reactive measurement (Kazdin, 1974). However, when behavior does change due to self-monitoring, the change is often temporary, with the behavior returning to baseline levels over time (Broden, Hall, & Mitts, 1971; Fixsen, Phillips, & Wolf 1972; Mahoney, 1974; Stuart, 1971). Kazdin (1974) investigated reactive self-monitoring in a series of laboratory experiments in which participants were given a sentence construction task. Self-monitoring alone increased participants' use of targeted pronouns, but in a feedback manipulation, participants who observed their own record of self-monitored responses performed the target response more frequently than participants who did not receive feedback. Kazdin suggested that these results indicate that self-monitoring and feedback contribute separately to behavior change. However, it is not

known whether the feedback manipulation effects would have extinguished without additional intervention (e.g., reinforcement). Kazdin stated that "self-monitoring enables the individual to discriminate the occurrence of the response to a greater extent than .... vague self-reports and casual observation" (p. 705). In fact, he postulates a "feedback loop" in which the individual adjusts his or her responses until a standard of performance is reached. However, the feedback manipulation in his experiment may have been simply an augmentation of the information available to the participant as a result of self-monitoring. Separate constructs are not needed.

Calpin, Edelstein, and Redmon (1988) asked clinicians in a rural mental health center to self-monitor the number of hours spent in direct client contact and later compared this condition to a self-monitoring plus assigned goals condition. Self-monitoring was considered self-generated feedback. The results indicated that most of the increases in the target behavior occurred during the self-monitoring phase and that the addition of goals produced only small additional increases. Whether improvements would have attenuated had the three-week self-monitoring phase been lengthened is unknown.

Winett, Neale, and Grier (1979) investigated daily feedback in comparison to self-monitoring for the reduction of household electricity use. Using households with large (> $200) monthly electric bills and high KWH usage (mean = 170 KWH), volunteer participants were randomly placed into one of four conditions: a daily written feedback group, a self-monitoring group, a control group of volunteers who did not receive treatment, and a group who agreed only to have their meters read. The self-monitoring

and feedback groups were both taught to read their meters in order to equate the groups on that dimension. The self-monitoring group received four weekly meter reading recording forms, which contained meter dials on which participants were to mark the position of the dials and to interpret the reading. A space was also provided to calculate the difference from the prior day's readings, and a graph on which to plot their daily KWH usage and the reduction goal in KWH chosen the first day. For the next consecutive 28 days, participants in this group received a note at their door that showed their expected use for the prior day in percentage terms. Expected use was based on a weather correction system that compared the control groups' prior day's KWH use divided by their baseline KWH average. The participants were then instructed to determine if their prior day's use had been above or below expected use.

The feedback group received a sheet at the door each day (no personal contact) with an ascending series of smiles or frowns that corresponded to the percentage decreases or increases in electricity consumption. The prior day's KWH consumption, its percentage increase or decrease from baseline (with the weather correction), the relationship of the decrease to a reduction goal chosen at the first meeting, and an estimate of the household's monthly electricity bill in dollars, based on its prior day's use, were all included on the feedback form.

All groups were equal in KWH use during baseline. Because the comparison groups never differed from each other at any phase of the study, their data were combined. During intervention, both the feedback and self-monitoring groups were consistently and significantly below the comparison group. The feedback group reduced consumption by

approximately 13% and the self-monitoring group by approximately 7%, also significantly different from each other. In kilowatt hours and dollars saved, the feedback group saved 19 KWH per day, 570 KWH per month, and about $23 per month (based on $.04 per KWH). The self-monitoring group saved 9 KWH per day, 270 KWH per month, and about $11 per month. This study adds support to the notion that feedback and self-monitoring are essentially two sides of the same coin. Both produced significant changes in electricity consumption. The feedback condition may have been more effective because there was less work involved for the participants (they did not have to read their meters), because more important information (money saved) was provided, or because there was a greater sense of being monitored by an outside agency.

Cases in Which Feedback Has Been Ineffective

In addition to the evidence from studies on self-monitoring in which feedback alone was insufficient to maintain a change in behavior, there are published reports in which feedback delivered by other agents is reported to be ineffective.

Winett and his colleagues tested a variety of techniques in combination with feedback in order to isolate the most effective treatment with the lowest cost (Winett, Kagel, Battalio, & Winkler, 1978). Volunteer households were assigned to one of five conditions: a high monetary rebate condition in which participants received conservation information, weekly written feedback on their electricity use, and monetary rebates amounting to a 240% price change in electricity ($.30 for each 1% reduction in weekly KWH used as compared to the previous summer's use); a low monetary rebate condition with the same structure as the high rebates except payments amounted to a 50% price

change; a weekly feedback condition in which participants received conservation information but no rebates; a conservation information-only condition; and a control condition. Meter readings occurred weekly for 4 weeks. The dependent variable was the percentage of reduction in electricity use. Only in the high rebate condition did participants significantly lower their electricity consumption, by 12%. Feedback without reinforcement was not effective. The authors supplied feedback once a week (mailed to the homeowner), compared current energy consumption to that of the prior year during the same time period, and used a standardized form letter. Any participants who were aware of neighbors' monetary rebates may have experienced resentful demoralization (Cook & Campbell, 1978, p. 55) which would have altered the way subjects responded to the treatment.

Ford (1984) compared the effectiveness of supervisory feedback to videotaped feedback and a combination of videotaped and supervisory feedback with teaching staff at a facility for the mentally retarded. The three participants were responsible for teaching self-help skills to severely and profoundly retarded adults. A modified multiple baseline across dining, dressing, and bathing skills was implemented so that each participant received each type of feedback in a different order from the other two participants. Participants were videotaped performing the three targeted skills with clients and their performance was scored on a Performance Rating Form. Supervisory feedback consisted of reviewing the form with the participants after each weekly taped session. Videotaped feedback consisted of reviewing the tape of the session without further discussion. Participants had the option of rating themselves on the Performance Rating Form, but

these were not collected or discussed. In the combined supervisory and videotaped feedback condition, participants watched the tape with the supervisor who reviewed each item of the rating form and used portions of the tape to highlight specific points. This condition resulted in the most rapid improvement and the highest sustained level of performance across all participants. The other types of feedback also resulted in gradual improvements but of a lesser magnitude. Videotaped feedback alone, or information, produced the least performance improvement, perhaps because the participant could not distinguish which behaviors to alter (inadequate stimulus control). Supervisory feedback was more effective, possibly due to more direct reinforcement, but with less direct information than can be supplied with videotape. The combined methods provided the strongest antecedent control (the most information) with the most reinforcement (supervisor contact).

Krumhus and Malott (1980) compared immediate versus delayed feedback to modeling and instructions in staff training of tutors in a remedial education program. College student volunteers were trained to use descriptive social reinforcement (vocal praise that names the response that the tutor is attempting to teach.) The dependent variable was the percentage of use of this technique per opportunity (following the completion of a task or a corrected error.) The experiment followed four conditions. The first condition, instructions, consisted of a handout and lecture. The modeling condition consisted of audiotapes of other tutors. The participant-tutors listened to these tapes and selected three instances of descriptive social interaction and three noninstances. One-half of the modeling sessions occurred immediately before the tutorial session and one-half

occurred immediately after. The third condition was feedback before the tutorial session versus feedback after the tutorial session. Feedback was provided on appropriate and inappropriate uses of descriptive social reinforcement. Three instances and three noninstances were provided on the tutor's most recent session; however, there were a few exceptions when the tutor failed to emit either three instances or noninstances. No graphic or quantitative information was supplied. The eight particpants were randomly assigned to two groups that determined the order of the conditions. The first group's order was baseline, presession modeling, presession feedback, and postsession feedback. The second group's order was baseline, postsession modeling, postsession feedback, and presession feedback. All participants were evaluated with a multiple baseline across subjects design. The results suggested that (1) modeling was very effective as a training procedure; (2) feedback following the modeling procedure produced slight improvements in performance; and (3) immediate feedback was no more effective than delayed (presession) feedback. The authors suggested that modeling was so effective that the added benefits of feedback might be negligable compared to costs when designing new training procedures. However, because modeling was always tested first, it is not known what the effects of feedback before or without modeling would have been. There may have been a ceiling effect in this case that would not exist in other cases. A direct comparison of the two procedures continues to be warranted.

Stoerzinger et al. (1978) increased the sorting and processing performance of six employees of a company that salvages donated household goods for resale. The dependent variable was the number of items processed per hour on the job each day. All six

employees were treated as a unit in an A-B design. Feedback consisted of a daily production graph for the group which depicted a daily goal and whether it was achieved. Performance increased by approximately 42% for three to four months after treatment began, which produced an average of $3600 worth of goods processed per week as opposed to $2500 per week during baseline. Performance declined to baseline levels in the last five weeks of the project. The authors speculated that high temperatures in the plant, employee vacations, and a decrease in verbal reinforcement from the plant supervisor contributed to this reversal. Feedback, in and of itself, was not a sustained functional reinforcer. Once again, one can note that informational feedback without further reinforcing contingencies did not maintain behavior.

Real estate agents were induced to increase the number of client contacts with the use of self-monitoring, anonymously coded but posted feedback, and token reinforcement (Anderson, Crowell, Sucec, Gilligan, & Wikoff, 1983). Sixteen real estate agents self-monitored their initial and follow-up client contacts for 20 weeks of baseline. These data were coded to ensure anonymity and were displayed without comment for each individual in this time period. In addition to this process feedback, outcome feedback of the number of pending sales, actual sales, pending listings, and actual listings were posted. The 15-week intervention consisted of a token reinforcement program to be exchanged for items ranging from a tank of gasoline to a pool table. Other experimental conditions consisted of anticipated withdrawal, in which the end of the program was announced 3 weeks prior to the withdrawal phase, and a reversal phase in which the token program was reinstituted.

Trends and levels for all measures changed with the various phases. Follow-up contacts remained fairly stable in "baseline" (self-monitoring and feedback), whereas initial contacts showed a gradual decline in this period. Both measures increased markedly in the token reinforcement phase, declined starting in the anticipated withdrawal phase, and fell below "baseline" levels in the withdrawal phase. Contacts returned to nearly intervention levels in the reversal phase. This study provides further evidence that feedback without additional contingencies will not have a sustained effect on behavior. Feedback was posted but coded so that individuals could compare his or her information to anonymous others. This information alone was not sufficient to maintain a high rate of client contacts. Whether this rate was higher than rates before self-monitoring and posting is unknown.

Informational Feedback Effective for Individuals

This section explores the effects of informational feedback, functioning either as reinforcers and punishers, delivered to individuals. In individualized feedback, each participant receives information about his or her own behavior. It is also possible to deliver feedback to groups (discussed in the next section), but a different set of contingincies operates because group behavior is described regardless of what the individual does. The current section will explore the contingincies operating on the individual.

Feedback as a reinforcer. In the following studies, informational feedback proved to be an effective stimulus for improving performance. Although feedback was provided without additional goals or contingencies, there may have been other inherent forms of

reinforcement built into the feedback. In some cases, there was direct contact with the researchers who may have provided praise or approval. At minimum, hypothesis-guessing and social desirability subject bias may have operated in cases with direct researcher contact. In other cases, the feedback described an amount of money saved by the participant, which is likely to be a potent reinforcer for most people. Finally, in some cases, the content of feedback focused more on the undesirable behavior than on the desirable behavior, although the outcome was an increase in the desired behavior. For example, feedback about the number of unsanitary kitchen habits observed is a direct description of an undesirable behavior, and when posted, handwashing (the desired behavior) increased. Participants in studies that explicitly described undesirable practices may have been motivated by negative reinforcement (avoidance of embarrassment, social undesirability) to improve. Feedback about the number of sales skills used focuses on desirable behavior and only implies undesirable behavior (lack of skill use) if results are low. Nevertheless, it is unclear why any of these variables should have operated in these studies and not in the studies described earlier in which feedback was found to be ineffective.

Leitenberg, Agras, Thompson, and Wright (1968) used reversal designs to demonstrate the efficacy of feedback in two phobic patients. A claustrophobic patient was given feedback about how long she was able to stay in a closet per trial and a knife-phobic patient was given feedback about how long she was able to keep a knife exposed per trial. The progress of both patients was impeded when feedback on these times was withdrawn. With the knife-phobic patient, the addition and removal of contingent verbal praise did not

alter the rate of progress. There were no additional therapies in either case. Feedback was not contingent on improvement, but was presented after both increased and decreased exposures. However, it can be noted that there were only 3 decreases in exposures of 22 feedback trials for one patient and 4 decreases in approximately 20 feedback trials for the second patient. Therefore, most of the feedback would have been "good news." The authors suggest that feedback allowed the patients to detect slight but successful changes that would have otherwise gone unnoticed. Both patients were cured of their phobias and were able to function normally.

Brown, Malott, Dillon, and Keeps (1980) compared sales training to feedback for increasing four sales behaviors of department store salespeople. These behaviors were approaching customers, greeting them, being courteous, and appropriately closing the sale. On-site observers recorded the frequency of these behaviors for approximately 6 hours per week. After baseline observations were collected, a 4-hour sales training program was provided on the dependent variables. The program consisted of videotaped demonstrations, pencil and paper worksheets, and behavioral rehearsal. Data were collected for 4 weeks following sales training. In the feedback condition that followed, salespersons were approached individually at the beginning of the observational session and showed a graph of their performance for all previous sessions. At the end of the session, an updated graph was shown for that day's performance.

All four sales behaviors were summed. The percentage of these totalled sales behaviors across the four study conditions were as follows: baseline, 49%; training, 59%; feedback, 85%; second baseline, 70%. Based on the 35% improvement for

feedback, versus a 9.6% improvement for training, the authors then estimated a cost-benefit analysis extended to the store's 264 store employees. The authors assumed that sales training would be unnecessary if written instructions and feedback were provided, and that overall results would approximate the current results. Although untested, customer service training costs were estimated to be $5,942, compared to $1,637 for a feedback system with training provided to management.

Feedback graphs have also shown efficacy with mentally retarded workers in a sheltered workshop who assembled hinges for boxes and assembled plastic rings (Jens & Shores, 1969). In a reversal design, feedback graphs were tested with each task for three study participants. Experimenters were careful not to make any evaluative comments after production. Graphs significantly accelerated the assembly of the rings, but not the hinges, although trends were in the desired direction. The researchers noted the evaluative comments made by the workers after reviewing their graphs ("I did great") and speculated that the graphs acted as conditioned reinforcers for performance.

Sulzer-Azaroff (1978) used a multiple baseline across subjects design to test written feedback and safety suggestions in 30 university research laboratories. Initial building inspections indicated the presence of many mechanical, chemical, and electrical hazards, which were noted on an observation sheet. Feedback was provided after five months (12 laboratories), nine months (11 laboratories), and twelve months (7 laboratories) of baseline. Feedback consisted of a copy of the observation form that indicated the location and nature of the hazards, along with a set of corrective suggestions. Twenty of the laboratories showed substantial improvement, eight remained unchanged,

and two worsened. The authors noted that of the ten that did not improve, some may have ordered safety supplies (e.g., fire extinguishers) that took several months to arrive. More important, perhaps, was the lack of consequences for compliance or noncompliance. Some laboratory supervisors argued that compliance to safety measures would make their jobs more difficult. These supervisors clearly did not respond to the information about hazards provided by feedback. Instead, they responded to the contingencies of the task. To be effective with these supervisors, feedback would have to be paired with other reinforcers or punishers strong enough to outweigh the punishing aspects of safety compliance.

Sulzer-Azaroff and de Santamaria (1980) conducted follow-up work to the study described above with six production supervisors in a small industrial organization. Observational recording and the form of feedback were similar to the Sulzer-Azaroff (1978) study, with the addition of congratulatory comments when safety hazards decreased. Inspections and feedback occurred twice a week. Substantial improvements were demonstrated in all six departments.

Hayes and Cone (1981) used a reversal (ABA) design to test the efficacy of monthly feedback on electricity consumption. Their design was further enhanced by the addition of a no-feedback control group. The efficacy of feedback for electricity conservation had already been demonstrated by the time this study was conducted. Hayes and Cone wanted to demonstrate that a monthly feedback form letter, consistent with both the timing and form of customers' electric bills, would be more effective, practical, and economical for intervention by electric companies.

Forty nonvolunteer consumers from Pawtucket, Rhode Island served as participants. Consumers were randomly divided into treatment and control conditions. The control group participants were never contacted, but their electricity information was given to the experimenters by the electric company. After two year-long baseline periods during which electricity information was collected for all participants, the treatment group participants received a form letter from their electric company each month for 4 months. The letters were sent one week following the meter reading, separate from the bill. The letters reported the percent change in consumption over the same month during the previous year (in the baseline phase) and showed the number of KWHs and actual dollar amounts involved. After the 4 months the treatment ended, returning to 3 months of baseline.

The results were analyzed in terms of a percent change score from the baseline year to the treatment year. The group that received feedback consumed 4.7% less electricity as compared to its baseline year, while the control group consumed 2.3% more. When feedback was withdrawn in the third phase of the design, the feedback group reversed its consumption pattern and used 11.3% more electricity than during the comparable period one year earlier. The control group used slightly less (.3%) than it had. Analysis of variance revealed that there was no significant effect for either group or phase, but that there was a significant group by phase interaction: the feedback group had lower consumption in the feedback phase than the control group, but the groups were equal in the baseline phase.

Hayes and Cone suggested that the results obtained in this study might not generalize

to other samples. In the old established community of Pawtucket, oil and gas were primarily used for heating, as opposed to electricity; therefore, electricity consumption rates were approximately half that reported in other studies. Feedback might then have been operating on other types of electricity consumption. Exactly what people modify when they reduce their electricity consumption is not known. Therefore, while monthly form letter feedback showed efficacy in this sample, the external validity was questionable and replication in other towns was recommended.

Seligman and Darley (1977) conducted a simple between-groups design to compare daily feedback of electricity consumption to a no-treatment control group. The setting was the same as used by Becker (1978). Forty homeowners were recruited to participate and were randomly assigned to conditions. Daily feedback was provided on a Lucite display attached to a window in which small plastic numbers could be inserted by a research assistant and read from inside the house. The numbers depicted the percentage of actual electricity consumption compared to predicted consumption based on baseline rates. For example, if the homeowner's predicted consumption was 10 KWH and he used 8 KWH, the display would read 80%. Feedback was displayed to homeowners four times a week for 1 month. No goals or standards were provided. The results indicated that while the two groups did not differ at baseline, the feedback group used 10.5% less electricity than the control group after treatment. This difference was significant at the .04 level.

The sanitation practices of nine kitchen workers in a university cafeteria were targeted for improvement by Geller, et al. (1980). The use of an A-B-A-C-A-D-A design compared the components of videotaped handwatching (in which participants' hands

and actions were videotaped as a control for the Hawthorne effect, phase B), sanitation training (phase C), and an individualized feedback checklist (phase D) on handwashing. In the feedback phase, individuals were shown checklists of the number and types of "critical antecedents" that should always be followed by handwashing. These critical antecedents included behaviors such as touching hair, face, floor, cleaning equipment, or handling unwashed food. The checklists indicated the number and type of these antecedents that had been performed the previous day as well as the percentage that had been followed by handwashing. Checklists were always presented in a positive manner. Videotaping and training had little effect on handwashing, but feedback had a prominent and consistent effect.

Feedback as an effective punisher. If feedback suppresses or decreases behavior, it is defined as a punisher. The following three studies indicate that feedback can be a punishing stimulus as well as a reinforcer.

Arkes, Christensen, Lai, and Blumer (1987) have stated that people are notoriously overconfident about their answers to general knowledge questions and that serious consequences can result when incorrect decisions are made based on these judgments. Feedback to reduce the overconfidence of participants who were answering general knowledge questions in a laboratory setting was tested in a 2 x 2 factorial design. Participants answered two sets of questions of equal difficulty, but which appeared to be either easy questions or hard questions. After each question, participants were asked to estimate their confidence of having chosen the correct answer on a scale from 50% to 100%. Feedback was provided to one-half of both the "easy" and the "hard" groups. Then

all the participants were asked to answer 30 final questions. Results indicated that there were no significant differences in either the number of correct answers to the first five questions or to the final 30 questions in either group, indicating that the degree of difficulty for all questions was approximately equal. However, the degree of confidence displayed was affected both by the perceived level of difficulty as well as by feedback. "Easy" questions elicited more confidence than "hard" questions. Participants without feedback were overconfident about their percentage of correct answers, estimating a larger number of them to be correct than were. Feedback resulted in underconfidence, or estimating a lower number of correct answers than were correct. In sum, the simple method of telling the participants the correctness of their answers reduced (punished) their reported confidence levels.

Supervisory verbal feedback was found to be an effective punisher of off-task behavior, but ineffective as a reinforcer for client care, in changing the behavior of direct care staff in an institution for the mentally retarded. Brown et al (1981) attempted to increase social interaction with clients and decrease off-task behavior of the staff by providing verbal feedback of staff behavior after hourly time sampling data collection. No approving or disapproving remarks were made. In a second condition, approval statements were provided in addition to the feedback. A multiple baseline design across staff and a reversal design to evaluate the second condition were implemented. Verbal feedback alone resulted in great reductions in off-task behavior (means = 25% of off-task behavior to 9% and means = 23% to 2% in morning and afternoon shifts). However, verbal feedback alone did not result in increases in social interaction with the clients.

Changes were temporary for one shift and nonexistant for the other shift. When approval statements were added, social interactions with clients in the afternoon shift increased from a mean of approximately 20% to 37%. The reversal phase demonstrated a return to baseline levels of 13%, which rose to 35% when approval statements were reinstated. In the morning shift (A B BC design), social interactions with clients went from a mean of 5% in baseline to a mean of 8% in the feedback phase to a mean of 14% in the feedback plus approval phase. The authors note that feedback alone was an effective punisher of off-task behavior and suggest that it may have been a punisher of the behavior targeted for increase as well, since little or no increase occurred. An alternative explanation is that information about off-task behavior was clearly helpful, because staff can recognize that any level of off-task behavior will be viewed as undesirable by supervisors. However, information about the amount of social interaction might not be as useful since staff might not know the optimal levels of social interaction, or they might not care. If clients themselves do not respond to social initiations (as is often the case with moderately and severely handicapped) and there is no reinforcement from supervisors, information about social interactions might be seen as useless until put into a larger context (paired with supervisor praise.)

Johnson and Frederiksen (1984) compared process feedback to outcome feedback with staff at a state mental health institution for geriatric patients. One important patient program was "Reality Orientation," which consisted of classes designed to orient patients to current events and correct perceptions of time and place. These classes were conducted by direct care nursing staff, who served as the participants in this study. The dependent

measures (and feedback) consisted of the daily total number of nursing staff contacts with patients in Reality Orientation classes (process feedback) and the number of correct, or oriented patient responses to questions posed by staff members (outcome feedback). Five treatment units participated in the study. Two units received process feedback in a multiple baseline across units design; two received outcome feedback in the same manner, and one unit served as a comparison group with no intervention. The graphed feedback was explained to each unit after baseline data were collected. Process feedback resulted in substantial increases in the rate of process behavior (patient contact hours) but had no effect on the intended outcome (higher orientation scores for patients). Outcome feedback had no effect on either patient contact hours or patient orientation scores. The comparison group's dependent measures remained constant across the course of the study. Because outcome feedback did not affect either dependent variable, one might speculate that information about task results is not a potent conditioned reinforcer. However, it must be noted that the increased patient contact in the "Reality Orientation" classes of the process feedback group did not increase reality orientation scores, indicating that "Reality Orientation" itself was an ineffective treatment. One would expect outcome feedback about decreasing patient scores to be punishing, and in fact, the number of patient contact hours actually declined for one unit in this condition, while it remained stable for the other unit. Staff could have easily interpreted their outcome feedback to mean "My efforts with the patients doesn't do any good." Hence, patient contact hours were punished and decreased.

Feedback Plus Other Independent Variables

Feedback has been combined with a host of other independent variables. The first two variables reviewed here will be posted, individualized feedback and posted, group feedback. Both types of feedback provide a social contingency (which can be positive or negative) by allowing feedback to be viewed by all members of the group. In the former case, reinforcement or punishment from other group members will have the greatest impact because each individual is compared to the others and scores can easily be ranked from "best" to "worst." Thus, even without back-up reinforcers, compliance to group norms can become a controlling factor. In the latter case, specific individual contributions to group scores is lost, which presumably weakens the reinforcing or punishing effect on the individual's behavior.

Feedback has also been used in an array of studies combined with other treatments (feedback "packages"). All independent variables are applied and tested as a unit. Because of the nature of field research, these research designs are often less than ideal; much internal validity is lost. Nevertheless, field studies often have high external and social validity and can contribute to a technology of behavior analysis.

Individualized Posted Feedback

In two experiments, Van Houten, Hill, and Parsons (1975) tested the effects of timing in-class work assignments, feedback on the amount of work accomplished, public posting, and teacher praise on the academic performance of fourth and fifth graders. In the first experiment with two classes of fourth graders, the authors employed an interaction design (A B BC B BC BCD BC) with B representing timing and immediate

feedback, that is, asking to students to count the number of words written on a composition task during a finite period of time; C representing the addition of public posting of each child's score; and D representing teacher praise to individual children for improvement. Timing plus immediate feedback nearly doubled the baseline rates of mean words written per minute. Public posting increased those rates (from 8.3 words per minute to 10.8 in Class A and from 3.5 words per minute to 5.4 in Class B). The subsequent removal of public posting reduced rates to the approximate level of the timing and feedback (B) condition, and reinstatement of public posting yielded an increase similar to the first public posting phase. The introduction of praise (BCD phase) increased the rate of Class A to 13.0 words per minute but had no additional effect on Class B.

In the second experiment, fifth graders completed reading comprehension and word meaning tasks in order to test an experimental package of all of the techniques used in Experiment 1. The design was a two-legged multiple baseline across tasks with a reversal. Results indicated that the number of comprehension problems worked correctly nearly doubled and the number of word-meaning problems correctly worked increased by 50% after the package was introduced. At reversal, the comprehension correct rate decreased to baseline levels and the word-meaning correct rate also declined. Taken together, these experiments demonstrated that a package of timing, individual self-generated feedback, public posting, and praise can be effective in improving children's academic skills. In Experiment 1, the elements of the package also showed differential effects, with public posting being the most effective.

Mental health technicians in an adult psychiatric facility received posted feedback on

whether they completed group therapy sessions with patients, completed individual therapy with patients, and completed assigned daily routine duties (Kreitner, Reif, & Morris, 1977). The intervention consisted of weekly posted interoffice memos, one for each job duty, that listed each technician's name the frequency of performance. Data were added to the posted memos as the weeks went by so that all previous performance data could be seen. The results indicated that for each targeted job duty, there was a marked improvement each time feedback was implemented.

Welsch, Ludwig, Radkier, and Krapfl (1973) used posted feedback in two separate A-B designs to increase the use behavioral techniques by staff in two wards of a state hospital. In the first design, the percentage of projects completed was posted daily for the entire ward. This improved the average percentage from 61.6% in baseline to 97.7% during feedback. In the second ward, several months after the first study was completed, a similar procedure added staff members' names to the chart along with their assigned projects. The staff members checked off the project when it was completed. This technique improved staff performance from 28.3% completed projects in baseline to 83.3% during feedback.

Maher (1982) used a two-legged multiple baseline across schools to test the efficacy of posted feedback on the completion of instructional programs by teachers of handicapped children. The two dependent variables consisted of the percentage of instructional programs completed weekly and the weekly percentage of teachers who recorded information for the programs. The school principal reviewed the instructional programs each week and compiled a publicly posted list of teachers' names and whether they had

performed either of the two behaviors for that week. After the intervention, completion of instructional programs rose from 8.8% to 82.1% in School A and from 17.4% to 93.3% in School B. The percentage of teachers who evaluated the programs weekly rose from 9.5% to 95.2% in School A and from 20.8% to 95.8% in School B.

Feedback to reduce charting errors in a clinic setting was tested by Frederiksen, Richter, Johnson, and Solomon (1981/82). Four types of charting errors were tracked throughout the study. Public feedback was provided on two types of errors in an A-B-A design. The results indicated that feedback worked specifically for the targeted errors and did not generalize to the other types of errors. Feedback reduced the other two types of errors in a second A-B replication. This study demonstrated that the effects of feedback are specific and will not necessarily generalize to other behaviors within the same class.

Newby and Robinson (1983) compared three interventions, group feedback, individual feedback, and token reinforcement, to baseline and withdrawal phases for drug store clerks on cashier precision, punctuality, and money check-out proficiency. In the initial phase after baseline, graphs of group performance were posted for each target behavior. In the next phase, individuals' names replaced group data. After a withdrawal phase, individual feedback was reinstated with the addition of a token reinforcement program. Grouped feedback did not result in any improvement over baseline, but individual feedback resulted in substantial improvements in all target behaviors, presumably due to social contingencies (praise, embarrassment). With the addition of token reinforcement, behaviors in all categories increased to nearly ceiling levels. The efficacy of the token program without feedback is not known. However, the conditioned

reinforcement hypothesis would suggest that individualized feedback without additional contingencies (tokens, social reinforcement or social punishment) would eventually cease to control behavior.

Ralis and O'Brien (1986) increased waiters' suggestions to customers that they purchase wine in a restaurant with prompts, goal-setting, and feedback using a multiple baseline across tasks design. After a six-week baseline, prompts were added to waiters' checks that said "ask" and "sold" and referred to either red or white wine, depending on which was targeted at the time. Red wine was suggested in Phase 1 and white wine in Phase 2. Corresponding to the prompting condition, goals were set to increase to 100% the suggesting of wine on all orders and the selling of wine on 20% of the orders. In the next condition, individualized posted feedback was added that depicted sales, suggestion percentages, and the number of glasses sold. Goals were not discussed during this condition; presumably, however, they were still in effect. Both interventions had an effect on suggestive selling, and there was correlational evidence that sales were affected as well. However, the effects of goal-setting lessened in each week of this condition for both types of wine (red wine, 59%, 46%, 38%; white wine, 70%, 60%, 44% over 0% in baseline for both). Feedback increased suggestive selling to a high of 67% for red wine and 61% for white wine. Red wine orders with meals went from 2% in baseline to 7%, 5%, and 2% in goal-setting, to 4%, 6%, and 7.5% in feedback. White wine orders went from 8% in baseline to approximately 16% in goal-setting to a high of approximately 20% in feedback. Follow-up data collected for six weeks in a return to baseline condition (no prompts, goals, or feedback) indicated that the percent of orders containing wine sales

was not maintained.

Zohar, Cohen, and Ahar (1980) implemented a program of posted, graphic feedback with workers in a metal fabrication plant in order to promote the use of ear protectors. In a between subjects design, workers in two noisy departments were given standard lectures about hearing conservation in noisy environments. Six randomly selected members of the experimental group received audiometric tests daily for one month, one test just before the start of their shift and one test just after. While the dependent variable for this study was the percentage of workers in each group wearing ear plugs, the information fed back to workers was the amount of temporary hearing loss that day (outcome feedback) and whether the worker had worn ear plugs (process feedback). One copy of the audiogram was given to the worker and another copy was posted on a bulletin board. The percentage of ear plug use in a baseline phase of the study varied between 30% and 50% for the experimental group and was approximately 10% for the control group, which did not change over the course of the study. The percentage of ear plug use steadily rose in the experimental group over 5 months and reached a plateau of approximately 85% to 90%. Of interest in this study is the fact that not every worker received the audiogram feedback, yet clearly most of the workers began to comply with the practice of wearing ear plugs. In addition, the use of outcome feedback (actual hearing loss), rather than process feedback (the percentage of workers wearing ear plugs) did prove to be quite effective. One could speculate that the outcome information was quite relevant to the workers and hence a potent controller of behavior. By comparison, one could speculate that the Van Houten et al studies on highway speeding would have shown more drastic

effects by including on their feedback signs not only the percentage of nonspeeders but the number of accidents on those roads as well.

Posted Feedback for the Group

Posted feedback to increase the use of behavior modification techniques by staff with mentally retarded clients has shown efficacy in separate investigations (Panyan, Boozer, & Morris, 1970; Welsch et al, 1973; Kreitner et al, 1977; Prue, Krapfl, Noah, Cannon, & Maley, 1980). The Panyon et al. study employed a multiple baseline design across halls in an institution for retarded children. Initial training in the skills led to a fairly high rate of use on each of four halls (63% to 79%) but which quickly dropped off in the coming weeks without follow-up. Posted feedback was implemented that showed, for each hall, the percentage of behavior modification sessions conducted out of a total possible along with the names of the staff members who conducted them. The use of percentages allowed ranked comparisons among hallways to be made. Results indicated that for each hallway, use of techniques improved to levels of 90-100%.

Prue et al (1980) increased the number of hours of staff treatment activity and the number of hours of client participation in treatment activities in a psychiatric hospital. Feedback consisting of weekly summaries of each unit's treatment activity was presented to the unit's treatment coordinator in one condition of this study. In a second condition, treatment coordinators specified treatment goals, and in a third condition, unit treatment activity was displayed in the lobby of the hospital where it was highly visible to employees, clients, and visitors. In general, each additional condition led to improved

performance. By the end of the last condition (public posting) there was a 163% increase in staff treatment hours and a 218% increase in client participation in treatment programs.

Chhokar and Wallin (1984) used graphic, publicly posted feedback with 58 employees in a heat exchanger manufacturing and repair plant to test how frequently feedback need be given in order to maintain safe behaviors. Employees were observed and scored according to a behavioral checklist of safe behaviors. An interaction design ($A-B-BC_1-BC_2-B-BC_2$) compared two levels of feedback (BC_1 = weekly posting; BC_2 = biweekly posting) to baseline (A), and training and goal-setting (B). Training consisted of a slide presentation of safe and unsafe practices, and goal-setting consisted of a 95% group compliance goal with safe practices. Mean baseline compliance was 65.21%. This rose to 80.92% when training and goal- setting were instituted. However, the goal of 95% was not reached until feedback was instituted (mean of BC_1 = 94.58%; mean of BC_2 = 96.78%; mean of the second BC_2 phase = 93.91%). The two levels of feedback, weekly versus biweekly, did not significantly differ from each other, indicating that feedback posted once every two weeks may be sufficient to maintain a high level of safe performance. However, the reversal phase in which all feedback was withdrawn (mean = 89.11%) indicated that feedback is a necessary component for maintaining the optimum level of safe practices. Whether feedback alone could have acheived this result without training or goal-setting was not tested.

Komaki, Collins, and Penn (1982) compared control by antecedents to control by

consequences in a poultry processing plant. The safety performance of 200 employees was monitored three times per week for 46 weeks in a multiple component, multiple baseline design (A-B-C-D-C) across four departments. Antecedent control (phase B) consisted of a safety training program, the posting of safety rules, and the supervisor discussing the rules at weekly safety meetings. Consequence control consisted of a feedback graph that was posted three times a week (phases C) and once a week (phase D). The supervisor also discussed safety scores at the weekly meetings. Results for antecedent control were mixed, indicating significant improvement in two of the four departments. All departments significantly improved in the feedback phase, above levels achieved in the antecedent phase. In addition, all workers were polled on their preferred method ("rules" or "graph") and 72% preferred the graph because "they could see how well they were doing."

Komaki, Heinzmann, and Lawson (1980) tested the efficacy of the different training and feedback components for safety in four sections of a city's vehicle maintenance division. After safety practices were defined, a multiple baseline design with a reversal phase (A-B-BC-B-BC) was implemented. Safety training (phase B) consisted of discussion, illustration, and posting of desired practices, while feedback (phase BC) consisted of a daily posted graph of the safety observations. Results indicated that the percentage of safe incidents increased from the baseline phase to the first training phase by 9%. Changes from the training-only (B) phase to the training plus feedback (BC) phase were significant in three of the four departments and in same direction in the fourth, for an overall mean improvement of 16% over training only and 26% over

baseline. The reversal phase to training-only phase indicated a decline in safety performance by an average of 9%, which improved in two departments when feedback was reinstated three to four times per week. In the other two departments, feedback was reinstated only one to two times per week, and performance did not significantly improve over the reversal phase.

Reber and Wallin (1984) worked with farm machinery manufacturing plant employees to increase the number of safe behaviors performed on a behavioral checklist (e.g., wiping up spills, wearing safety goggles). Using a multiple baseline across the three departments of the plant, the study was designed to test the differential effects of training, goal-setting, and group feedback. The order of conditions was (a) baseline; (b) training only ; (c) training and goal-setting; and (d) training, goal-setting and feedback. Baseline measures indicated a mean of 62.2% of the items on the safety checklist were performed. Each component of the program significantly improved safe practices (training mean = 70.85%; training and goal-setting mean = 77.54%; training, goal-setting, and feedback mean= 95.35%). The authors concluded that feedback is a critical component for achieving maximal performance. The design of this study, however, did not allow for the test of other orders or combinations of the conditions.

Van Houten and Nau (1981; 1983) have tested the use of publicly posted feedback on highway speeding. In one study (1981) comparing posted feedback to increased police surveillance, feedback was found to be 10 times more effective than surveillance and ticketing in controlling speeding. Driver speed was measured with a concealed radar unit. Using a counterbalanced reversal design on two streets, highly visible police surveillance

was alternated with feedback signs that indicated the percentage of drivers travelling at or below acceptable limits during the previous week, as well as the best to date. Police surveillance occurred for 30 minutes per day in one condition and 60 minutes per day in the next condition. Feedback signs reduced speeding for all levels of speeders, with the greatest reduction within the higher speed categories. Police surveillance did not reduce the percentage of speeders.

Van Houten and Nau (1983) conducted a series of five experiments to further investigate the parameters of feedback signs. Experiment 1 systematically varied the criterion used to define speeding on the posted feedback sign and found that a lenient criterion which allowed for the posting of high percentages of nonspeeding drivers was more effective than a stringent criterion. This finding lends credence to the notion of group norms as a controlling factor in group behavior. Experiment 2 examined the lasting effects of feedback by measuring the distance beyond the feedback sign of reduced driving speeds and found it to be approximately 6 km. Experiments 3 and 4 compared the effectiveness of an unmanned, parked police car and a police air patrol speeding program (respectively) to posted feedback. Both programs produced larger effects than feedback initially, but unlike posted feedback, the magnitude of their effects attenuated over time. Experiment 5 compared the traditional speeding enforcement program (surveillance and ticketing) to a warning program in which drivers were given a flier describing the number and types of accidents that had occurred on that road during the previous year. The warning program produced marked reductions in speeding while the traditional program did not. The warning program combined with posted feedback produced an even

greater reduction in speeding than either program alone.

The results of these studies are interesting given the fact that the "feedback" may have been given to people who did not perform the initial behavior (driving on that road) for which the information was posted, and yet, the information was highly effective. A conditioned reinforcement model of information continues to make sense under these circumstances. The information tells the driver that "most people" (a high percentage) were not speeding. Thus, adherence to group norms may account for the program efficacy.

A nonequivalent groups design was used to test informational feedback with 196 aircraft technicians in a helicopter maintenance unit (Catano, 1976). Technicians were required to fill out cards on each completed job describing the purpose of the job, downtime, man-hours, part identification numbers, serial numbers, parts replaced, etc. This system allowed for an up-to-date inventory. However, the error rates on these cards were between 50 and 65%. The experimental group received outcome feedback about organizational changes made as a result of this collected data. Errors for the experimental group significantly decreased after feedback, and were significantly fewer than the control group's rate, which did not change over time. Interestingly, the technicians were never told that their error rates were decreasing. When technicians were told that their feedback (job cards) to the main office were important because changes were based on them, the accuracy improved. Thus, outcome information regarding their work improved productivity.

Karan and Kopelman (1987) used outcome feedback regarding the number of vehicular accidents in a large nationwide package forwarding company. Using a

nonequivalent control groups design with the three facilities of the company, Facility A received feedback on the number accidents during the current year to date, the previous year, the number of days since the last accident, and each shift's relative standing (percentage improvement/decrement on past performance.) Facilities B and C served as a control group. The frequency of accidents in Facilities B and C rose 24.43% during the intervention phase (42 weeks) but decreased in Facility A by 2.05% over baseline. This improvement, compared to the previous year's accident rate, saved the company $28,129. It would have been impossible to directly observe each driver's behavior in order to determine safe driving practices. The use of outcome feedback was nevertheless effective in reducing accidents rates.

Feedback "Packages"

Psychologists have moved more and more into the private sector and have taken their techniques and methodologies with them. Difficulties arise, however, when companies are unable or unwilling to adequately test the use of techniques in these new settings. The result is a number of studies that provide limited information due to uncontrolled variables and/or threats to validity. However, as Kazdin (1982) points out, the ability to make causal inference depends on several factors that are possible in pre-experimental designs, such as "the type of data that are obtained, the number of assessments, whether information is available about past and future projections of performance, the types of effects that are achieved by the intervention, and the number and hetergeneity of the subjects. When several of these conditions are met, pre-experimental designs can rule out selected threats to validity" (pp. 101-102). In addition, research with poor internal

validity can have high social validity due to high returns in cost-benefit analysis (Wolf, 1978). For these reasons, the following studies are reviewed here.

A process approach to electricity conservation was taken by Luyben (1984) who modified conservation practices directly. Instead of giving outcome feedback on the amount of electricity consumed (KWH), he provided feedback to college faculty members about whether window blinds were properly dropped and tilted to reduce heat loss through the windows and to deflect heat back into the room.

The study took place in two campus office buildings in upstate New York with 52 faculty members involved. Maintenance personnel who cleaned the offices daily collected the data and eventually provided the feedback. After a baseline phase of 10 and 11 days for the two buildings, two conditions were implemented. In the first condition, prompting, the college president announced the new conservation policy in a special bulletin delivered to each faculty member, which was followed up a few days later by notices in the campus newspaper. In the second condition, one-half of the faculty offices in each building received feedback in addition to the prompt. Maintenance staff left notes on the desks of faculty members in the prompt plus feedback groups, either thanking them for complying with the program or requesting for them to do so.

The intervention phase lasted for 27 days. The data were analyzed by comparing the median performance for each condition (prompt; prompt plus feedback) at each phase (baseline; treatment) for each office building. The prompt-only groups increased their compliance to drop-and-tilt by a median of 11% in one building and 41% in the other building. The prompt plus feedback group increased their compliance by 56% in one

building and 65% in the other. No statistical analyses were performed.

While a clear pattern of improvement is shown from phase to phase in this study, design flaws are evident. By not staggering the feedback intervention across the two buildings to form a two-legged multiple baseline across settings, a threat of history cannot be ruled out. Furthermore, a reversal phase would have contributed to internal validity by demonstrating the effects of the removal of feedback in comparison to the prompts that were given at the beginning of the treatment phase and could not, by definition, be withdrawn. Next, the use of phase medians instead of time series analysis obscures the fact that there were increasing trends for both conditions in one of the buildings after the treatments were implemented. (This is evident only by viewing graphs.) While it might be hypothesized that feedback has a cumulative effect, one would have to question why prompts, given at the beginning of the treatment phase only, would have a cumulative effect over time. Since the effect did not hold true for the other building, it is likely that diffusion of treatment or compensatory rivalry (Cook & Campbell, 1979) took place, weakening internal validity even further. In sum, while on the surface the both treatments appear to show efficacy, design flaws weaken the interpretability of these results, since threats to validity cannot entirely be ruled out.

In a treatment designed to improve gasoline conservation, Runnion et al (1978) used an A B C design on a group of 195 textile mill truck drivers who drove over 6,000,000 miles per year for their company. After a one-week baseline measuring miles per gallon, an intervention package consisting of instructions on how to save gasoline, publicly posted graphed feedback for the individual and the fleet, and personal

commendation letters to randomly selected drivers who showed improvement, was instituted. This phase lasted one year. In the next phase of the study, individual feedback letters to each driver were added, along with publicly posted feedback charts with the best drivers listed (recognition), and a weekly drawing for small prizes. This phase lasted another year. Results indicated that the first intervention increased miles per gallon from 5.73 to 6.02, an increase of approximately 5%. In the second year of intervention, miles per gallon dropped from 6.02 to 5.97, which was still an increase of approximately 4% over baseline (5.73). This study is beset by many of the same problems as the Luyben study. The use of mean levels, instead of looking at trends, is misleading. First of all, the one data point in baseline may be spurious, so that later "treatment" effects might really be the effects of regression or cyclical variation. This is indicated because there is a sharp decrease in the miles per gallon towards the end of the first year. Second, while the mean of the second year's intervention package is 5.97, less overall than the first year mean of 6.02, there is a steady increase in miles per gallon across weeks. The mean level is much higher than the data points at the beginning of the year, fairly representative of the middle, and much lower at the end of the year. Failure to point this out indicates that the first intervention package is better than the second, which in fact is not necessarily the case. (A direct comparison was not performed.) Nevertheless, this study had high social validity to the company. As the authors pointed out, even a fractional improvement in miles per gallon translates into many dollars in a year's time. In fact, during the second year of the study, the company saved enough money to run its entire fleet at no cost for one month. Company personnel also stated that they believed the drivers had greater

job satisfaction due to greater recognition and interest, and that drivers became more responsive to requests for improvements in safety, preventive maintenance, and efficiency.

On a more global level, Rothstein (1980) investigated the use of televised feedback to reduce the gasoline consumption of an entire community. Every weekday evening in the summer of 1979, the local 6 p.m. news for Midland-Odessa, Texas (population 300,000) ran a graph that showed the number of gallons of gasoline consumed since 3 p.m. the previous day. The graph was on the screen for 30 seconds, during which time the announcer gave a conservation tip, offered commendation, and tried to encourage competition between the two cities. The number of gallons of gasoline consumed daily was tabulated from a random sample of the 190 gasoline stations in the area. Results showed decreases of 24.8%, 27.2%, and 37.6% in the first three weeks of feedback. In a reversal phase the graph was not shown for the following two weeks. Fuel consumption returned to 75.2% of baseline by the second week. When the graph was returned to television, consumption again dropped to 33.7% in the third week. A follow-up three months after the last graph was on the air indicated that the towns were still using only 84.8% of their original baseline consumption. The possibility that rising gasoline prices contributed to the reduction of consumption is ruled out by the use of the reversal design because consumption varied with the treatment.

Four textile machine operators increased their productivity by a mean of 7.5% when each received a weekly graph of her average number of hours per day and week gained or lost (Dick, 1978). Improvements were noted and praised by crew foremen. Internal

validity could have been increased through the use of a multiple baseline across subjects design; instead, all operators received feedback at the same time. Nevertheless, this simple project resulted in annualized savings to the company of $3400.

Rowe (1981) decreased the amount of time it took for independent automobile appraisers to appraise and report financial estimates of damages to automobiles. Although they were required to do this within 24 hours of a request, there were no contingencies for doing so, and fewer than 40% did. The feedback system consisted of a 98% compliance goal, a posted graph that depicted the percentage of appraisers who met the 24-hour deadline, copies of the graph sent to the claims department supervisor, and prompts from the employees who assigned the appraisals. Within 10 days the compliance rate to the 24-hour period was 100%. Financial benefits to the company were not assessed; however, the faster settlement of claims resulted in fewer customer complaints regarding slow payment of damages.

The reduction of materials waste in a production department was achieved with a combination of goals, feedback and praise (Eldridge, Lemasters, & Szypot, 1978). Workers were required to make a judgment about the quality of sheets of material and discard poor quality material. Returned orders of defective material had to be unpacked, inspected, and repacked by the operator who had shipped it; therefore, there was a high rejection rate of material that was in fact acceptable. Piece yield was calculated by dividing the actual number of units packed by the theoretical maximum possible from perfect material. Baseline data indicated a 57.6% average piece yield. A feedback sheet was designed to calculate piece yield and was given to workers, who were told to strive for

a goal of 75% piece yield. Piece yield increased to the stated goal which saved the company $105,000 in reduced waste in the first year. In addition, productivity also improved, enough so that a second shift was eliminated.

Runnion, Johnson, and McWhorter (1978) used an A-B-C-D design to test feedback and reinforcement on reducing the truck turnaround time of a textile mill. Intervention consisted of weekly feedback of truck turnaround time sent in a letter to each plant manager (condition B), biweekly feedback in the same manner (condition C), and monthly feedback (condition D). The letters also contained praise, prompts and suggestions for reducing turnaround time. Tangible group and individual reinforcers were also provided. The average results for each condition were: baseline, 67 minutes; weekly feedback, 39.1 minutes; semi-weekly feedback, 37.2 minutes; and monthly feedback, 38.3 minutes. Increased efficiency resulted in internal transportation of 12% more materials.

A combination of posted feedback, daily goal-setting, and verbal praise reduced the number of "high bobbins" in a textile spinning department (McCarthy, 1978). A reversal design (A-B-A-B) demonstrated that the average number of high bobbins across phases were 55.9, 14.2, 8.7, and 8.1, respectively. The use of averages obscures the trend lines that were in the desired directions with each introduction of the intervention.

Frost, Hopkins, and Conard (1982) also used a reversal design (A-B-A-B) to test feedback and verbal reinfrocement on machine-paced production. Six employees in a light manufacturing department were assigned to package rosin bags. Prep time was the amount of time it took to ready bundles of rosin bags for filing; fill time was the amount of time it took for the bags to be filled. Total cycle time was the sum of both. Employees received

graphed, posted feedback on the amount of total cycle time. Mean total cycle time went from 33.92 minutes to 28.78 minutes, with further decreases and increases corresponding to the second A and B conditions. Fill times showed no changes over conditions, but prep time changes were considerable, with feedback and reinforcement procedures generating an overall 26% increase in productivity.

Summary and Hypothesis

There can be little doubt that "feedback" has proven to be an effective treatment in a wide range of applications and that the term has become an accepted part of behavioral terminology (Kopelman, 1982; Prue & Fairbank, 1981; Rapp, Carstensen, & Prue, 1983). However, Balcazar et al's (1985-86) review of performance feedback states that consistent effects of feedback as the sole treatment were observed in only 28% of the total number of feedback applications, mixed effects in 57% of its applications, and no effects in 15% of its applications. These figures rose when other treatments (reinforcement, goal-setting) were added to feedback. Many dimensions of feedback have been identified (Ford, 1980; Herold & Greller, 1977; Ilgen, Fisher, & Taylor, 1979; Larson, 1984; Queglieri & Carnazza, 1984-5), including the credibility and power of the source, whether feedback was self-monitored, its mode of transmission (oral, written, graphic; public or private), process type or outcome type, and its valence, frequency, specificity, amount, accuracy, and timing. Given these many dimensions that are rarely held constant across studies, it is not surprising that mixed effects are found.

The efficacy of feedback, like reinforcement, will differ depending on the learning histories of the participant, the dimensions of the feedback stimulus, its schedule of

implementation, and whether it is paired with effective back-up reinforcers. Information alone, without reinforcement, appears to lose effectiveness when that information does not lead to reinforcement. Hence, behavior that initially can be controlled by feedback may extinguish if there are no consequences for attending to it.

In the current study process feedback on the sales skills of three telemarketers in a thrift institution was examined. The purpose of the study was to attempt to separate the informational, instructional, and reinforcing components of feedback by only providing written information about performance. Feedback was individualized and private, and no comments about performance were supplied, including suggestions for improvement. In this study, therefore, an attempt was made to discover whether information alone, without goals, peer pressure via public posting, praise, or any other additional contingencies, could consistently function as a reinforcer.

A multiple baseline design across sales skills was employed. It was expected that skills would show an initial increase in frequency which would decline without additional reinforcement. A reinforcement phase targeting two skills at a time and alternating between them was expected to result in increases in frequency of skill use for the targeted skills.

Method

Participants

The institution was a thrift (or savings and loan) that ranked in the top 100 of thrifts nationwide in net income and commercial loans as of December, 1986, and in the top 300 for deposits and assets as of June 30, 1987 (American Banker, 1988).

The study participants were three female telemarketers employed by a savings and loan institution in a midsize city in the Northeast with a population of approximately 900,000. The telemarketers' jobs consisted of answering phone calls from current and potential customers in order to describe and sell bank accounts, provide service for account problems, and answer customers' questions. The telemarketers also made calls to current customers for special promotions (e.g., checking accounts for new mortgage customers) or to remind customers of certificates of deposit about to come due. These calls were made with the aid of a script. The telemarketers were responsible for maintaining their own sales records, generating correspondence with the customer, and following up on any unresolved problems with customers.

The telemarketers had less than one year's combined experience in either telemarketing or banking, although they had all completed a two-week bank course in teller and bank operations training. They also completed a brief sales training course just prior to the opening of the department, which will be described in detail on page 57. The telemarketers did not know that they were participants in an experiment until the end of the experiment, at which time they were debriefed in person. (Written informed consent was not required because the telemarketers were not deemed to be at risk.)

Apparatus

The telemarketing department functioned as an independent branch of the bank. Its office, which measured 16 feet by 21 feet, was located so that there was no possibility of any face-to-face customer contact. The department consisted of three telemarketers (D., T., and M.), a supervisor, and a clerk-typist. Each telemarketer was situated at a work station and provided with noise-reducing partitions, a desk, telephone, file, and computer terminal, which had the capacity of accessing and amending customer files.

The telemarketers were each equipped with a Northern Telecom SL-1® Business Communications System telephone. This system included an Automatic Call Distribution (ACD) feature that allowed incoming calls to be automatically directed to the telemarketing station which had been idle the longest (and was ready to accept calls.) This feature therefore allowed calls to be evenly distributed among telemarketers. Telemarketers did have the capacity to turn off the ACD unit during breaks or while performing other duties. In general, however, each telemarketer received approximately the same number of calls each day as any other telemarketer.

Each station was furthermore linked to an ACD Management Reporting feature so that various aspects of telephone usage could be recorded on an hourly basis. While this feature recorded many variables of interest to the supervisor of the department, those used in this study were (a) the number of inbound calls per day per telemarketer, (b) the cumulative duration of these calls each day, and (c) the average length of the calls per day per telemarketer. In addition, telemarketers recorded the number of sales they made per day on a paper-and-pencil log. Logs were checked for accuracy by the departmental

supervisor, since these sales records were used to determine departmental efficiency.

Each telemarketer was equipped with a Radio Shack tape recorder on which she recorded all of her inbound calls from customers. Only the telemarketer's side of the conversation was recorded; the customer was not recorded to protect confidentiality. The conversations were recorded on standard cassette tapes.

Preexperimental Procedures

Sales training. At hiring and prior to the start of this study, all telemarketers had been trained in the use of specific telephone sales skills. Two of the three telemarketers received training which consisted of a 3-day behaviorally-oriented workshop including a slide presentation of sales skills as an overview, lectures, and video-taped role-playing. Each telemarketer participated in problem solving and memorized specific sales strategies (e.g., suitable target accounts for potential cross-selling.) Telemarketer "M." was hired later and did not participate in the sales training workshop. (This difference appeared to have no impact on the results of the study.) The workshop was held in a conference room on bank premises during regular working hours.

All three telemarketers completed a self-instructional manual detailing the use of sales skills which corresponded to the sales training workshop. The workshop and manual were tailored to the skills required in bank telemarketing. They included training in product knowledge, telephone etiquette and voice skills, and specific sales skills.

Selecting the targeted sales skills for inclusion in study. The sales skills to be targeted in the study were selected from the sales training material. These skills are delineated below.

Definition and scoring of identifiable sales skills. A preliminary data sheet was devised to collect data from the tape recorded telephone calls. This data sheet had a space to identify the purpose of each call and tallying the frequency of 14 sales skills.

The 14 sales skills were divided into five categories: (a) sales techniques, (b) benefits; (c) handle objections; (d) ask for sale; and (e) establish follow-up. "Sales techniques" included (1) using the customer's name (in order to establish rapport with the customer); (2) asking the customer leading questions (in order to discern other potential sales opportunities with the customer); and (3) attempting a cross-sell (that is, the introduction of an account in addition to the one the customer has inquired about.)

The category called "Benefits" included (4) describing account benefits (i.e., describing the account in terms of the advantages for the customer, as opposed to a "features" description which simply outlines the facts of an account); and (5) using the "benefit tie-down technique" (in which the telemarketer states the account benefit and ends it with a question to which the customer will agree; for example, "This is a great rate, isn't it?")

The category called "Handle Objections" included skills that a telemarketer could use to counter the objections raised by a customer. Objections might include low interest rates on savings accounts, high fees or penalties, or problems specific to a customer's account. There were three skills in this category: (6) using the "feel-felt-found" technique [in which the telemarketer offers empathy (e.g., "I understand how you feel"), validates the customer's concern ("other customers have felt the same way"), and offers a solution ("however, they found that if they . . .")]. (7) offsetting objections with correct

information (in the case where the customer had incorrect information), and skill (8) was whether the telemarketer offset objections with account benefits (when shortcomings of the account could not be changed, the telemarketer stressed other good points about the account.)

There were four sales skills under the category called "Ask for the sale." Any of the four methods could be used to ask the customer to open the account. These included: (9) the direct technique, [in which the telemarketer directly asks the customer to open the account ("May I begin the paperwork?")]; (10) the impending event technique [in which the telemarketer stresses an upcoming event or deadline which would compel the customer to open the account now ("Our special ends on April 10; why don't you take advantage of it before it's too late?")]; (11) the assumption technique (in which the telemarketer assumes the customer will open the account and begins the paperwork without asking the customer for the sale directly;) and (12) the agreement technique [in which the telemarketer stresses the points which have been agreed upon by the customer, so that opening the account becomes the logical conclusion ("As you stated, Mr. Jones, this is a flexible account that gives you a competitive interest rate. I think you'll agree that this is right for you. Let's begin the paperwork.")]

The final category, "follow-up established," indicated whether the telemarketer had suggested the possibility of future contact with the customer. With skill (13), passive follow-up, the telemarketer invited the customer to contact the department again for further information (e.g., "Please feel free to call me back if you decide you want to open the account.") With skill (14), active follow-up, the telemarketer stated that she would

contact the customer and requested the necessary information (e.g., "If I may take your name and address, I'd be happy to send you a brochure about this account.")

Rationale for the four chosen sales skills. Poor interobserver reliability, the telemarketers' infrequent use of some of the skills, and their high rate of other skills made it necessary to eliminate ten of the 14 sales skills from consideration in the study. Using the customer's name, attempting to cross-sell, and describing benefits were selected on the basis of acceptable interrater reliability scores. The fourth skill selected was asking for the sale. Because three of the four ways of asking for the sale (the impending event technique, the assumption technique, and the agreement technique) were used infrequently, asking for the sale was defined as including any of the four methods.

When the telemarketer asked leading questions (skill 2), she was attempting to discover the customer's financial needs and other sale possibilities. The interobserver agreement scores for this skill, however, were highly unreliable. Without being able to hear the customer's side of the conversation, it was difficult to determine if the telemarketers' questions were simply gathering the required information for the purpose of the call, or whether the questions were actually probing the customer to reveal more information than might have been necessary to transact the business at hand. This skill was therefore eliminated from the study. An alternative idea, to simply count every question asked, would not have been a valid measure of the skill. Many routine questions are asked once an account is being opened, but this might reflect absolutely no "sales initiative" on the part of the telemarketer.

"Benefit tie-down statements" (skill 5) were so infrequently used by the

telemarketers that reliable scores could not be obtained. In addition, it is directly linked to skill 4, describing benefits, because it is a method of just attaching a rhetorical question to the end of benefit statement. A multiple baseline design requires independent behaviors. Therefore, this skill was eliminated from the study.

The entire category of "objections" had to be eliminated because the customer was not heard, making it impossible to know whether the customer had had any objections. Even when the telemarketer could be heard answering questions or supplying information, it was impossible to know what prompted these statements, or, if it seemed fairly evident that there had been some sort of objection, whether the telemarketer was making the most appropriate response.

While interobserver agreement was good, the entire category of "established follow-up" was also eliminated because the telemarketers consistently made appropriate follow-up statements. With both passive and active follow-up near the ceiling level of 100%, it was not practical or necessary to intervene on these skills.

In summary, then, the four sales skills targeted for this study were: using the customer's name in the conversation, cross-selling (i.e., introducing additional accounts to the customer in a sales presentation for another account), describing account benefits, and asking for the sale.

Final Experimental Procedures

Daily procedures of the telemarketers. The telemarketers recorded their conversations with customers by setting their tape recorders on "record" in the morning and setting the "pause" button. When they were signalled for a telephone call, they turned

off the pause button and answered the call. They recorded only inbound calls; outbound calls made to customers were omitted. ("Inbound calls" come into the department from the customer via a toll-free telephone number. In addition, telemarketers made outbound calls to current customers to remind them of maturing certificates of deposit or other areas of interest. These outbound calls were made with the aid of a script and were not included in the current study.) The cassette tapes of the calls were then sent twice a week by overnight mail service to the experimenter who scored them for the use of sales skills. (Transcripts of sample phone calls can be found in Appendix A.) Photocopies of the sales logs and telephone software data sheets were also sent.

Scoring the tapes. The data were collected for approximately 35 weeks. Several steps were involved in collecting the data from the cassette tapes of phone calls. The experimenter determined (a) the purpose of the customer's phone call (by inference from the telemarketer's statements), (b) the opportunities which existed for the telemarketer to use each of the four sales skills, and (c) the number of times each of the four sales skills was used. Each step will be considered.

Determining the purpose of the phone call. The experimenter used the Telemarketer Sales Score Form (Appendix C) revised to reflect the four targeted sales skills. The name of the telemarketer and the date of the tape were recorded at the top.

Because the opportunities to use the various sales skills were dependent upon the nature of the customer's call, it was necessary to first identify why the customer had called. Because the customer could not be heard, the experimenter inferred the purpose of the telephone call from what the telemarketer said to the customer.

Four major types of call purposes were identified: (a) an account inquiry, in which the customer has called to either open an account or ask for information; (b) a service call, in which the customer has called to have a problem solved (e.g., resolve a checking mistake, ask for a branch phone number); (c) a bank survey, in which competing banks call to inquire about current interest rates from this bank; and (d) inbound calls which require the use of a script on the part of telemarketers. These calls included inquiries on mortgages, home equity lines of credit, or a phone call from a customer who had received a message to call the department about a maturing certificate of deposit (CD) or individual retirement account (IRA.) Only account inquiry calls, which indicated the possibility of a sale, were scored for the use of sales skills. Service calls were not included due to the bank's philosophy that the telemarketers should not try to sell accounts on these types of calls. Calls in which scripts were used were also omitted, since the script provided the sales skill which the telemarketer had only to read.

Identifying opportunities to use sales skills and their use. After the purpose of the call was identified, the experimenter determined whether the telemarketer had the opportunity to use any of the four targeted sales skills. It was next determined whether the telemarketer had actually used any of the skills, and if so, how many times. Whether the telemarketer had the opportunity to use a skill depended partially on the type of account for which the customer had called. Therefore, correct identification of opportunities depended on correct identification of call types.

Determining the opportunity for using the customer's name. If the telemarketer knew the customer's name, she should have used it during the phone conversation at least

once. (She should not necessarily ask for it. Many customers prefer to remain anonymous when they call for information.) The opportunity for using the customer's name was always scored "yes" because it was not always possible to determine whether the customer had given his or her name to the telemarketer. While this method includes error (the telemarketer might be scored as having the opportunity to use the name when in fact she did not know the name), it was preferable to relying on inference from the telemarketer's remarks about whether she knew the name. This method is likely to yield an underestimate of the telemarketer's ratio of name use to opportunity.

Determining the opportunity for cross-selling. Cross-selling is the skill of prompting the customer to open additional accounts. It is a difficult skill to perform because the telemarketer must make a smooth conversational transition from the account being discussed to one which is not being discussed. It is an important skill, however, because cross-selling expands the bank's account base without having to draw in new customers. The opportunity for cross-selling relied on two factors. The first factor was whether the customer was thought to have funds at his or her disposal with which an additional account could be opened. It was assumed that the customer did have such funds in all cases except for maturing certificates of deposit (CD's) which were not yet up for renewal. In such a case the customer already has a CD account with the bank but does not have access to that money until the term of the account is up. It was reasoned that the telemarketer could not cross-sell in these cases because the customer did not have available funds with which to open additional accounts. However, if the customer called while the CD account was in the 7-day grace period between the end of the term and automatic renewal, cross-selling was

possible. The customer, during this 7-day period, has the option of renewing the account at the same term, closing the account and withdrawing the money from the bank, or reinvesting the money into different terms or different accounts. The telemarketer at this point could suggest many alternatives to the customer in order to both maintain the banking relationship and expand it (e.g., withdraw some of the money from the CD to open a money market account for more liquidity and maintain the rest of the funds in a CD for the higher interest rate.) (See Table 1 for a list of accounts and their corresponding opportunities to cross-sell and ask for the sale; see Appendix B for a description of typical bank accounts.)

The second factor determining the opportunity for cross-selling a second account was whether the sales presentation on the first account was already quite lengthy or complicated, making cross-selling impractical. This was true for only one account, the IRA loan. The IRA loan was a complicated product to describe; to ask the telemarketers to cross-sell on an IRA loan call would have been impractical for the telemarketers and confusing to the customer.

Determining the opportunity for using benefit statements. There was always an opportunity to describe accounts in terms of benefits; therefore opportunity was always scored "yes."

Determining the opportunity for asking for the sale. After telemarketers describe the details of an account, they should ask customers to open it, with two two exceptions. First, if the customer already had the account but was calling about its renewal, as in the case of maturing CD's or IRA's, it would not make sense for the telemarketer to request it

Table 1

Opportunities to Use Two Sales Skills According to the Call Request

	OPPORTUNITIES TO USE SALES SKILLS	
	CROSS-SELLING	ASK FOR THE SALE
Certificate of Deposit - new	Yes	Yes
Certificate of Deposit - renewal (not in 7 grace day period)	No	No
Certificate of Deposit - renewal (in 7 grace day period)	Yes	Yes
Checking	Yes	Yes
Direct Deposit	Yes	Yes
Individual Retirement Account (for renewals, follow same rules as for CD's)	Yes	Yes
IRA loan	No	Yes
Loans (personal, auto, home improvement)	Yes	Yes
Savings	Yes	Yes
Visa	Yes	Yes

to be opened. The second exception was when the customer asked for the account to be opened, in which case the telemarketer had no chance to ask for the sale. This case could easily be determined because there was little or no sales presentation, the telemarketer did not ask for the sale, and yet the customer opened the account.

Determining the use of each skill. After the opportunity to use each skill was determined, the experimenter listened for the use of each skill and noted it if it occurred. A check mark was placed every time the name was used. In addition, the name of the attempted cross-sell account was noted, and the "benefit" and "asking for the sale" phrases were written down on a separate piece of paper, noting for which call they occurred. This

was done for purposes of interobserver agreement.

Interobserver agreement. Three scorers were trained to listen to the tapes and determine the type of phone call as well as the opportunity for and use of the sales skills. Each scorer listened to the tapes of only one telemarketer. After training, interobserver agreement data were calculated on (1) the type of call, (2) the opportunities to use each skill, and (3) the use of each skill. Agreements were calculated for positive instances only, by dividing the total number of agreements by the total number of agreements plus disagreements to obtain the agreements ratio. Because the scoring of opportunities and skill usage depended on correctly identifying account types, agreement on opportunities and skills were calculated only on account types on which the experimenter and scorer agreed.

After training, scorers independently listened to tapes corresponding to 25% of the total number of data points for each telemarketer. For example, if a telemarketer had a total of 48 data points, the scorer listened to tapes corresponding to 12 data points. Each data point consisted of data from three days; therefore, in this example, the scorer listened to approximately 36 days of telemarketing calls. The data points selected for independent scoring were randomly chosen using a random numbers table.

The Independent Variables

Performance feedback. The first independent variable was informational feedback to each telemarketer regarding her use of the sales skills. The information was provided for three of the skills (customer name, asking for the sale, and cross-selling) in a multiple baseline across sales skills design. The fourth skill, describing benefits, never received

the intervention due to poor interobserver agreement at the time of the intervention.

All information was calculated on the basis of the recorded conversations and delivered by the experimenter twice weekly to the departmental supervisor by telephone. The supervisor wrote down the information on preprinted forms for each telemarketer (Appendix D) and gave them out. In this way the experimenter remained "one step removed" from the telemarketers, while the supervisor, who was instructed to remain neutral about the contents of the feedback, maintained a more natural role within the department. The telemarketers were aware that the feedback was generated from their recorded conversations and calculated by the experimenter, who had previously been a paid consultant for the bank and continued posing in this role.

At the top of the forms were the dates for which the information pertained. The forms stated, for each skill, how many opportunities the telemarketer had had to use the skill, how many times she had used it, and the resulting percentage, or sales skill ratio (# uses/# opportunities x 100). It was always possible that the ratio of skill use could exceed 100%. This was due to the fact that while a phone call represented one opportunity to use any given skill, the telemarketer might use that skill multiple times within that call. For example, using the customer's name three times in one phone call would equal 300%. Ceiling levels, or points beyond which the telemarketers could not improve, were not defined. It was assumed that if feedback improved skill use to the maximum (undefined) point, it would continue to act in a maintenance function, just as reinforcement maintains any modified behavior.

In accordance with the multiple baseline design, the forms described only the first

skill (customer name) for 11 semiweekly feedback trials. At that point, asking for the sale was added to the forms for the next 18 semiweekly feedback trials, and then cross-selling was added for the remaining 25 semiweekly feedback trials.

Due to the time constraints involved in sending the cassette tapes and scoring them, feedback was delivered with a two- or three-day delay between the end of the feedback period and the delivery of feedback. Phone conversations taped on Monday, Tuesday, and Wednesday were sent by overnight delivery service to the experimenter who received them on Thursday. The experimenter listened to the tapes, scored them on Thursday and Friday, and called the departmental supervisor on Friday (as early in the day as possible, depending on how many phone calls were to be scored) and delivered the information regarding the previous Monday's, Tuesday's, and Wednesday's sales performance. Meanwhile, the telemarketers were recording their conversations on Thursday and Friday, with those tapes being sent Friday evening to arrive on Monday. The experimenter scored those tapes on Monday and Tuesday and delivered the information regarding those days on Wednesday. This cycle continued for the duration of the experiment. Telemarketers who were out of work due to sick days or vacations received the information when they came back to work. No one was out for more than one week at a time. If the supervisor was out, the experimenter gave the information to the clerk-typist. The tapes were not sent on time one day in January because snow prevented it, and there was also a temporary rescheduling of tape delivery and feedback during the Thanksgiving and Christmas holidays due to bank closure.

Reinforcement. A reinforcement phase was added to performance feedback after 40

semiweekly feedback trials (approximately 20 weeks), while the informational feedback continued. The reinforcement consisted of a point system, with points being traded for prizes from a gift catalog. Reinforcement (number of points earned) was delivered on a preprinted form (Appendix E) by the departmental supervisor who got this information by phone from the experimenter. The approximate dollar value of each point was twenty cents. Due to practical constraints of the department, reinforcement was given on a weekly basis (Mondays), and, due to the time required to deliver and score the tapes, with the added constraint of weekly delivery of points, reinforcement was given for the period of behavior which had ended nine days prior to reinforcement (e.g., if the behavior occurred during the week of April 13-17, and reinforcement was delivered on Monday, April 27.)

To earn points, each telemarketer was required to improve her previous week's (or best) performance. Specific individualized goals were given to the telemarketers on the form (Appendix E) when reinforcement was delivered. These goals were based on the average weekly rate of increase between the baseline phase and the feedback phase for each skill, for each telemarketer. This figure, which remained constant, was derived from the following formula:

$$\frac{(\%\text{ skill use in feedback phase} - \%\text{ skill use in baseline phase})}{[(\#\text{ of feedback days}/5]}$$

For example, on the "asking for the sale" skill, Telemarketer "D.'s" baseline percentage of use was 31.91%; her percentage of use during the feedback phase was 88.37%. There were 63 feedback days. Therefore the formula is:

$$\frac{88.37 - 37.91}{[(63/5)]} = \frac{56.46}{12.6} = 4.48 \text{ (rounds to 4)}$$

Telemarketer "D." was required to increase her percentage of skill use on "asking for the sale" by 4 over her previous best performance. The new goals were determined by adding a constant number of percentage points to the current level of skill usage. A simple numerical increase was not used in an effort to equalize the amount of change necessary to earn the reinforcer across telemarketers. For example, if one telemarketer used the customer name skill 50 times in one day, an increase of five times would be an increase of 10%. For a telemarketer who only used the skill ten times per day, an increase by five would represent an increase of 50%. Thus, each telemarketer and each skill could have a different required rate of increase, but it matched her own baseline-to-feedback rate of improvement. Table 2 shows the constants (in parentheses) which were to be added to each goal if the goal for that skill were met. If the goal was not met, the goal remained the same as for the previous week.

Two skills (customer name and cross-selling) were randomly targeted to be reinforced at one time in order to contrast with the third skill (asking for the sale) for which performance feedback alone was presented and the fourth skill (benefits) which remained as a baseline control (no intervention.)

Twenty-five bonus points were also awarded weekly to the telemarketer who showed the most improvement. Improvement was calculated for each telemarketer by subtracting the previous ratio of skill use from the current ratio for each skill, adding the ratios together and dividing by two to obtain the average improvement across the two skills.

Table 2

Required Average Weekly Skill Increase (ratio of weekly improvement between baseline and feedback) to Receive Reinforcement

	% use in Baseline	% use in Feedback	# days in Feedback	Average Weekly Increase
Telemarketer "D."				
Customer name	0	83.00	81	5.12 (5)
Ask for sale	31.91	88.37	63	4.48 (4)
Cross-sell	25.17	40.38	21	3.62 (4)
Telemarketer "M."				
Customer name	100.00	154.01	90	3.00 (3)
Ask for sale	17.57	73.67	72	3.90 (4)
Cross-sell	20.57	35.25	24	3.06 (3)
Telemarketer "T."				
Customer name	0	118.18	87	6.79 (7)
Ask for sale	12.16	44.96	69	2.30 (2)
Cross-sell	13.27	13.68	24	.09 (1)

The telemarketer with the highest average received the bonus points for that week. By these calculations the most improved telemarketer was not necessarily the one with the highest rate of skill use, because there may have been ceiling levels at which there could

be no more improvement. It was planned that the skills to be reinforced would alternate. However, the reinforcement phase was abandoned after 9 weeks because there was little change in behavior with the reinforcement of the first two skills.

Discussion of Data Analysis

The Dependent Measures

The dependent variable for each of three participants was a ratio of sales skills used per opportunity, based on three consecutive days of phone calls, for each of four sales skills. In addition, a sales-to-calls ratio in three-day periods, calculated by dividing the number of sales made per three days by the number of calls accepted per three days, for each telemarketer, was obtained from the Automatic Call Distributing Unit Management Reporting feature of the SL-1® and the pencil-and-paper logs described in the "Apparatus" section.

Rationale of Data Analyses

Analyses on sales skill ratio data were performed with use of the Gottman-Williams Time Series Analysis Programs for the Social Sciences (Gottman, 1981; Williams & Gottman, 1982). Interrupted time series analysis removes any serial dependency from the data and performs t-tests for the intercepts and slopes of two phases. Thus, it treats each data set as an A-B design.

Significant differences may occur for intercepts, slopes, or both. A significant difference in both intercept and slope would indicate that there was a change (either increase or decrease) in the elevation of the intercepts across phases, and that the direction and/or rate of change differed across phases. A treatment effect may be assumed,

because the data sets are best described by two different functions. A significant change in intercept but not slope may indicate that the overall level of performance in the phases differed, but that their predicted rates of change would not differ. However, the direction of the change between intercepts must be considered by reviewing the regression lines on the graphs. A significant change (increase or decrease) in intercept from phase A to phase B does not necessarily mean that the level of skill use changed. A positive slope across phases A and B would result in a higher intercept in the B phase. A significant intercept change might also be an artifact of different (though nonsignifcant) slopes. This particular result was usually the case in this study. A significant change in slope, but not intercept, would suggest that the rates or direction of change between phases differed. This pattern of results did not occur in this study. The reader may refer to Gottman (1981, pp. 45-50) for a more detailed discussion of intervention effects.

The current data sets contained two planned interventions and one unplanned intervention for two legs of the design (customer name and cross-selling), one planned and one unplanned intervention for one leg of the design (asking for the sale), and one unplanned intervention for the last leg of the design (benefits). The unplanned intervention (the IRA loan program) overlapped with the planned interventions and partially confounded the design. Furthermore, because of the nature of a multiple baseline and its stepwise interventions, the design of each leg, due to the unplanned comparison, differed from every other leg. As such, none of the legs could be considered an A-B design. If A denotes baseline, B denotes feedback, C denotes the IRA loan program, and D denotes reinforcement, the following designs applied: A-B-BC-BCD-BD (customer name);

A-B-BC-B (ask for sale); A-C-BC-BCD-BD (cross- selling); or A-C-A (benefits).

In order to accommodate the requirements of time-series analysis for an A-B design, each leg of the design was subjected to multiple analyses using different contiguous portions of the data sets to comprise the A-B comparisons. Because it appeared, according to the graphed data, that the IRA loan exerted a powerful influence on responding, the data were analyzed twice: once ignoring the IRA data points as though they were not an influence (because this was an unplanned comparison), and once as though the IRA loan phase were a planned comparison. For example, in using the customer's name, the baseline phase was compared to the entire feedback phase up to the start of the reinforcement program in one comparison, as though the IRA loan program would have no impact. In a second comparison, the baseline phase was compared to the feedback phase only up to the start of the IRA loan program. Next, the first portion of the feedback phase without the IRA program was compared to the second portion of the feedback phase that included IRA points. This process continued so that all phases were analyzed both including and excluding the IRA from consideration in the phases. Only contiguous data points between phases were analyzed.

These analyses sometimes resulted in extremely short phases; for example, in using the customer's name, the reinforcement-plus-IRA loan phase contained only four data points. According to Gottman (1981, p. 359), as long as the sum of the autoregressive coefficients is ≤ 0.6 (as was the case in all but one of these analyses), interrupted time series analysis can be used with even very small samples. In addition, imbalance between the number of observations pre- and post-intervention makes very little difference

(Gottman, 1981, p. 359).

Using the slopes and intercepts provided by the Gottman-Williams program, regression lines were fit onto to graphs of skill use ratios. The graphs, in combination with the *t*-tests provided by the program, served as the basis upon which decisions about intervention effectiveness were made.

Additional analyses were performed in order to test the independence of the four skills. For each intervention that went into effect, there existed a control condition of an untreated skill in at least one other baseline. For example, when feedback began for using the customer's name, the other three skills remained untreated. When reinforcement went into effect for the using the customer's name and cross-selling, asking for the sale and benefits statements remained untreated. The strength of the multiple baseline design is in reducing threats to internal validity by showing that behaviors change only in response to treatment. If treatments are staggered in time, then the likelihood of other events (threats to validity) being responsible for behavior change is highly improbable. However, in order for this process to work, behaviors must be independent. This means that they must not be correlated so that a change in a treated behavior will result in a change in an untreated behavior. If independence is not demonstrated, internal validity is weakened, because one cannot be certain whether observed changes are due to a treatment effect or a threat to validity.

In order to conduct tests of independence, the portions of the untreated skill data which corresponded in time to the treated skill data were analyzed as though there had been a treatment. If no changes in function (intercept, slope) were observed in the

untreated skill, independence could be assumed. If changes were observed in the untreated skill, independence could not be assumed and/or threats to validity could not be ruled out.

This system of analysis led to 26 comparisons per telemarketer. While there is justifiably a concern of having made a Type I error, significance tests are reported at both the .01 and .05 levels.

Correlations were performed between sales skills ratios and actual sales to determine whether there was any relationship between the two measures.

Interobserver Agreement

Interobserver agreement was calculated using the point-by-point agreement ratio method for occurrences (Kazdin, 1982, p. 53). This method is an exact method of computing agreement because agreement is evaluated on a response-by-response basis (Kazdin, 1982, p.54).

Agreements were calculated for each telemarketer for (a) the type of customer inquiry; (b) the opportunities to use each of the four skills (Table 3); and (c) the actual use of the skill (Table 4).

Median agreement scores on the type of phone call were .88 for Telemarketer "D." (range = .50 to 1.00); .83 for Telemarketer "M." (range = .65 to .91); and .88 for Telemarketer "T." (range = .63 to 1.00).

The agreement ratios for the opportunities to use each skill are presented in Table 3. The skills of "using the customer's name" and "stating benefits" were omitted because, by definition, the telemarketers always had these opportunities.

Table 3

Median Interobserver Agreement Ratios on Skill Opportunities

	Ask for Sale	Cross-Sell
Telemarketer "D."	.88	.91
Telemarketer "M."	.94	.95
Telemarketer "T."	.93	.93

Agreement ratios for skill use within each phase are depicted in Table 4. Agreement ratios were not affected by any particular phase of the study; however, the low frequency of skill observation within each phase lowers reliability of the data.

Table 4

Median Interobserver Agreement Scores on Occurrences by Experimental Phase

Skill and Phase	Telemarketer D.	Telemarketer M.	Telemarketer T.
Customer Name			
baseline	no observation	.50 (1)	no observation
feedback	1.00 (2)	.90 (5)	.71 (4)
IRA loan & feedback	.92 (5)	.92 (5)	.90 (7)
IRA loan & rein.	no observation	.92 (1)	.83 (1)
reinforcement only	1.00 (1)	.96 (4)	.55 (2)
Ask for Sale			
baseline	no observation	.40 (5)	no observation
feedback	.75 (2)	.50 (1)	1.00 (2)
IRA loan & feedback	.80 (4)	.76 (6)	.69 (8)
feedback only	1.00 (1)	.77 (4)	.75 (4)
Cross-Sell			
baseline	.75 (3)	.75 (5)	.88 (2)
baseline & IRA loan	.50 (3)	.71 (1)	no observation
feedback	.76 (2)	.66 (4)	.67 (3)
IRA loan & rein.	no observation	.67 (1)	1.00 (1)
reinforcement only	.33 (1)	.60 (3)	.61 (4)
Benefits			
baseline	1.00 (3)	.75 (7)	.91 (4)
IRA loan	.67 (5)	.77 (6)	.69 (8)
baseline	.78 (1)	.78 (4)	.73 (3)

Note: The number in parentheses represents the total number of interobserver agreement observations in the phase. There was one observer per telemarketer.

Results and Discussion

Overall Results for Telemarketers "D.," "M.," and "T."

Figures 1, 2, and 3 depict the multiple baselines for Telemarketers "D.," "M.," and "T." Planned interventions (feedback and reinforcement) are denoted by the dotted lines between phases, whereas the unplanned intervention (the IRA loan program) is boxed off by solid lines. Each data point represents a 3-day period.

The purpose of a multiple baseline design is to (a) determine whether a change in behavior occurred at the time of the intervention; (b) determine whether any observed change from A to B can be replicated at another point in time (i.e., leg), increasing the construct validity of the treatment; and (c) rule out threats to validity by ensuring that changes did not occur in the corresponding data points for untreated, control behaviors.

The three multiple baselines presented here can partially permit these determinations by visual inspection alone. The variability of the data in some phases, however, is problematic to visual inspection. Because of this variability, and because of the possibility of naturally occurring cyclical variation, interrupted time series analysis was applied. We thus have two methods to determine the effects of the interventions. The first is visual inspection of the data; the second is statistical analysis. The focus of this section will be to review the graphic data for each intervention and determine whether statistical analysis supports the conclusion that is derived visually. Each intervention (feedback, reinforcement, and IRA loan program) will be discussed for each skill.

The Gottman-Williams time-series analysis program yields two intercepts and slopes for each series run. The "large" intercept and slope are the actual parameters for the data; the "small" intercept and slope are the corrected estimates for the time series. Following the example provided by Gottman (1981, p. 363), the regression lines drawn

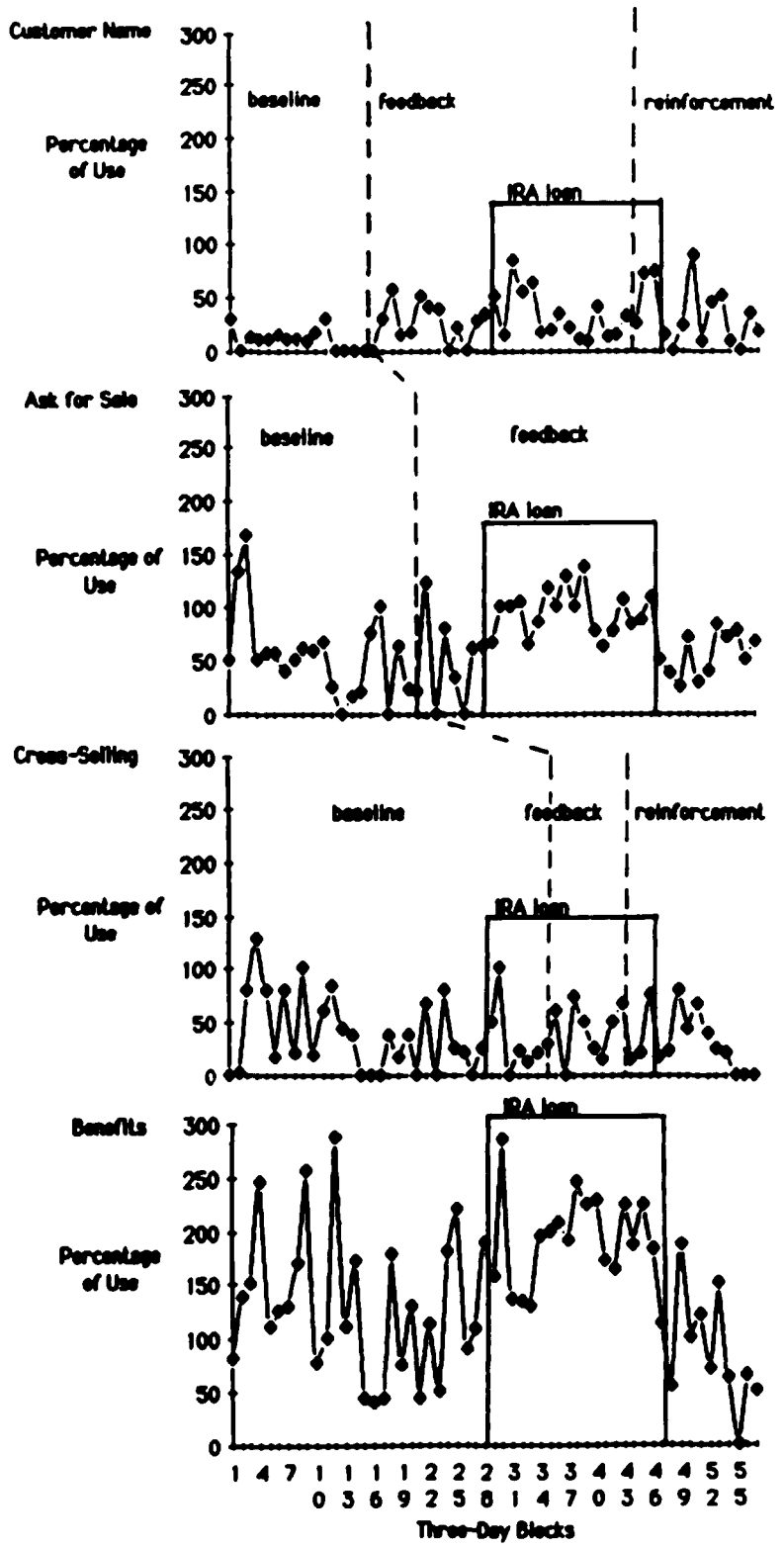


Figure 1. Feedback, reinforcement, and IRA loan program effects for Telemarketer "D."

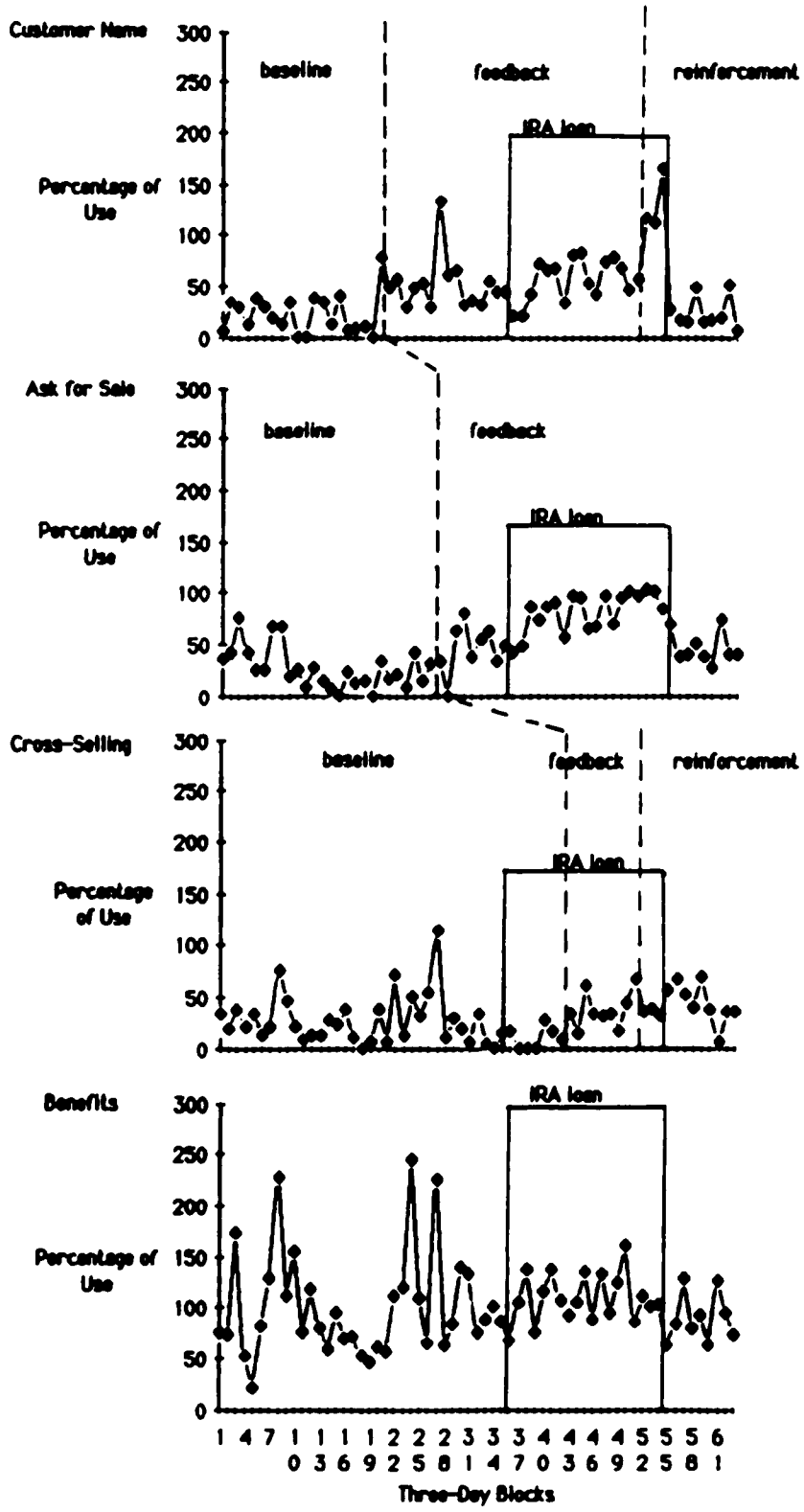


Figure 2. Feedback, reinforcement, and IRA loan program effects for Telemarketer T1.*

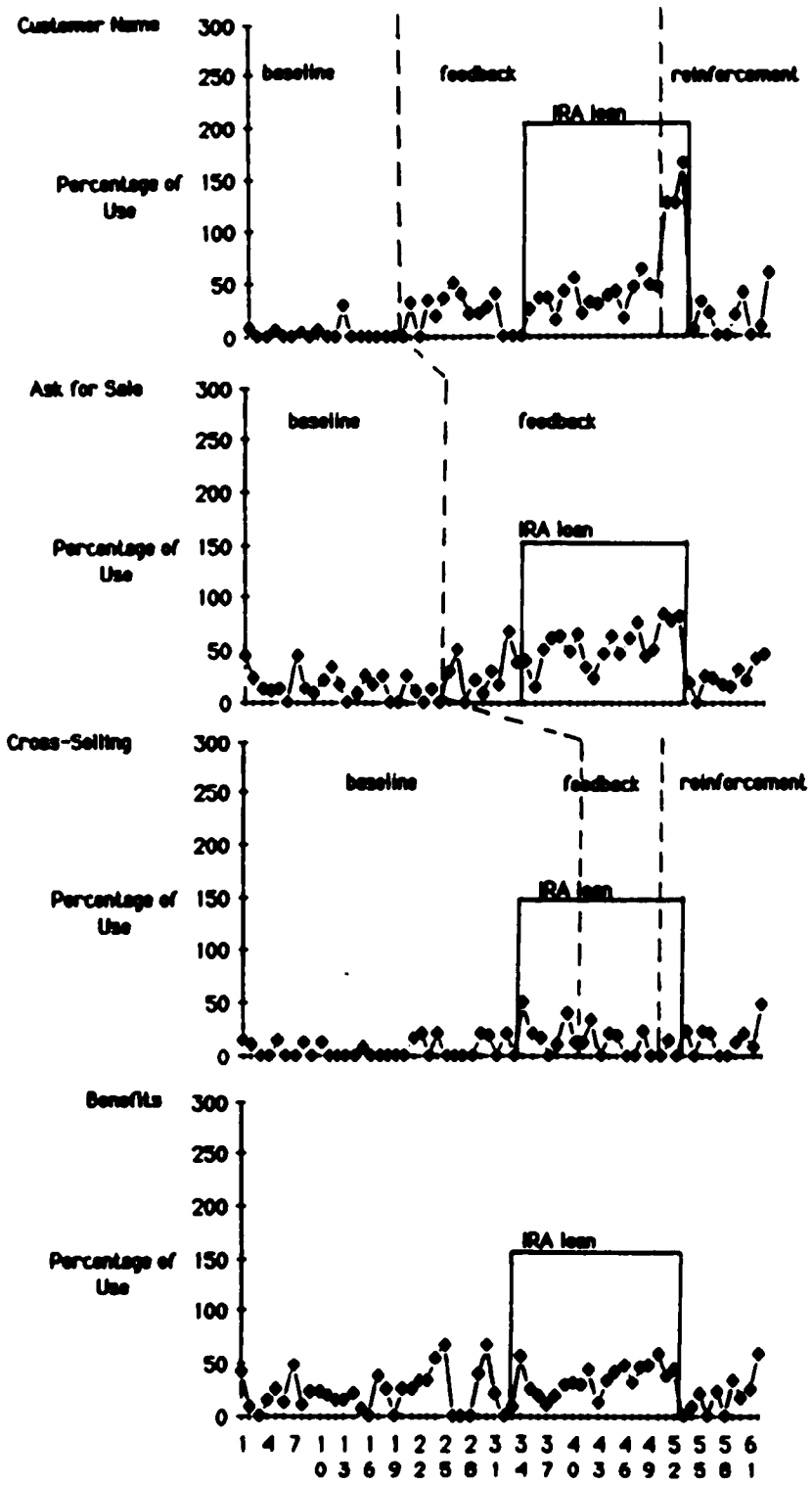


Figure 3. Feedback, reinforcement, and IRA loan program effects for Telemarketer "T."

in the figures included here were computed with the standard formula using the "large," uncorrected (i.e., actual) parameters. The results reported in the tables are the corrected estimates for intercept and slope. This difference must be considered when comparing the parameters as described in the tables with the regression lines as drawn in the figures.

Feedback Effects

Feedback effects were tested for each skill by comparing the feedback phase to baseline. The data were statistically analyzed twice: once ignoring the IRA loan points as though they were not an influence (because this was an unplanned comparison), and once as though the IRA loan phase were a planned comparison. All phases were analyzed both including and excluding the IRA program from consideration. Only contiguous data points between phases were analyzed.

Using the Customer's Name

Figure 4 depicts the three graphs of "using the customer's name" for Telemarketers "D.," "M.," and "T.," including all IRA data points. Visual inspection of each graph reveals level increases at the point of intervention. For each telemarketer, the slope in the baseline phases was close to zero or negative. There were no significant changes in slope when feedback was introduced; however, there were significant changes in intercepts (see Table 5). The low or negative slopes in the baseline phases would not lead one to predict increases in intercepts obtained in treatment. Thus, both the visual inspection and the statistical analysis suggest that there was a treatment effect for this skill for all three participants.

This conclusion does not change when baseline phases are compared to feedback up to the point of the IRA loan program (Figure 5 and Table 6). There are clear level increases

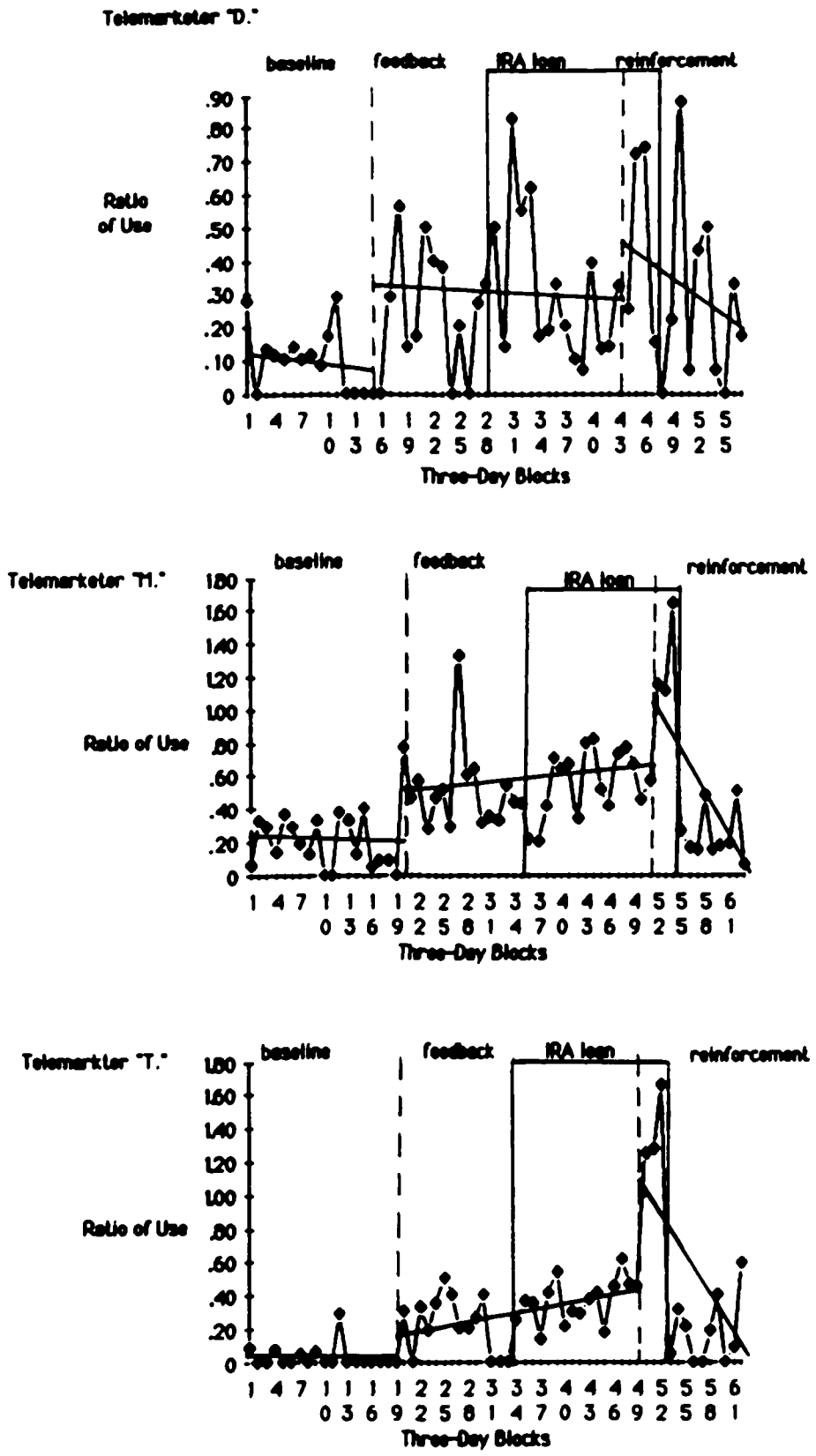
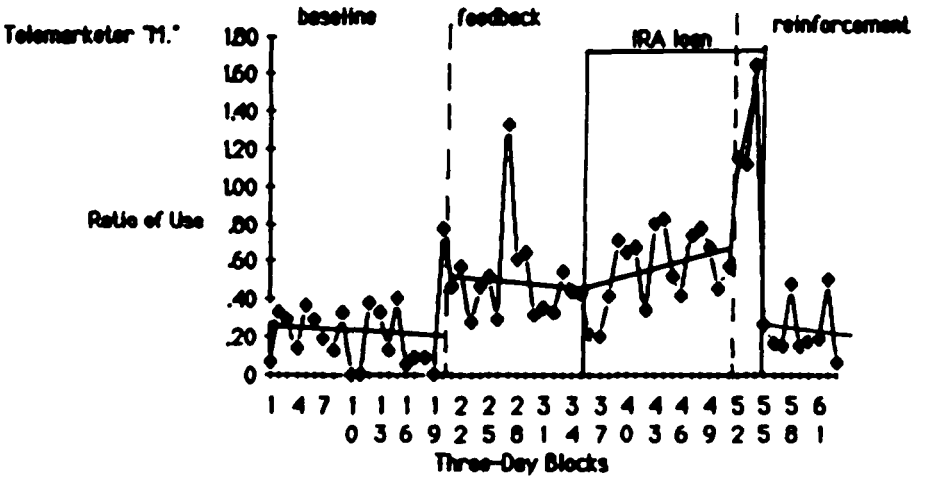
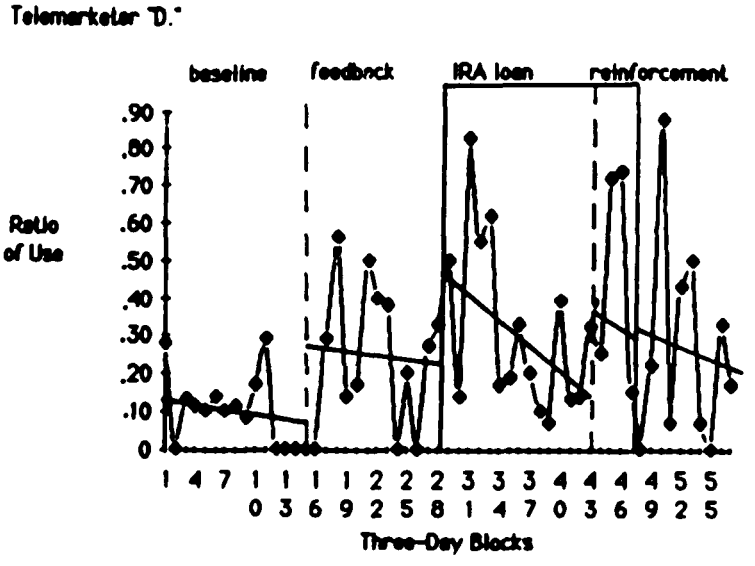


Figure 4. Effects of feedback and reinforcement for all telemarketers on "using the customer's name."





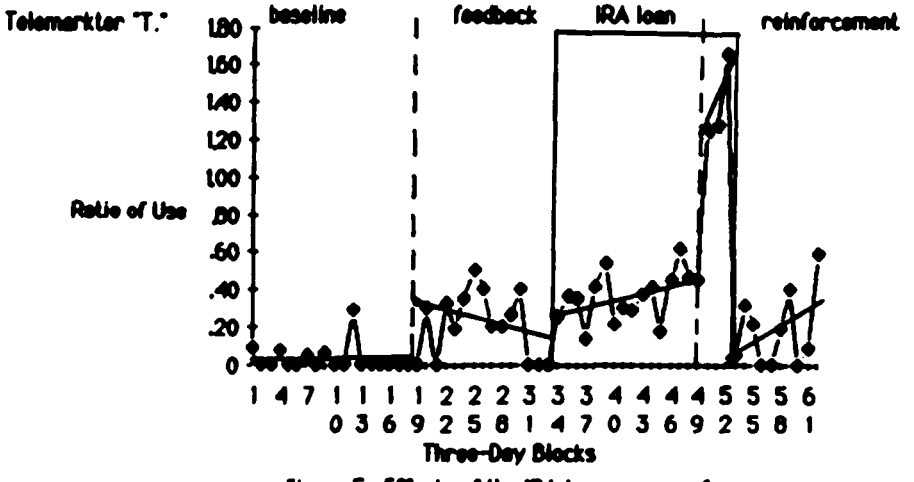


Figure 5. Effects of the IRA loan program for all telemarketers on "using the customer's name."

Table 5

Intercepts and Significance Tests Between Baseline and Feedback

For "Using the Customer's Name" - Includes all IRA and Non-IRA Data Points

Telemarketer	Intercept in Baseline	Intercept in Feedback	t	df	Significance Level
"D."	.097	.252	5.270	37	$p < .01$
"M."	.233	.477	7.413	45	$p < .01$
"T."	.026	.231	9.039	32	$p < .01$

Table 6

Intercepts and Significance Tests Between Baseline and Feedback

For "Using the Customer's Name" - Includes Points to the Beginning of the IRA Program

Telemarketer	Intercept in Baseline	Intercept in Feedback	t	df	Significance Level
"D."	.122	.247	4.462	22	$p < .01$
"M."	.272	.649	8.814	29	$p < .01$
"T."	.025	.307	12.090	27	$p < .01$

between baseline and feedback phases. While there were no significant changes in slope, the significant increases in intercepts, along with visual inspection, support the notion of a feedback effect for this skill, because one would not predict changes in intercept based on the low or negative slopes obtained in baseline.

Asking for the Sale

Figure 6 depicts "asking for the sale" for each telemarketer. Table 7 describes the intercept differences for this skill when all data points were used in the analysis (including IRA loan points); Table 8 describes the intercept differences when the IRA loan program points were excluded from the analysis in the feedback phase.

First consider the analysis in which all data points were included (Table 7). Graphically, for each telemarketer, the slope of the function describing the use of this skill during baseline sessions was negative, while it was non-negative post-intervention; however, these apparent changes in slope were nonsignificant. Analysis of changes in intercept are consistent with an increase in the use of this skill for two of the three telemarketers ("M." and "T.") The graphed data for Telemarketer "D." also suggest that asking for the sale increased during the feedback phase; however, due to a decrease in intercepts from baseline to feedback, this interpretation is not supported by the time series analysis.

Next consider this skill with the IRA loan program points excluded from the analysis (Table 8). The graphed data show no change for Telemarketer "D." at the point of the feedback intervention, given the variability in data prior to and after intervention. There was no significant change in slope and a significant decrease in intercept for this participant. Thus neither the graphed data nor the statistical analysis reflect a treatment effect without the influence of the IRA loan program.

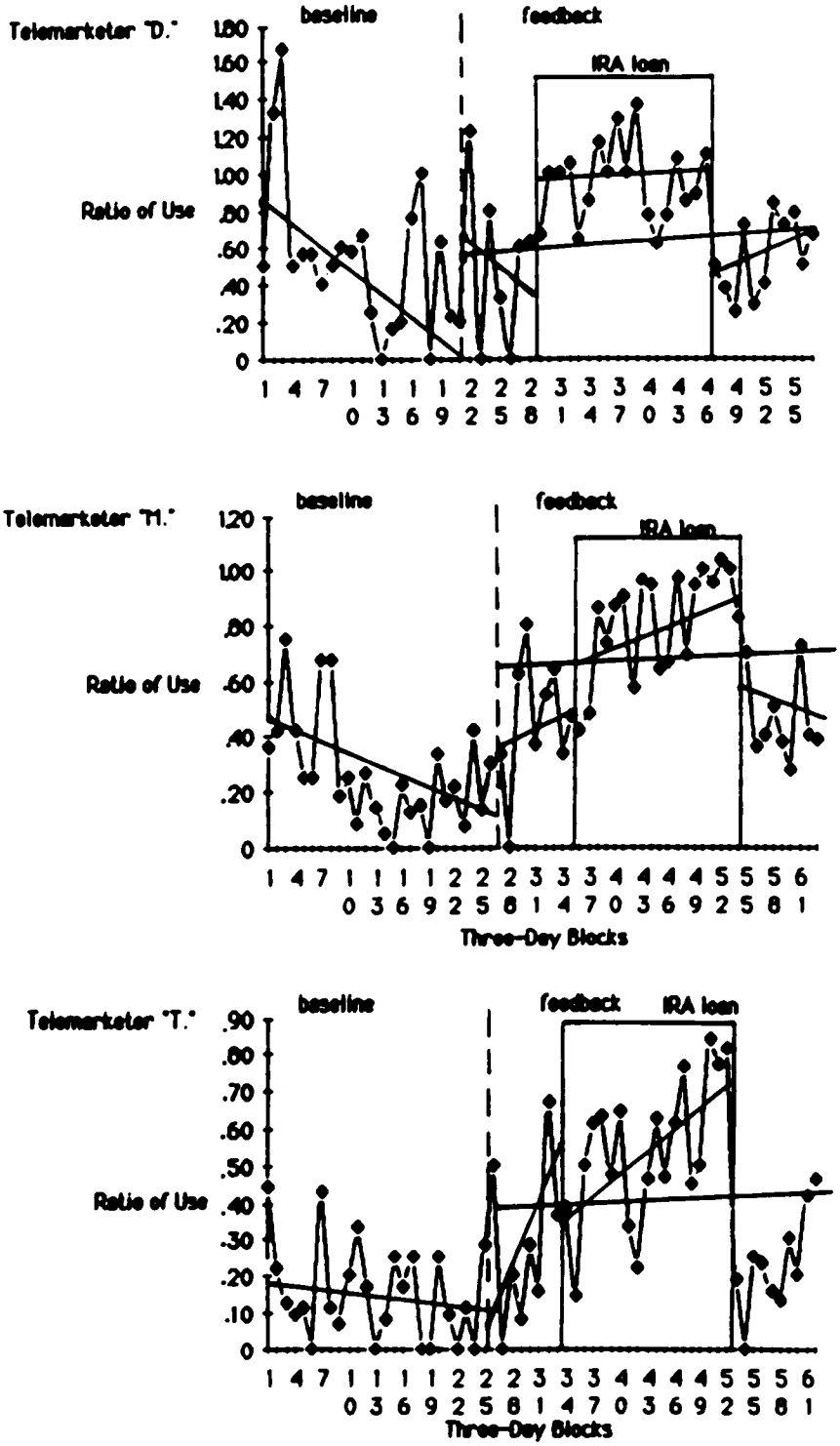


Figure 6. Effects of feedback and IRA loan program for all telemarketers on "asking for the sale."

Table 7

Intercepts and Significance Tests Between Baseline and Feedback

For "Asking for the Sale" - Includes all IRA and Non-IRA Data Points

Telemarketer	Intercept in Baseline	Intercept in Feedback	t	df	Significance Level
"D."	1.156	.656	9.632	51	$p < .01$
"M."	.222	.336	4.137	57	$p < .01$
"T."	.084	.245	6.507	56	$p < .01$

Table 8

Intercepts and Significance Tests Between Baseline and Feedback

For "Asking for the Sale" - Includes Points to the Beginning of the IRA Program

Telemarketer	Intercept in Baseline	Intercept in Feedback	t	df	Significance Level
"D."	1.162	.826	4.311	22	$p < .01$
"M."	.394	.323	1.929	29	n.s.
"T."	.208	.181	.969	27	n.s.

The graphed data for Telemarketers "M." and "T." are ambiguous due to the variability of the data and the small number of points in the feedback phase. There were no significant changes in either slope or intercept. Thus, no treatment effect is indicated.

In conclusion for this skill, it appears that the IRA loan program exerted a powerful effect which accounts for the observed changes between baseline and feedback. Without the

IRA loan program included in the analysis, no treatment effect appears.

Cross-Selling

Figure 7 depicts the graphed data on cross-selling for the three telemarketers. Table 9 describes the statistical analysis between baseline and feedback including all IRA loan program points in the baseline condition. Table 10 describes the results of comparing only the IRA loan program points in baseline to the feedback phase. Because the IRA loan program completely overlapped with the feedback condition, there was no way to separate these two interventions for purposes of analysis.

Figure 7 for Telemarketer "D." shows a negatively sloping baseline followed by an increase in slope during feedback. This pattern suggests that feedback resulted in an increase in cross-selling for this participant. However, the statistical analysis (Table 9) does not support this conclusion. The apparent change in slope from negative to positive was not statistically significant, and the level of skill use, and defined by intercepts, actually decreased. Therefore no treatment effect is indicated.

Let us contrast these results for Telemarketer "D." with those obtained for her when only the IRA program points are used in baseline for comparison to feedback (Table 10, Figure 7). While there is a much variability in both phases, the slope of the baseline phase appears to be flat, while the slope of the feedback phase appears to be positive. These apparent changes in slope, however, were not statistically significant. There was a significant increase in intercepts. Due to the extreme variability of the data in both phases and no clear demarcation of a treatment effect, no treatment effect is indicated.

Telemarketer "M." shows a different pattern (Table 9, Figure 7). While the slope of the baseline phase appears negative and the slope of treatment appears positive, statistical analysis reveals no significant change in the slopes; nor is there a change in the

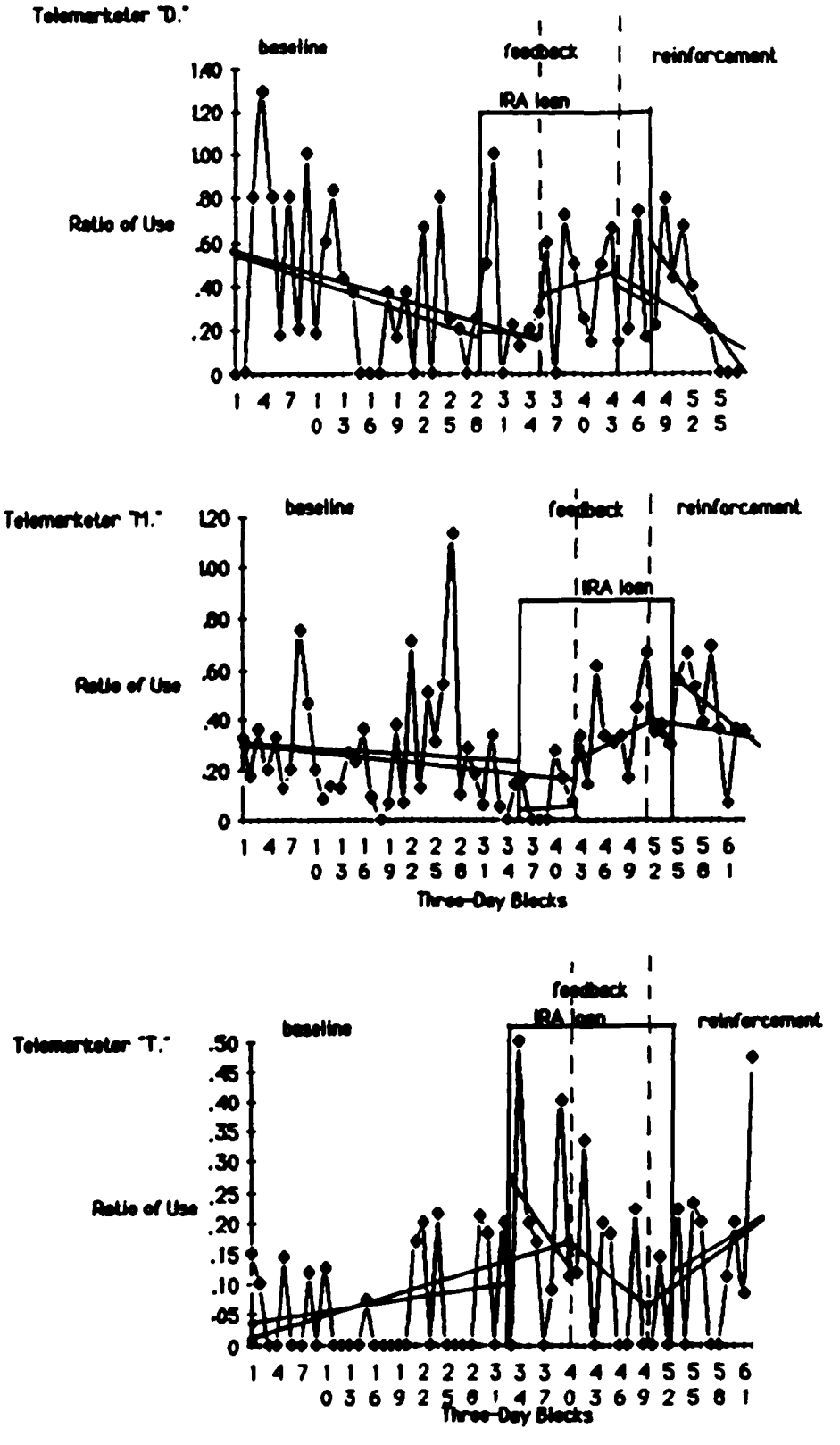


Figure 7. Effects of feedback, reinforcement, and IRA loan program for all telemarketers on "cross-selling."

Table 9

Intercepts and Significance Tests Between Baseline and Feedback

For "Cross-Selling" - Includes IRA and Non-IRA Data Points in Baseline

Telemarketer	Intercept in Baseline	Intercept in Feedback	t	df	Significance Level
"D."	.504	.347	3.085	37	$p < .01$
"M."	.297	.245	1.555	45	n.s.
"T."	-.004	.231	13.959	43	$p < .01$

intercepts. Visual inspection does not suggest a treatment effect for feedback, but rather a decrease in performance during baseline and a reverse (increase) beginning at approximately data point number 34. Because neither visual inspection nor statistical analysis suggested a treatment effect, one will not be assumed.

When only the IRA loan points in baseline are compared to feedback for Telemarketer "M." (Table 10 and Figure 7) a significant increase in intercepts occurs. Both visual inspection and statistical analysis reveal a much lower intercept in the baseline phase, which would have led to a significant increase to occur at feedback.

Telemarketer "T." reveals a different pattern of results (Table 9 and Figure 7). Visual inspection suggests that feedback decreased the use of this skill. While there was a significant increase in intercepts, the slope in the baseline phase is positive while the slope in the feedback phase is negative. The increase in intercepts could therefore be explained by the change in direction (though nonsignificant) of the slopes.

When only the IRA points in baseline are considered (Table 10 and Figure 7), there

Table 10

Intercepts and Significance Tests Between Baseline and Feedback

For "Cross-Selling" - Includes Only IRA Program Points in Baseline

Telemarketer	Intercept in Baseline	Intercept in Feedback	t	df	Significance Level
"D."	.028	.493	6.315	9	$p < .01$
"M."	.024	.308	6.544	10	$p < .01$
"T."	.193	.234	.951	10	n.s.

was no significant change in intercept or slope. Visual inspection suggests that any changes in cross-selling performance are due to the IRA loan (to be discussed later). Thus, no treatment effect is indicated.

In summary for the three telemarketers, while statistical analyses sometimes suggested the possibility of a feedback effect, visual inspection of the data do not reveal a treatment effect for this skill.

Reinforcement Effects

Reinforcement effects were analyzed by comparing the pattern of data in the feedback phase to the pattern of data in the reinforcement phase which followed for using the customer's name and cross-selling. Statistical analyses were considered in conjunction with the visual inspection. Because of the interference of the IRA loan program, there was a slightly different pattern of analysis for the two skills. We shall first discuss using the customer's name, feedback as compared to reinforcement, including all data points in both phases (disregarding the possible influence of the IRA program). Table 11 describes

these results. We will then discuss the segments of the data which include only the IRA program points in both phases (Table 12) for this skill. Because the IRA program completely overlapped with feedback for cross-selling, we will discuss feedback as compared to reinforcement without regard to the IRA program (Table 13) and save discussion of the IRA loan program effects for the next section.

Using the Customer's Name

Refer once again to Figure 4 and to Table 11 for the visual and statistical results of feedback compared to reinforcement for using the customer's name for the three telemarketers. The graph for Telemarketer "D." (Figure 4) indicates a zero or slightly negative slope during the feedback condition, and a sharply negative slope in the reinforcement phase. There also appears to be a momentary enhancing effect at the onset of the reinforcement phase, which is not maintained. The statistical analyses (Table 11) indicated that the apparent change in the slope was not significant, while the increase in intercept was significant.

A similar picture emerges when we compare the effects of feedback to reinforcement when only the IRA points are evaluated (Figure 5 and Table 12). The pattern of results is similar to that just discussed, however, there is no significant difference in intercept or slope between the last portion of feedback and the first portion of data points for reinforcement.

The graph for Telemarketer "M." also suggests an initial enhancement by reinforcement of using the customer's name (see Table 11 and Figure 4). However, when the IRA program ended, the level of skill use dropped sharply. These results are reflected in the statistical analyses by significant changes in both intercept and slope. Figure 5 and Table 12 reveal significant increases in both intercept and slope when only the IRA points

in both phases are compared. We thus conclude that there was a treatment effect which was not maintained when the IRA loan program ended.

The results for Telemarketer "T." are quite similar to those of Telemarketer "M." (see Figure 5). Both graphic and statistical analysis support the notion of an initial reinforcement effect which was not maintained when the IRA loan program ended. There was a significant increase in intercepts regardless of whether the IRA program points were included (Tables 11 and 12) and there was a signficant increase in slope when only the IRA points were analyzed (Table 12).

Table 11

Intercepts and Significance Tests Between Feedback and Reinforcement for "Using the Customer's Name" - Includes IRA and Non-IRA Data Points in Both Phases

Telemarketer	Intercept (and slope) in Feedback	Intercept (and slope) in Reinforcement	t	df	Significance Level
"D."	.336	.485	3.775	36	$p < .01$
"M."	.429	1.008	12.296	37	$p < .01$
slope:	.004	-0.087	1.929	37	$p < .01$
"T."	.144	.883	15.797	37	$p < .01$

Table 12

Intercepts and Significance Tests Between Feedback Phases and Reinforcement for "Using the Customer's Name" - Includes Only IRA Program Points in Both Phases

Telemarketer	Intercept (and slope) in Feedback	Intercept (and slope) in Reinforcement	t	df	Significance Level
"D."	.575	.547	.419	13	n.s.
"M."	.519	.634	2.113	13	$p < .05$
slope:	.003	.266	4.845	13	$p < .01$
"T."	.335	1.031	19.322	13	$p < .01$
slope:	.014	.283	7.466	13	$p < .01$

Cross-Selling

Refer to Table 13 and Figure 7 to compare the effects of feedback to reinforcement for cross-selling. This analysis includes all the data points in the feedback phase and all the data points in the reinforcement phase. The graph for Telemarketer "D." reveals a slope close to zero or positive in the feedback phase, which was not significantly different from the negative slope in the reinforcement phase. The change in the direction of slopes, however, could account for the significant increase in intercepts regardless of whether there was truly a treatment effect. Visual inspection of the data reveals no treatment effect whatsoever. Therefore, no treatment effect is indicated.

Table 13

Intercepts and Significance Tests Between Feedback and Reinforcement for "Cross-Selling" - Includes all IRA and Non-IRA Data Points in Reinforcement

Telemarketer	Intercept in Feedback	Intercept in Reinforcement	t	df	Significance Level
"D."	.260	.528	3.912	16	$p < .01$
"M."	.254	.485	4.750	15	$p < .01$
"T."	.316	.046	8.871	16	$p < .01$

The pattern of results for Telemarketer "M." is similar to that of Telemarketer "D." (see Table 13 and Figure 7). The graphed data suggest that reinforcement led to a gradual decline in the use of this skill. A positive slope in feedback is not significantly different from the negative slope in reinforcement, although there is a significant increase in the intercepts. Because there is no graphic support for the notion of a treatment effect, no reinforcement effect is indicated.

The results for Telemarketer "T." (Figure 7) are exactly the opposite from those of the other telemarketers. The graphed data suggest that reinforcement reversed the downward trend in skill use. The statistical analyses, however, do not support this, due to nonsignificant slope changes and a significant decrease in intercepts. Therefore no treatment effect is indicated.

In summary for the effect of reinforcement, there is evidence that reinforcement had a temporary impact on using the customer's name, which ended with the IRA program. There is no evidence that there was any effect of reinforcement on cross-selling.

IRA Loan Program Effects

The effects of the IRA loan program were analyzed by visual inspection and by statistically comparing the IRA phase to the phase immediately preceding it for each of the four skills (Tables 14, 15, 16 and 17), and then again to phase immediately following it for each of the four skills (Tables 18, 19, 20, and 21).

Using the Customer's Name

Let us begin with "using the customer's name" (Table 14 and Figure 5) for Telemarketer "D." There was a significant increase in intercept, but no significant change in slope, between the feedback phase without the IRA loan and the beginning of the IRA loan (also in the feedback phase). Visually, there appears to be a temporary increase in the level of skill use when the IRA program went into effect, followed by a decrease. Therefore an effect of the IRA loan program is indicated.

No effect appears, however, for Telemarketer "M." (Table 14 and Figure 5). There are no significant changes in intercept or slope between feedback and the IRA program, nor

---

Table 14

Intercepts and Significance Tests Between Feedback and the IRA Loan Program for "Using the Customer's Name"

Telemarketer	Intercept in Feedback	Intercept in IRA Loan Phase	t	df	Significance Level
"D."	.396	.565	4.090	22	$p < .01$
"M."	.548	.466	1.690	25	n.s.
"T."	.301	.244	1.879	24	n.s.

---

can any difference can be detected visually. Therefore, no effect is indicated.

The visual analysis for Telemarketer "T." suggests an increase in skill use due to a negative slope in feedback and a positive slope with the IRA program (Figure 5). The overall level of skill use also appears somewhat higher due to no zero occurrences in the IRA phase. Nevertheless, there is no statistical difference between the two phases (Table 14) to support the visual analysis, and therefore no IRA program effect is indicated.

Asking for the Sale

There are inconclusive results for the IRA loan program in "asking for the sale" across the three telemarketers (Table 15 and Figure 6). Graphic analysis for Telemarketer "D." (Figure 6) suggests an IRA loan effect due to an overall level increase in skill use, and a negative slope in feedback and a positive slope in the IRA phase. In spite of this compelling graphic evidence, statistical analysis does not support these conclusions. There were no signficant changes in either slope or intercept. Therefore the effect of the IRA loan cannot be determined with certainty in this case.

Table 15

Intercepts and Significance Tests Between Feedback and the IRA Loan Program for "Asking for the Sale"

Telemarketer	Intercept in Feedback	Intercept in IRA Loan Phase	t	df	Significance Level
"D."	.605	.682	.945	20	n.s.
"M."	.470	.907	6.552	21	$p < .01$
"T."	-.039	.361	11.200	21	$p < .01$

The graphic results for Telemarketer "M." (Figure 6) also suggest a possible treatment effect due to an increase in overall level of skill use during the IRA phase. On closer inspection, however, it appears that the performance may be part of an overall increase in skill use that began at approximately data point number 14. The significant intercept change does not support the graphic analysis, and therefore the exact effect of the IRA loan program cannot be determined.

The graphic results for Telemarketer "T." (Figure 6) also do not suggest an effect of the IRA loan program, but rather a steadily increasing trend from approximately data point number 22. Nevertheless, the statistical analysis reveals a significant increase in intercepts between feedback and the IRA program, with no significant change in slope (Table 15). The results are again inconclusive with regard to the IRA program.

Cross-Selling

The preceding phase of the IRA loan program in cross-selling was baseline (Table 16). These analyses will refer to the uncontaminated baseline and the baseline plus the

---

Table 16

Intercepts and Significance Tests Between Baseline and the IRA Loan Program for "Cross-Selling"

Telemarketer	Intercept in Baseline	Intercept in IRA Loan Phase	t	df	Significance Level
"D."	.503	.173	5.531	29	$p < .01$
"M."	.263	.047	5.539	36	$p < .01$
"T."	.031	.365	19.465	34	$p < .01$

IRA loan program. The feedback phase will not be included. Refer to Figure 7 for graphic analysis of these phases.

Statistical analysis for Telemarketer "D." reveals a signficant decrease in the intercepts between the uncontaminated and contaminated baselines in cross-selling (Table 16 and Figure 7). The slope in the uncontaminated baseline is negative, while the slope in the contaiminated (IRA) baseline is close to zero. These apparent differences are nonsignificant. Graphically, there is no evidence that cross-selling decreased with the onset of the IRA loan. Therefore, no IRA effect will be assumed.

Telemarketer "M." shows a similar pattern (Table 16 and Figure 7). There were no increases in skill use associated with the onset of the IRA loan that would suggest an IRA loan effect. There was a significant decrease in intercepts and no signifcant changes in slope. Because neither graphic nor statistical evidence suggested a treatment effect, no effect will be assumed.

It appears that the IRA loan program had an impact on the use of cross-selling (Figure 7) for Telemarketer "T." due to a sharp level increase. The statistical analysis (Table 16) revealed a significant increase in intercepts, with no significant change in slope. Because both the visual and the statistical analyses suggest an effect, it will be assumed that there was an impact of the IRA loan on the use of this skill.

Stating Benefits

In comparing the first baseline of "stating benefits" to the IRA loan program for Telemarketer "D." (Figure 8), the overall level of skill use increased with the onset of the program. Statistically, however, there were no significant changes in intercept or slope (Table 17), so that the statistical analyses do not support the graphic analysis.

According to graphic analysis for Telemarketer "M." (Figure 8), the IRA loan

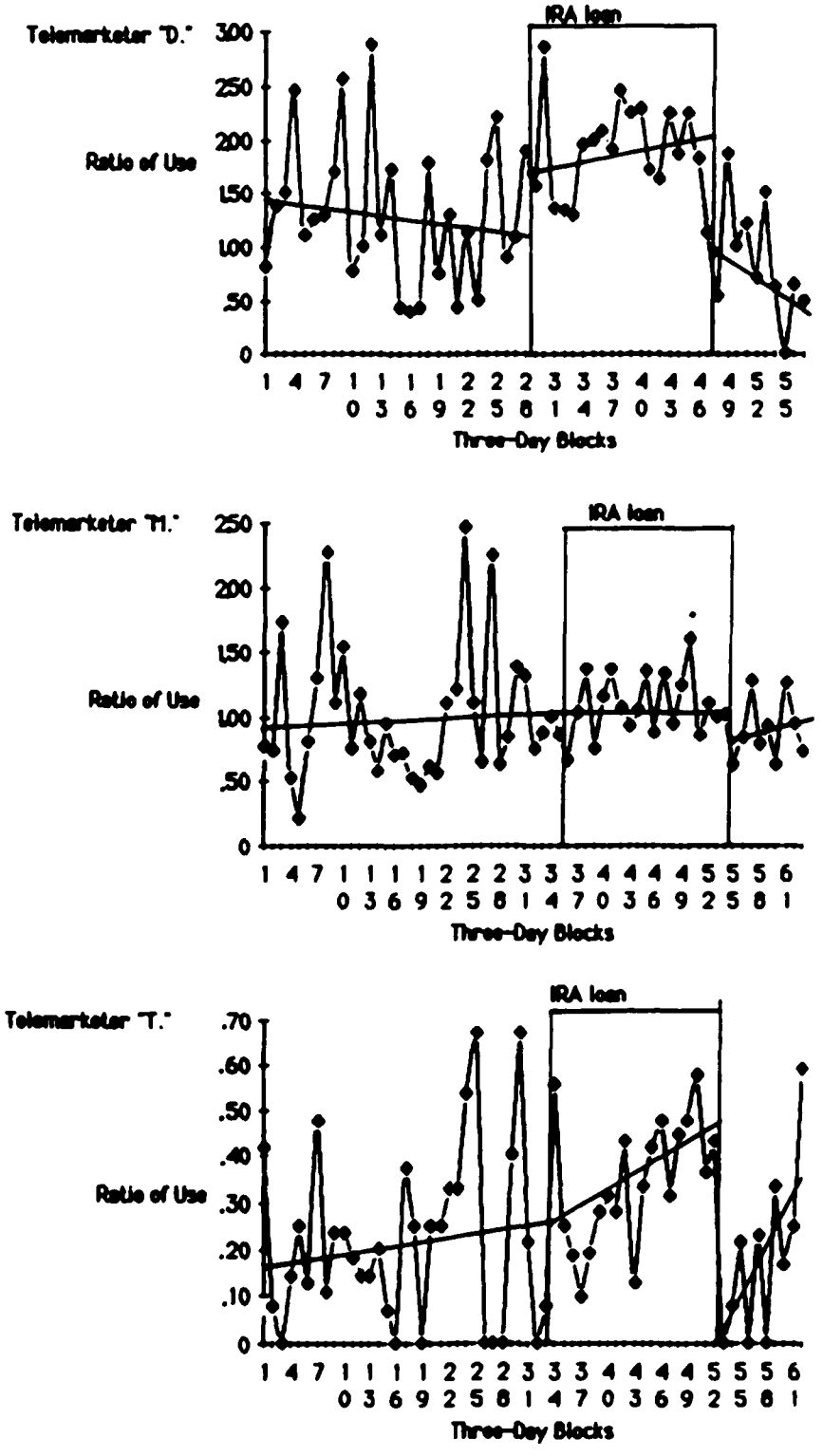


Figure 8. Effects of the IRA loan program for all telemarketers on "describing benefits."

program did not appear to increase (or decrease) the overall level of skill use. Statistical analysis revealed no significant change in slope or intercept. Thus we will conclude that there is no impact of the onset of the IRA loan program on the use of this skill.

Table 17

Intercepts and Significance Tests Between First Baseline and the IRA Loan Program for "Benefits"

Telemarketer	Intercept in Baseline	Intercept in IRA Loan Phase	t	df	Significance Level
"D."	1.681	1.898	1.990	41	n.s.
"M."	.916	.971	.828	48	n.s.
"T."	.139	.208	2.787	46	$p < .01$

The graph for Telemarketer "T." (Figure 8) reveals an apparent change in skill use through an overall level increase and an increasing slope over time during the IRA phase. While there was a signficant increase in intercept, however, the lack of a significant slope change when they are both positive suggests the effects seen during the IRA loan phase could be predicted from the functions during baseline. We thus cannot conclude that the IRA loan program increased the use of this skill for this telemarketer.

The next section will discuss the comparison of the IRA loan to contiguous phases which followed it for each of the four skills.

Using the Customer's Name

Refer to Table 18 and Figure 5 for "using the customer's name." Both the graphic depiction (Figure 5) and the statistical analysis (Table 18) suggest that there was no

Table 18

Intercepts and Significance Tests Between the IRA Loan Program in Reinforcement (A-Phase) and Reinforcement without IRA points (B-phase) for "Using the Customer's Name"

Telemarketer	Intercept (and slope) in A-Phase	Intercept in (and slope) in B-phase	t	df	Significance Level
"D."	.347	.326	.382	8	n.s.
"M."	.608	.269	4.755	6	$p < .01$
slope:	.537	-.004	7.571	6	$p < .01$
"T."	-.003[a]	-.124	2.156	7	$p < .05$

[a]Sum of autoregressive parameters was greater than .07; view with caution

impact of the end of the IRA program on skill use for Telemarketer "D." Telemarketers "M." and "T.," however, did exhibit sharp drops in skill use (Figure 5) when the IRA program ended. These conclusions are supported by significant intercept changes, and, for Telemarketer "M.," a significant change in slope.

Asking for the Sale

The end of the IRA loan phase appeared to have an impact on all telemarketers by decreasing skill use, at least temporarily (Table 19 and Figure 6). The level of skill use for Telemarketer "D." decreased after the end of the IRA loan, which is supported by a significant decrease in intercepts (no signficant change in slope). Telemarketers "M." and "T." also exhibited decreases in skill use at the end of the program, as supported by a significant decrease in intercept (no change in slope). Interestingly, after a sharp drop in

Table 19

Intercepts and Significance Tests Between the IRA Loan Program and Feedback Phases Following it for "Asking for the Sale"

Telemarketer	Intercept in IRA Loan Phase	Intercept in Feedback Phase	t	df	Significance Level
"D."	.666	.521	2.105	23	$p < .05$
"M."	1.034	.716	5.604	22	$p < .01$
"T."	.304	.008	9.675	23	$p < .01$

skill use for Telemarketer "T.," there was a steady increase until the end of the study.

Cross-Selling

The graphic data on cross-selling (Figure 7) comparing the end of the IRA loan program during reinforcement and the reinforcement phase without the IRA program are quite difficult to interpret due to the very short phase (three data points) of the IRA-plus-reinforcement condition. There is no graphic evidence that the end of the program had any impact on the use of cross-selling. Statistically, Telemarketers "D." and "M." both had significant increases in intercepts (Table 20). Based on the graphic analysis, however, no effect of the end of the IRA program will be assumed.

Stating Benefits

Refer to Table 21 and Figure 8 for discussion of the impact of the end of the IRA program on "stating benefits." Graphically, the decrease seen for Telemarketer "D." appears to be part of a decreasing trend starting around data point number 45. There was, however, a significant decrease in intercept, which suggests an effect (Table 21).

Table 20

Intercepts and Significance Tests Between the IRA Loan Program and Reinforcement Phase Following it for "Cross-Selling"

Telemarketer	Intercept in IRA Loan Phase	Intercept in Reinforcement Phase	t	df	Significance Level
"D."	.411	.644	2.793	8	$p < .05$
"M."	.469	.699	3.275	6	$p < .01$
"T."	.229	.110	2.073	7	n.s.

Table 21

Intercepts and Significance Tests Between the IRA Loan Program and the Baseline Phase Following it for "Stating Benefits"

Telemarketer	Intercept in IRA Loan Phase	Intercept in Reinforcement Phase	t	df	Significance Level
"D."	2.135	1.407	5.516	23	$p < .01$
"M."	1.447	1.032	6.084	22	$p < .01$
"T."	.188	-.028	9.358	23	$p < .01$

Telemarketer "M." also shows very little change graphically with the end of the program (Figure 8) and showed a significant decrease in intercept. For these two telemarketers, no effect will be assumed.

Telemarketer "T." showed the clearest impact of the end of the program (Figure 8) with a sharp decrease in skill use. This is supported statistically with a significant

decrease in intercept. However, like her performance on asking for the sale (Figure 6), her performance increased again during the last few weeks of the study.

Analysis of the Independence of Sales Skills

Figures 9, 10, and 11 refer once again to the complete multiple baselines for each telemarketer. Solid vertical lines have been added to depict the control conditions for each intervention. "Asking for the sale" served as the control for feedback and reinforcement in "using the customer's name." "Cross-selling" served as a control for the feedback interventions in "using the customer's name" and "asking for the sale," and "benefits" served as the control condition for the three feedback interventions on the other skills and the reinforcement condition. If the skills were independent, there would be no change in performance in the control condition associated with the intervention in the treated skill. If, on the other hand, there were observed changes in the untreated skills associated with the intervention in the treated skill, independence cannot be assumed.

The analysis for independence followed the same format as the analysis for treatment effects. First the graphs were visually inspected for evidence of change. However, due to the variability of the data and the short length of the phases, it was not possible to determine whether the changes were due to nonindependence or were a result of within-subject variation. In order to attempt to answer the question, the intercepts and slopes for data segments in the control conditions that corresponded to the treatment phases of the treated skill were analyzed for change. If intercept and/or slope differences were found, it was assumed that the skills were not independent. The direction of change in response to treatment, if skills were not independent according to statistical analysis, cannot be gleaned from visual inspection, due to the variability in the data.

Analysis of Independence of Skills for Telemarketer "D."

Figure 9 and Table 22 refer to the analysis of independence of skills for Telemarketer "D." All intercepts changed significantly in the control conditions with the exception of "benefits" as a control for the feedback intervention in "asking for the sale." There was one significant slope change as well, which was in "benefits" as a control for feedback in cross-selling. On visual inspection, no clear trends emerge that would suggest nonindependence. However, on the basis of the statistical analysis, independence cannot be assumed.

Analysis of Independence of Skills for Telemarketer "M."

Figure 10 and Table 23 refer to the analysis of independence for Telemarketer "M." In her case, significant changes in intercept occurred in "asking for the sale" for both interventions in "customer name." Changes also occurred in "cross-selling" in response to the feedback intervention in "customer name." Graphically, the only data segment that appeared affected was in "cross-selling," for the feedback treatment for "using the customer's name." Because this was statistically supported, nonindependence will be assumed for this skill. While visual inspection does not necessarily indicate that independence has been violated, independence cannot be assumed for the skills, except for "benefits," because of the statistical evidence which suggests nonindependence.

Analysis of Independence of Skills for Telemarketer "T."

Figure 11 and Table 24 refer to the analysis of independence of skills for Telemarketer "T." Again, in most cases, there were significant intercept changes in the untreated skills for data segments which corresponded to interventions in treated skills (no significant slope changes). Exceptions were in "cross-selling," in response to the feedback intervention in "using the customer's name," and in "benefits" in response to the feedback

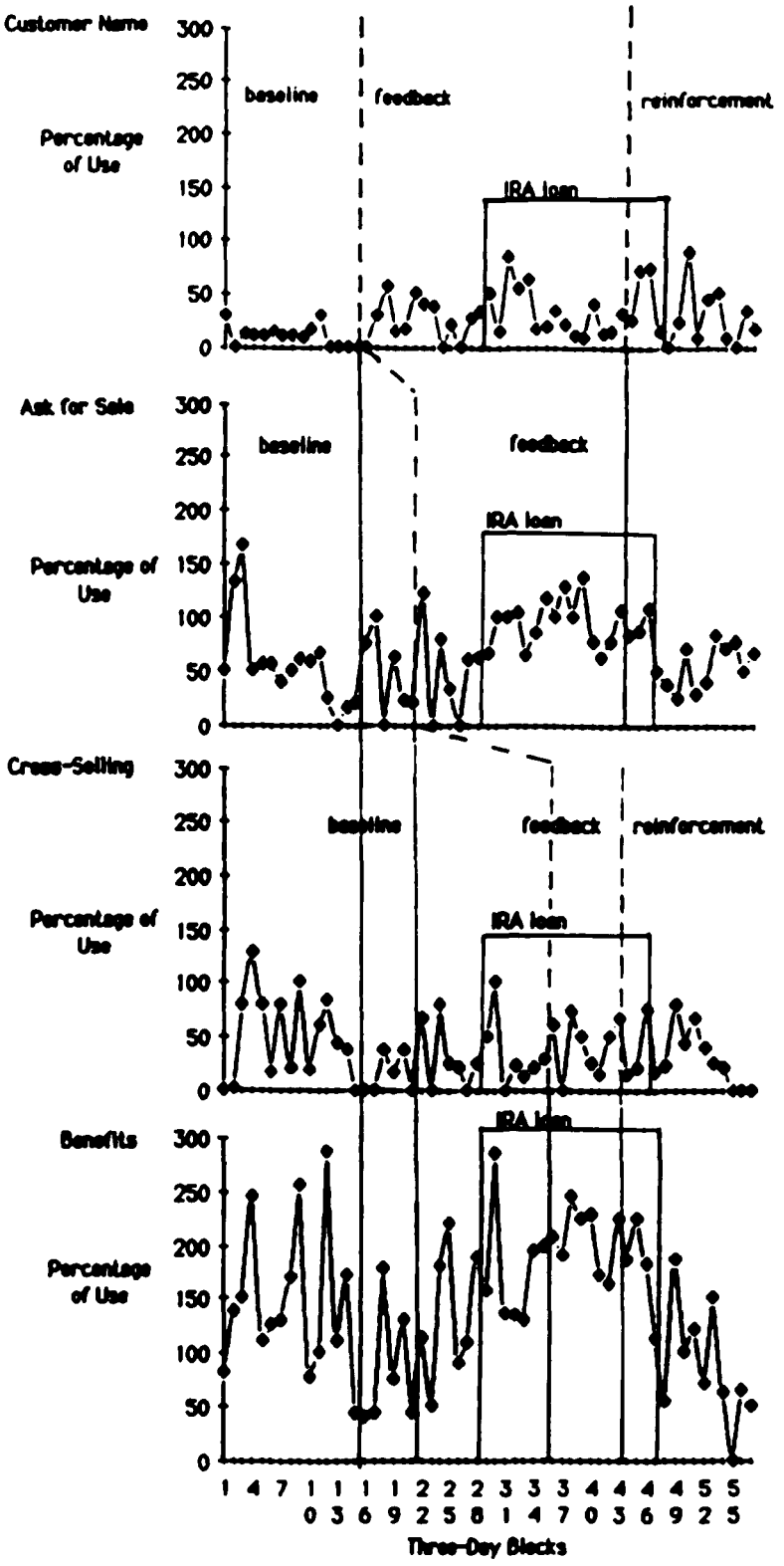


Figure 9. Control conditions for feedback, reinforcement, and IRA loan program for Telemarketer "D."

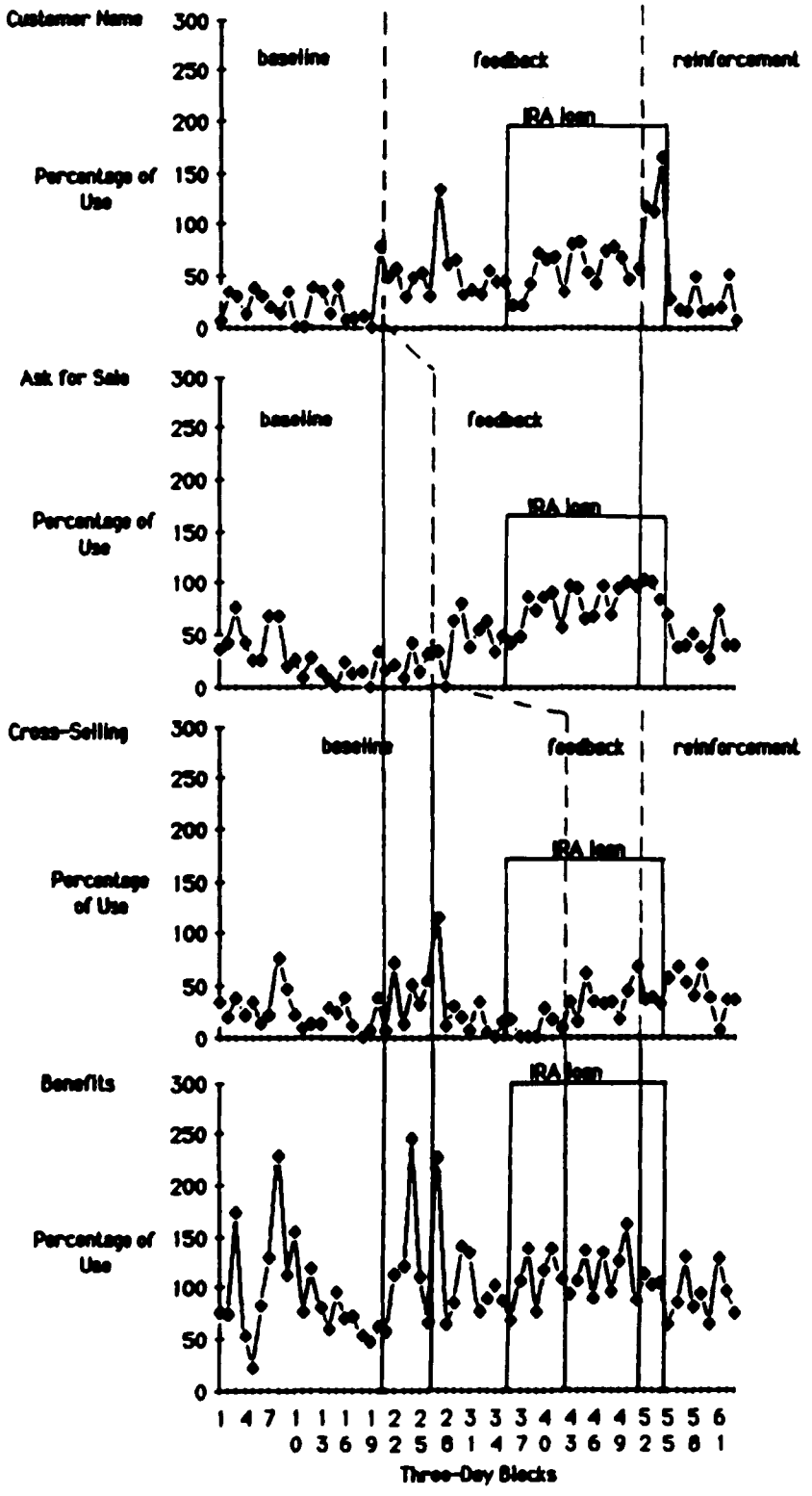


Figure 10. Control conditions for feedback, reinforcement, and IRA loan program for Telemarketer T1.*

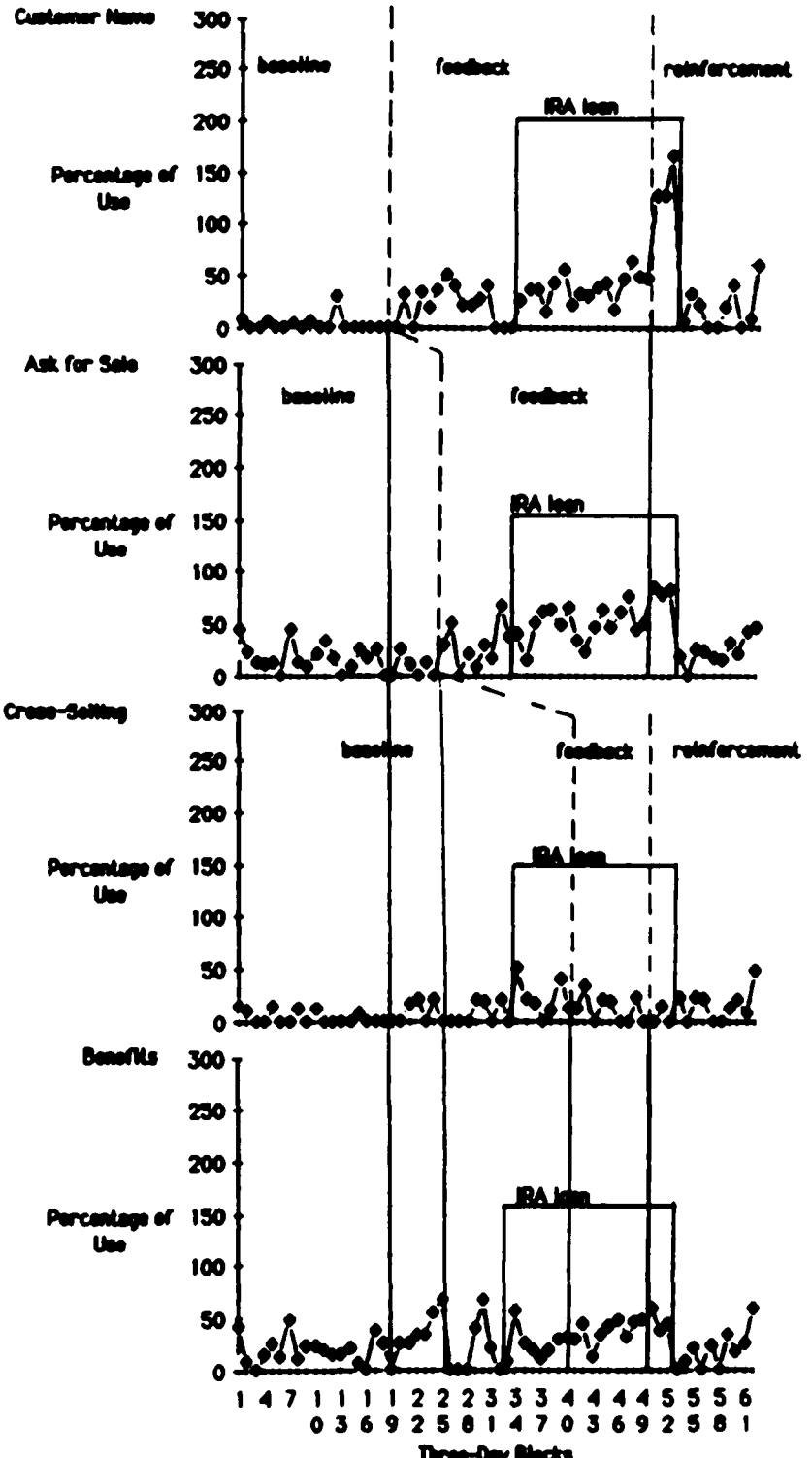


Figure 11. Control conditions for feedback, reinforcement, and IRAloan program for Telemarketer "T."

intervention in "asking for the sale." Graphically, the only skill that shows a possible influence is "benefits," in response to feedback in "using the customer's name." Because the statistical results indicate the possibility of nonindependence, independence cannot be assumed, even though no clear picture emerges through visual inspection.

Table 22

Analysis of the Independence of the Four Sales Skills for Telemarketer "D."

Control Skill	Intercept of skill ratios in first phase	Intercept of skill ratios in second phase	t	df	Significance Level
Ask for sale[a]	1.289	.478	9.856	15	$p < .01$
Ask for sale[b]	.402	1.021	9.743	30	$p < .01$
Cross-sell[a]	.614	.154	6.181	22	$p < .01$
Cross-sell[c]	.127	.802	8.193	7	$p < .01$
Benefits[a]	2.126	.747	9.258	22	$p < .01$
Benefits[c]	1.143	1.177	.130	7	n.s.
Benefits[d]	3.030	2.528	3.704	13	$p < .01$
slopes	- .371	- .322	2.427	13	$p < .01$
Benefits[b]	1.262	2.115	8.727	51	$p < .01$

[a]Control for customer name's feedback intervention

[b]Control for reinforcement phases in customer name and cross-selling

[c]Control for ask for sale's feedback intervention

[d]Control for cross-selling's feedback intervention

Table 23

Analysis of the Independence of the Four Sales Skills for Telemarketer "M."

Control Skill	Intercept of skill ratios in first phase	Intercept of skill ratios in second phase	t	df	Significance Level
Ask for sale[a]	.547	.151	10.213	20	$p < .01$
Ask for sale[b]	.529	.795	5.545	31	$p < .01$
Cross-sell[a]	.395	.289	2.449	20	$p < .05$
Cross-sell[c]	.529	.504	.404	16	n.s.
Benefits[a]	1.024	.954	.560	20	n.s.
Benefits[c]	1.775	2.092	1.555	9	n.s.
Benefits[d]	1.539	1.615	1.033	10	n.s.
Benefits[b]	.915	.966	.897	57	n.s.

[a]Control for customer name's feedback intervention

[b]Control for reinforcement phases in customer name and cross-selling

[c]Control for ask for sale's feedback intervention

[d]Control for cross-selling's feedback intervention

Table 24

Analysis of the Independence of the Four Sales Skills for Telemarketer "T."

Control Skill	Intercept of skill ratios in first phase	Intercept of skill ratios in second phase	t	df	Significance Level
Ask for sale[a]	.205	.137	2.319	19	$p < .05$
Ask for sale[b]	.140	.445	8.403	31	$p < .01$
Cross-sell[a]	.080	.084	.216	26	n.s.
Cross-sell[c]	.231	.001	6.479	8	$p < .01$
Benefits[a]	.139	.322	5.045	27	$p < .01$
Benefits[c]	.073	.048	.305	8	n.s.
Benefits[d]	.137	.272	4.824	13	$p < .01$
Benefits[b]	.110	.213	4.383	56	$p < .01$

[a]Control for customer name's feedback intervention

[b]Control for reinforcement phases in customer name and cross-selling

[c]Control for ask for sale's feedback intervention

[d]Control for cross-selling's feedback intervention

Summary of Time Series Analysis for the Telemarketers

In this experiment the effects of informational feedback and reinforcement-assisted feedback on three telemarketing sales skills were studied. However, a serious confound existed due to the IRA loan program, the effects of which were too powerful to ignore. Therefore, the effects of feedback and reinforcement are qualified by the impact of this

program. Table 25 summarizes the effects of feedback for each telemarketer. Both visual and statistical analysis suggested that feedback increased using the customer's name for all telemarketers. There was also evidence for increases in asking for the sale for two of the telemarketers, but only when the IRA loan data points were included in the analysis, suggesting that if feedback had any effect at all, it was obscured by the powerful effect of the IRA loan program. We must therefore conclude that observed effects were a result of the IRA loan program. Nor were any effects demonstrated on "cross-selling." An effect was statistically supported for Telemarketer "T." on "cross-selling" by a significant increase in intercept (Table 9), but not supported graphically due to high baseline levels of skill use when the IRA loan program began.

Table 25

Summary of Results of Feedback for All Telemarketers: Changes in Skill Use

	Telemarketer "D."	Telemarketer "M."	Telemarketer "T."
Customer Name			
Include IRA points	increase	increase	increase
Exclude IRA points	increase	increase	increase
Ask for Sale			
Include IRA points	no change	increase	increase
Exclude IRA points	no change	no change	no change
Cross-Selling			
Entire baseline	uninterpretable	no change	no change
IRA-only baseline	no change	no change	no change

Reinforcement appeared to cause increases in using the customer's name, but in two of these cases this was qualified by the IRA loan program (see Table 26). When the IRA loan program ended, even though reinforcement was still in effect, skill use dropped dramatically. There was no graphic support for a reinforcement effect on cross-selling for any of the telemarketers.

Table 26

Summary of Results of Reinforcement for All Telemarketers: Changes in Skill Use

	Telemarketer "D."	Telemarketer "M."	Telemarketer "T."
Customer Name	increase	increase[a]	increase[a]
Cross-Selling	no change	no change	no change

[a]Significant drop in skill use after IRA program ended

Sales Percentages

Sales percentages were determined by dividing the number of sales made per three days by the number of calls accepted per three days and multiplying by 100. These data were obtained from the Automatic Call Distributing Management Report Unit and the pencil-and-paper logs described in the Apparatus section. First, sales percentages were examined as a function of the study phases. Next, correlations were computed (a) between sales percentages and the mean length of calls and (b) between sales percentages and sales skill usage.

No interobserver agreement data were obtained for either the number of sales or the number of calls received. The number of sales was determined from the telemarketers' pencil and paper logs. These logs can be assumed to be fairly accurate because the

telemarketing department used them to verify new account openings. It is always possible that a certain percentage of customers who agreed to a sale on the phone never actually sent in a check to open the account, but this is unlikely because the electronic "paperwork" always began immediately while the customer was still on the line. Loans taken over the phone constituted a "sale" regardless of whether that loan was subsequently approved by the Consumer Lending Department of the bank. Monthly loan approval rates were approximately 28%, which increased to 53% during the IRA loan program period.

The number of phone calls received at each workstation was automatically recorded by the Management Reporting Unit of the telephone system. These figures contain an unknown amount of error because personal calls, wrong numbers, in-house calls, and calls transferred across workstations to other telemarketers were all recorded by the equipment. Therefore, the sales percentages can be considered deflated because the divisor (the number of calls) was inflated by error.

The mean length of calls was determined by summing the number of seconds of each call and dividing by the number of calls. This information was obtained from the print-outs of the Management Reporting Unit. This measure also contains error, due to the time spent on personal calls, wrong numbers, and so forth.

Sales Percentages Across Study Phases

Telemarketer "D." Figure 12 shows the sales percentages in time blocks which correspond to the three-day blocks which were used in skill assessment. A decreasing trend at baseline (mean = 15.98) was reversed when feedback for the customer name was introduced. Mean sales percentages increased with the introduction of feedback for each

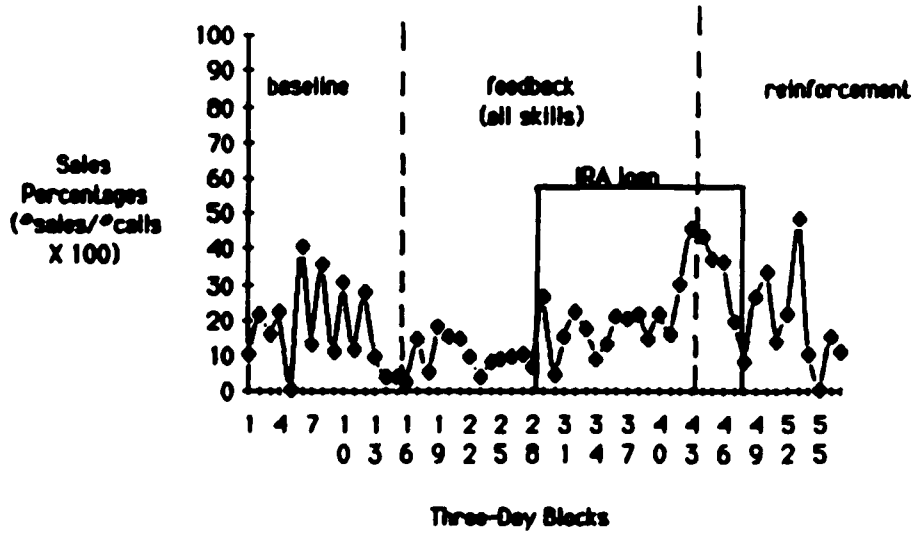


Figure 12. Sales percentages in three-day blocks for Telemarketer "D."

skill (customer name = 17.99%; ask for sale mean = 18.93%; cross-selling mean = 23.31%). Reinforcement resulted in a slightly lower mean (23.02%) The IRA loan program, which overlapped with all feedback and reinforcement phases, can be viewed as a separate study phase. The mean sales percentage before the IRA loan, including baseline, was 13.9%; the mean percentage during the IRA loan program was 22.39%, and the post-IRA loan program mean was 18.71%.

Telemarketer "M." Figure 13 indicates a baseline mean of 8.05%. Mean sales percentages increased with the introduction of feedback for each skill (customer name mean = 16.56%; ask for sale mean = 16.63%; cross-sell mean = 21.52%). Reinforcement resulted in a slightly lower mean percentage (20.11%). Mean sales percentage prior to the IRA loan program, including baseline, was 9.92%. During the IRA loan program, the mean rose to 30.62%, which dropped to 17.57% after the program ended.

Telemarketer "I." Figure 14 indicates that, after baseline (mean = 6.93%), no overall shift in level or trend was exhibited (customer name mean = 10.47%, ask for sale mean = 10.81%; cross-sell mean = 12.86%) until the IRA program was introduced, at which time the level increased and an upward trend began. The mean of sales percentages up to the point of the IRA loan program, including baseline, was 7.21%. The mean of the IRA loan program was 13.33%, which dropped to 9.08% when the IRA loan program ended.

Sales Percentages Correlated with Sales Skills Percentages

Table 27 summarizes the correlation coefficients for each telemarketer between sales percentages in three-day blocks and the use of each sales skill for the length of the study. For Telemarketer "D.," there were significant correlations between sales

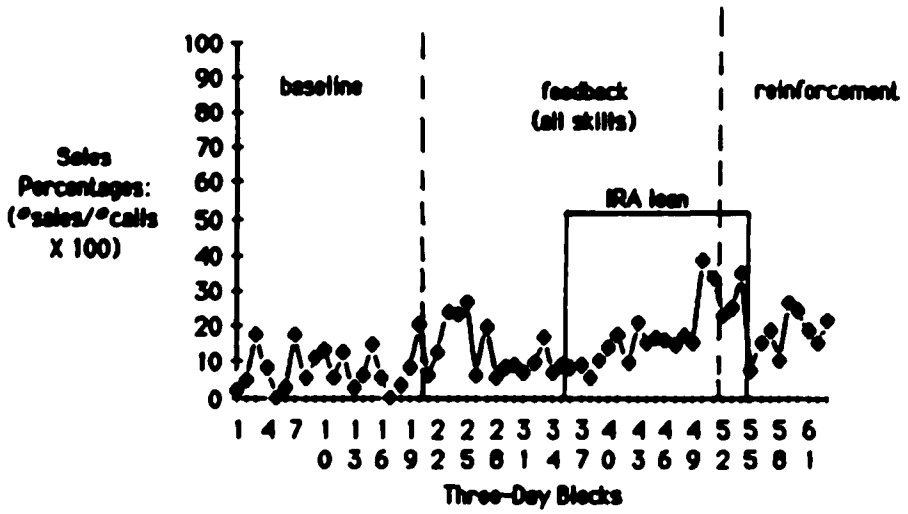


Figure 13. Sales percentages in three-day blocks for Telemarketer "T1."

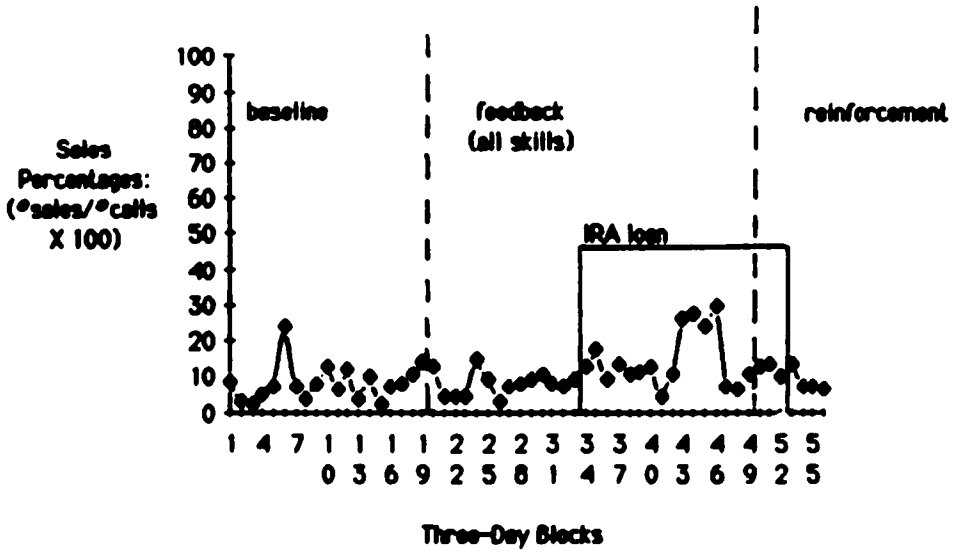


Figure 14. Sales percentages in three-day blocks for Telemarketer "T."

Table 27

Correlation Coefficients of Sales Percentages and Skill Use Percentages

	# of Correlated Pairs	Cust. Name	Ask Sale	Cross-Sell	Benefits
Telemarketer "D."	57	.37*	.25***	.08 n.s.	.31**
Telemarketer "M."	63	.44*	.44*	.38 *	.35*
Telemarketer "T."	62	.69*	.45*	-.11 n.s.	.34*

*p < .01

** p < .05

***p < .10

percentages and the use of skills for using the customer's name, asking for the sale, and stating benefits. For Telemarketer "M.," all sales skills were significantly correlated with sales percentages at the .01 level. For Telemarketer "T.," there were significant correlations at the .01 level for using the customer's name, asking for the sale, and stating benefits.

Sales Percentages Correlated with Mean Length of Calls

For each telemarketer, sales percentages were correlated with the mean length of phone calls in the three-day blocks across the length of the study. The correlation coefficients and the number of correlated pairs for Telemarketer "D." was .65, N = 57; for Telemarketer "M." it was .69, N = 63; and for Telemarketer "T." it was .75, N = 62. All were significant at p < .01.

CONCLUSIONS

The results for informational feedback and token reinforcement obtained in this study support the point of view stated in the introduction that feedback is an environmental stimulus that may acquire several functional properties (discriminative stimulus, reinforcer, punisher) depending on the learning history of the individual and current contingencies for performance. Unlike many previous investigations, this study represents an attempt to separate the informational component of feedback from other components typically combined with feedback such as praise, instructions, peer pressure, or graphic presentations of data. While the current results may reflect a host of uncontrolled variables (discussed below), the positive effects obtained for feedback in the first of three skills point to the potential of such a treatment. The fact that these effects were sometimes short-lived (i.e., an initial increase in behavior which then showed a decreasing trend) supports the notion that feedback without additional contingencies for performance will undergo extinction and eventually cease to control behavior.

Interobserver Agreement

While interobserver agreement ratios were within acceptable ranges for the type of customer inquiry and opportunities to use each skill (which were based on the type of inquiry), failure to consistently reach acceptable levels ($\geq .80$) in five cases must be addressed.

The inability to reliably detect skill use can be attributed to the following seven factors: (1) cassette tapes and tape recorders of poor quality that sometimes broke or did not function properly, resulting in tapes that ran at high or low speeds; (2) background

noise in the department (e.g., other voices, printer) that may have masked the telemarketer's voice; (3) the telemarketer speaking too softly or too quickly; (4) having to infer what the customer was saying because only the telemarketer's voice was recorded; (5) confusion on the scorers' part about skill definitions; (6) scorer fatigue; and (7) low skill use frequency by the telemarketers. That the first three factors played a role is indicated by agreement ratios of .83, .87, and 1.00 for using the customer's name. Reliably detecting this skill is simply a matter of hearing it. These three factors might easily be corrected in future studies with better equipment and steps for noise reduction.

It was necessary for legal and ethical reasons to only record the telemarketer's voice without recording the customer. However, this made it more difficult for scorers to determine what type of inquiry was being made, and resulted in lower agreement ratios for the type of call inquiry and opportunities to use each skill. It was not a serious impediment, however, since these ratios remained within acceptable limits. However, calls on which there was not agreement on the type of inquiry were not scored for agreement on skill use, since by definition opportunities to use the skills would not agree. Omission of those calls resulted in the loss of data.

The last three factors, scorer confusion, scorer fatigue, and low skill use frequency, present greater difficulty. Confusion about skill definitions arose from two sources. First, two skills, asking for the sale and cross-selling, sometimes overlapped. If the telemarketer asked for the sale of a cross-sold product or service, it was not intuitively obvious how this should be scored. Initially, "asking for the sale" was scored if the telemarketer asked the customer to buy the product that represented the initial customer

inquiry, while "cross-selling" was scored if the telemarketer introduced or discussed any additional product. However, these rules did not always make sense. If the telemarketer steered the customer to a more appropriate account (cross-sold) and then asked for the sale of that product while abandoning the initial product, it made sense to score both cross-selling and asking for the sale, particularly for purposes of feedback. Because the skills were not entirely independent, there was a lowered ability to distinguish between them and score them properly.

A second source of confusion was the nature of the sales conversation. While products were typically described by standard features (e.g., rates, terms, penalties), comments and responses to questions were often novel and unique. While many samples of instances and noninstances of skills were provided, each scorer had to determine whether a unique statement constituted a skill, and if so, which skill it was. Scoring instructions requested that the scorer monitor the content of the statements independent of the telemarketer's tone or affect. However, sometimes an inflection of the telemarketer's voice could sway a scorer's judgment; e.g., "And you want to renew the C.D?" (asking for the sale) versus "and you want to renew the C.D." (confirmation of a customer's statement; no skill scored).

Scorer fatigue was a source of error. The poor quality of the tapes, the quantity that had to be scored quickly for feedback, the many long pauses while the customer was speaking, and the often monotonous nature of the conversation (e.g., taking a loan application) all contributed to this fatigue. Voice-activated tape recorders might have helped; however, there was often so much background noise that their use could have been

more confusing. While future research might consider using a random sample of tape-recorded calls, it was not possible in this study. In the beginning of this study, very few sales calls were coming into the department; so few, in fact, that days were aggregated into three-day blocks in order to get a minimum of ten calls per data point. This situation changed, with calls coming in every few minutes during some periods of the study. However, the study procedures were already in motion, and could not be changed.

The last factor, low frequency of skill use, contributed to lower interobserver agreement ratios. Kazdin (1982, p. 56) points out that the extent to which scorers agree is partially a function of the frequency of the participant's performance of the behavior. If the behavior is frequent, agreement ratios will tend to be high due to chance. Conversely, low frequency behaviors contribute to low agreement scores because each instance of the behavior contributes to a large percentage of the agreement ratio. For example, if a behavior occurs ten times and scorers agree on nine of the observations, the agreement ratio is an acceptable 90%, with the disagreement constituting 10% error. However, if a behavior occurs only twice and there is one disagreement, the agreement ratio is only 50%, with 50% error. Thus, low frequencies of skill occurrences such as cross-selling contributed to low agreement ratios.

There are two important implications for low interobserver agreement. The first is that it lowers the statistical conclusion validity (Cook & Campbell, p. 43) of the experiment. If data are unreliable, then one is uncertain of the true state of performance, whether there was an intervention effect, or its magnitude. Inasmuch as this was the case, the experiment lost internal validity. The second implication for this study is that it

lowers the construct validity of feedback. The information contained in the feedback was literally the data that were collected. If those data were unreliable, feedback was false. On the other hand, feedback was known to be false for using the customer's name (the skill with the best reliability coefficients), and yet signficant effects were obtained. This suggests that feedback, if accurate, may be a more powerful intervention than the results of this study suggest.

Discussion of the Use of Interrupted Time Series Analysis

Interrupted time series analysis was used in order to remove serial dependency and analyze the data for significant changes in intercepts and slopes between contiguous phases. There were several problems with the use of this type of analysis. The use of regression analysis which results in intercepts and slopes, rather than mean levels of performance, is problematic. While significant intercept changes suggest an overall change in the mean level of performance between phases, without significant slope changes it is not possible to interpret many results. For example, cases in which the intercept has decreased in the second phase, but the trend line appears to increase, cannot be interpreted without a significant slope change. A second problem, which is critical to the interpretation of single subject designs, is that the use of regression analysis indicates whether the functions in two phases are alike or different, but does not indicate the precise point at which the change occurred. According to the logic of single subject designs, the observed change must occur at the point of intervention in order to be interpreted as a treatment effect and rule out threats to validity. While regression analysis can describe different functions between two phases, it does not indicate at which

point the change in function occurred. While it would be possible to re-run the analyses multiple times in order to find the exact point of change, this procedure would greatly increase the likelihood of a Type I error. Third, interrupted time-series analysis can compare only two contiguous phases at a time, and cannot handle three or more contiguous phases. Re-running the analyses increases the likelihood of a Type I error. Fourth, interrupted time-series analysis perform t-tests on the intercepts and slopes across phases, and cannot account for more than one independent variable. In this study, the IRA loan program acted as an unplanned variable, the effects of which could not readily be factored out.

Summary of Feedback and Reinforcement Effects

Feedback

Feedback appeared to increase skill use for using the customer's name. The strong impact of the IRA loan program and the evidence for dependence of sales skills confounded the feedback procedure, however, such that there were no working control procedures to rule out threats to validity. Thus, one cannot say with certainty whether feedback caused the observed increases in customer name use, or whether another unknown variable was responsible for the observed changes.

Using the customer's name. Feedback reliably increased the use of the customer's name for all telemarketers (see Figure 4). Pilot testing and formal baseline data collection had occurred for 6.5 months prior to the first feedback intervention, during which time the telemarketers tape recorded their telephone calls daily. Feedback may have had its impact on this skill because the constrast between baseline and intervention

was very salient. The telemarketers, who had begun to complain about the lack of consequences for their tape recording, were very glad to finally receive feedback. Thus, the contrast between baseline and treatment may have contributed to the observed increase.

Asking for the sale. While statistical evidence suggests an effect for feedback on "asking for the sale," graphic evidence suggests that the impetus for change in the feedback condition was the IRA loan program. (See Figure 6.)

Cross-selling. Cross-selling skills decreased during baseline for Telemarketers "D." and "M." and showed little change for Telemarketer "T." who was at floor levels of skill use. The IRA loan program began before feedback was introduced and appeared to exhibit its own control over the behavior in two cases (see Figure 7.) There was statistical, but not graphic, evidence for a feedback effect for Telemarketer "D.;" no evidence for feedback for Telemarketer "M.;" and statistical and graphic evidence of a suppressing effect of feedback for Telemarketer "T."

Reinforcement

Reinforcement in the form of a point system for prizes was shown to increase the use of customer name for all telemarketers, but it had no effect for cross-selling. The end of the IRA loan program appeared to have little impact on the use of cross-selling, but it interacted with reinforcement for using the customer's name. When the IRA loan program ended, "using the customer's name" decreased sharply for Telemarketers "M." and "T."

IRA Loan Program Effects

The IRA loan program clearly had a strong impact with both its onset and end in 17 of 24 cases (see Table 28), as indicated by significant intercept changes. The IRA loan program probably had such a strong impact on the use of skills because of the sheer volume of calls which came into the department about it, and that the volume increased as

Table 28

Changes in Intercepts Due to the Onset and End of the IRA Loan Program

IRA loan compared to	Telemarketers		
preceding phases	"D."	"M."	"T."
Customer name	sig. increase	n.s.	n.s.
Ask for sale	n.s.	sig. increase	sig. increase
Cross-selling	sig. decrease	sig. decrease	sig. increase
Benefits	n.s.	n.s.	sig. increase
IRA loan compared to following phases			
Customer name	n.s.	sig. decrease	sig. decrease
Ask for sale	sig. decrease	sig. decrease	sig. decrease
Cross-sell	sig. increase[a]	sig. increase	n.s.
Benefits	sig. decrease	sig. decrease	sig. decrease

[a] The increase in intercept is an artifact of a sharper downward in trend line in the second phase and should not be taken to mean that the skill level increased. See Figure 7.

tax season (April 15) approached. On some days the calls were almost exclusively IRA calls. Through sheer repetition, each telemarketer developed her own style of describing the account. Although no formal scripts were written, each telemarketer began describing the account in her own habitual fashion, usually incorporating several benefit statements and asking for the sale. Once the customer's name was learned, that was incorporated also. The end of the program resulted in decreases in many skills except for cross-selling, which increased in intercept for two of the telemarketers. It can be surmised that once the program ended, the telemarketers were able to turn their full attention back to other types of accounts.

One explanation for the decrease in skill use at the end of the program may be a post-reinforcement pause, if the end of a very successful advertising campaign can be considered the reinforcer. The telemarketers had worked overtime and weekends towards the end of this campaign and were tired. Their voices became flat and unenthusiastic. Interestingly, Telemarkter "T." left the job six weeks after the end of the IRA program for another career not associated with banking, telemarketing, or sales.

Independence of Sales Skills

The sales skills, for the most part, were not independent of one another. Table 29 depicts each skill and whether it was independent of other skills according to significant changes in the time series analysis. Independence is indicated if changes between corresponding phases in the targeted skill were nonsignificant.

Only 8 of 24 tests resulted in independence according to statistical analysis. Feedback for using the customer's name (or some other, unidentified variable) was

related to increases in all other skills except for cross-selling (for Telemarketer "T.") and benefits (for Telemarketer "M."). Feedback for "asking for the sale" and "cross-selling" were related to mixed results (see Table 29), but reinforcement for "using the customer's name" and "cross-selling" (or some other, unidentified variable) was related to decreases on the untargeted skills for all telemarketers.

Table 29

Independence of Sales Skills Summarized for All Telemarketers: Changes in Intercepts

Skills	Telemarketers: "D."	"M."	"T."
Ask for Sale			
Independent of customer name feedback?	increase	increase	increase
Independent of reinforcement?	decrease	decrease	decrease
Cross-selling			
Independent of customer name feedback?	increase	increase	skill ind.
Independent of ask for sale feedback?	increase	skill ind.	increase
Benefits			
Independent of customer name feedback?	increase	skill ind.	increase
Independent of ask for sale feedback?	----------	skill independent	--------
Independent of cross-sell feedback?	increase	skill ind.	increase
Independent of reinforcement?	decrease	skill ind.	decrease

Sales Percentages

Sales Percentages Correlated with Sales Skills Percentages

Sales percentages were significantly and positively related to the use of sales skills in 10 out of 12 cases (see Table 27). Sales percentages and cross-selling were not significantly correlated for two telemarketers.

Due to the nature of correlational analysis, it is unknown whether skill use led to sales, whether getting a sale increased the use of skills, or whether some third factor, such as the type of account inquiry, product knowledge, or the passage of time, led to increases in both. Telemarketer "D." had the highest sales percentages (overall mean = 17.42), followed by Telemarketer "M." (overall mean = 13.49), and then Telemarketer "T." (overall mean = 9.39). Yet when sales percentages were correlated with sales skill ratios, Telemarketer "D." had the weakest correlation coefficients in each skill category. Telemarketer "T.," with the lowest sales percentages, showed the strongest relationship between sales and skill usage. This result indicates that skills might have been used by Telemarketer "T." only after she was certain that a customer would buy the product. In other words, a sale led to skill use. For all telemarketers, while it is unknown what proportion of sales can been attributed to skill use, clearly other unmeasured factors (for example, product knowledge) played an important role in closing a sale.

Sales Percentages Correlated with Mean Length of Calls

Sales percentages were significantly and postively related to the mean length of calls. It is unknown whether "keeping the customer on the phone" resulted in sales, or whether a sale resulted in a longer phone conversation. The second alternative is the more likely

explanation, mostly because the process of taking loan applications over the phone contributed greatly to the length of calls and always represented a sale. However, there may have been cases in which the telemarketer, by describing product features and benefits at length, was able to interest the customer in the product and sell it. It should also be noted that the process of selling products that represent sizable investments for customers often took place over several phone calls while customers investigated other options, discussed it with spouses, and so forth. As the number of phone calls to discuss one account increases, the mean length of phone calls will decrease, even though one sale is still represented by several calls.

Summary and Conclusions

Informational feedback appeared to be a viable intervention for increasing the use of sales skills in a telemarketing department. Significant, positive effects were obtained with feedback in the first of three skills in spite of many uncontrollable factors which may have competed with the intervention for behavioral control. Reinforcement was successful in 1 out of 2 skills, but in those cases the effects were quite temporary. Factors that may have competed with the treatments are described below.

One aspect of the interventions is how the telemarketers may have viewed them. The face validity of feedback to the telemarketers may have been weak. While the experimenter's view of their behavior (on which feedback was based) came from tape-recorded phone calls, their view of their own behavior may have been quite different. This experiment omitted outbound calls, calls on mortgages, home equity lines of credit, student loans, and account servicing calls. When the telemarketers thought

about their own behavior and skill use, they probably included these calls in their own overall estimation, even though they all knew upon what the feedback was based. Skill use ratios would not appear accurate if all of these additional calls were considered by the telemarketers. Therefore the feedback may not have seemed accurate to the telemarketers, and this may have weakened any reinforcing effect the feedback might have. In addition, feedback was delivered, by necessity, two to five days (counting weekends) after the behaviors had been performed. Studies in applied settings have indicated that feedback is more efective when delivered as soon as possible after the response (Balcazar et al., 1985-86). If it had been possible to deliver feedback daily or without a delay before the beginning of the next phase, there might have been a much more powerful effect. Reinforcement was also based on the skill ratios, and therefore the value of the reinforcer may have also been reduced for these same reasons (low face validity and delay of delivery).

Besides the perceived inaccuracy of feedback, there were cases in which feedback actually was inaccurate. Because the customer could not be heard, it was assumed that the customer always gave his or her name (a deliberately false assumtion), so that the telemarketer always had the opportunity to use the name. This method always led to an underestimate of the skill use ratio, because more opportunities were assumed to have occurred than actually did occur. Therefore feedback for this skill was always inaccurate. Interestingly, this skill was the one that had the most clear and consistent positive effect. Thus even somewhat inaccurate feedback can exert control over behavior.

Some skills were clearly more difficult to perform than others. For example, "using

the customer's name" is simply a matter of injecting it into a phrase or question, whereas "cross-selling" means that the telemarketer must determine an appropriate place in the conversation to introduce an additional account and make the transition gracefully. This was not difficult for some accounts, for example, checking, which is naturally linked with a variety of other services, e.g., overdraft protection. But on other accounts, such as rate requests on certificates of deposits, it was difficult enough to determine whether the customer was considering opening a new account, was considering closing an account and taking it elsewhere, or had one elsewhere and was considering bringing it to this bank. Attempting to cross-sell in this situation was difficult at best. The differences that exist between skill difficulty levels may account for the failure of feedback, especially if more in-depth skill training was required in order for the telemarketers to perform the skill proficiently.

The lack of back-up reinforcers for feedback meant that telemarketers were actually in an extinction condition with feedback. It was suggested in the introduction that feedback without additional consequences would eventually cease to control behavior. This appeared to frequently be the case when sales skills increased in level with the introduction of feedback, only to decrease across time until either the IRA loan program or reinforcement was introduced (see Figures 6 and 7.)

According to time series analysis, there were effects for the reinforcement intervention in three cases. Part of these effects may have been due to the IRA loan program. In either case, it should be pointed out that the telemarketers stated that they were not happy with the reinforcers, which were points for prizes from a gift catalog (a

program that the bank already had in place for tellers and employees in other departments). The telemarketers wanted cash, which was against bank policy. Therefore it was somewhat surprising that there were any reinforcement effects at all.

A last issue in this experiment strikes at the very nature of sales behavior and salesperson-customer interactions. While it is assumed that the use of sales skills prompts customers to buy, this is not always the case. Sometimes the most eloquent use of sales skills goes unreinforced (no sale), and at other times customers buy products without any prompting at all on the part of salespeople. This observation was also made by Brown et al. (1980) when they trained store salespeople to improve their customer service skills. They noted that not only did customers frequently ignore sales skills, sometimes the use of these skills were punished (e.g., customers angrily saying they did not want to be bothered). Thus, coexisting with feedback and reinforcement in this study was a competing, noncontingent reinforcer for the telemarketers: customers opening accounts. The lack of a contingency between sales and sales skills may have competed with feedback and reinforcement for the use of those skills. This may have contributed to the decline in skill use after initial elevations when feedback was implemented.

In Balcazar et al's. (1985-86) review of the feedback literature, four outcomes of feedback effects were identified: (a) consistent effects, in which significant effects were found in all subjects, settings, or behaviors targeted for change; (b) mixed effects, in which significant changes were identified in some, but not all, of the targeted subjects, settings, or behaviors; (c) no effects, in which feedback resulted in no differences from baseline or control condtions; and (d) unknown effects, which consisted of studies that had

baselines of only one or two data points and therefore could not be evaluated. The authors reviewed 47 studies in which feedback without additional treatments was tested. While there was not consistency across studies for feedback characteristics identified by Fairbank and Prue (1982) (feedback source, privacy, participants, content, mechanism, and frequency), results of the meta-analysis indicated that 13 (28%) of the 47 studies had consistent effects; 27 (57%) had mixed effects; and 7 (15%) had no effect. These results fall clearly in line with the current investigation which resulted in mixed effects.

Balcazar et al. have concluded that "feedback is by no means uniformly effective" (p. 75). After reviewing how neutral stimuli can become discriminative stimuli or conditioned reinforcers through the establishment of a relationship between the stimulus and primary consequences, the authors stated that "feedback will function to prompt or reinforce improved performance if and only if it is related to some more primary consequence" (p. 76). The current study supports this position from the demonstration that while informational feedback without any additional contingencies increased the use of sales skills, the difference between baseline and feedback in many cases was not marked, and skill improvements were not necessarily maintained over time. The lack of a more robust effect was probably due to the lack of more "important" consequences, for example, supervisor evaluations, bonuses, or even additional sales training. Future research may address the feedback-reinforcement relationship by providing response-contingent feedback both with and without reinforcement in an alternating treatments design.

From the point of view of the applied researcher attempting to discover robust treatment effects that are effective in a variety of settings or for a variety of behaviors, it

will most certainly be necessary to combine informational feedback with more potent treatments. While feedback in the current study may have shown more powerful effects if all uncontrolled variables could have been controlled, in most applied settings it is not possible to do this. Thus, in future applications, appropriate steps should be taken to ensure that the improvements gained when feedback is implemented will be maintained. The addition of public posting, data graphed in comparison to a baseline, little delay between performance and feedback, and desired back-up reinforcers on an appropriate schedule of reinforcement will exert stronger control over behavior.

Appendix A

Transcripts of Calls (Telemarketer's Voice Only)

Benefit statement are in **bold print.**

Cross-sell attempts are in *italics.*

Asking for the sale is <u>underlined.</u>

Account: Rates inquiry for certificates of deposit

"Good afternoon, [name of bank and service], [M.] speaking, may I help you?"

"O.K., which CD's were you interested in?"

"O.K. The 18-month account, today's rate, is 6.40. It **compounds daily** and yields a 6.61."

"O.K, we have . . we have a 4-year, and we also have an account that goes six-to-ten years. And the rates on both those accounts is 7%, with a yield of 7.25. Those are our **highest rates** and longest terms available."

"A four year or a six-to-ten year."

"Right. Right."

"On the 4-year the minimum is 500."

"Yeah, these are **insured by the, oh, the FSLIC,** the Federal Savings and Loan Insurance, up to $100,000 each."

"Um-um."

"Would this be for a new account?"

"O.K., <u>what we can also do, if at this time we reserve an account number for you, and if you would like to open the account through mail</u>, we can **guarantee that interest rate**

for seven days."

"O.K. You're welcome. Bye-bye."

Account: IRA

"Good afternoon, [name of bank and service,] [T] speaking, may I help you?"

"Mm...mm."

"Ok. Ok. Well, we do have different terms of accounts for that. Uh, six months, ok, and there's different interest rates, ok, and you have different terms. Like a six-month account, right now, the interest rate is 6.2, and that's with a yield of 6.3. You understand what I mean by "yield."

"Right. You don't start that until, when you withdraw it's when you claim it as income."

"When you start withdrawing."

"Mm, because --"

"Well, we have a 4-year term, and a term for 6 to 10 years, and both of them are at the same rate now, which is 7%, with a yield of 7.25."

"Oh, we could do it over the phone. What I could do is just mail out all the necessary documents for you to sign, and you'll sign them and just mail them back to me along with your check, and then I deposit - pardon me?"

"Ok. 'Cause you're going to deposit how much now?"

"Ok. Or what we could do is, where are you getting the money from?"

"They could contact - they have this form they can fill out and mail it here to [bank name.] Ok, which is called a trustee-to-trustee transfer, and they can mail it directly here and then our bank can directly deposit that to the account that you want to open."

"If you'd like to do it that way."

"Mm-mm. You can call back, too, you know, there's no problem. You can call back, if the interest rates are -- you never know, they might be going up. You can call and check the rates."

"Ok, you're welcome. Bye-bye."

Account: automobile loan

"[Name of bank and service,] [D.] speaking, may I help you?"

"Hello?"

"Ok, let me see."

"Is that any better? . . . ok."

"Is it a new car or a used car?"

"Ok. Uh, let me see. We've got a couple of different rates available. Um, are you going, what kind of a down payment are you going with?"

"Ok, so you're looking at maybe 20% down or greater?"

"Ok. Um, we'll go up to 36 months, and the variable rate is 9.92; the fixed rate is at eleven-and-a-half right now. Ok? What year is this car?"

"Ok. Did you want to go ahead and apply?"

"Do you want to go . . . the rates?"

"Ok. That's, that's what you're looking at. As far as our rates go, ok? With a variable, it has a cap on it, ok? The highest it'll go . . ."

"Now let me add it up here."

"The highest it'll go is 14.42. Ok. And you would not see that adjustment until the very last payment. Ok? Um, and **you know of course it could go down.**"

"Ok, and on which rate?"

"The variable or the fixed?"

"Ok. Can you hang on?"

"Thanks, just a second."

"Thanks for holding on. Uh, the payment - now this is calculated without insurance - is figured at $257 and 83 cents. Ok?"

"Yeah, that's as high as they'll go on a used car loan."

"Uh-huh?"

"Well, um, it's a good idea all the way around. Because, you know, you've got that deductability there. Um, the interest on the home equities tend to be a little less as well. Like ours is fixed at seven-and-a-half right now, for the calender year 87. And, um, you know . . ."

"That's a good way to look at it. It's a *revolving line of credit.* Ok? And you only use- you're only paying on the amount of money you use." [continue to cross-sell home equity line of credit.]

Account: IRA loan

"[Name of bank and service] [D.] speaking, can I help you?"

"Mm-mm. . . ok . . .ok. The loan is to be repaid over a maximum of four years, ok? Um, and, now do you know how the program actually works?"

"All right. Um, we have advertised in the paper that you can borrow the money to fund your IRA at 4% interest. Ok? Now, on that, you can, with that, paying 4% on the loan, ok? Um, you would be earning nothing on your IRA until the loan is paid back. Ok? Um, just to give you an idea what a payment would be . . ."

"On $4000 over four years, at 4% interest, the monthly payment would be $90.31. Ok?

Now that's one you can go about it, or, you can go with paying 4% above the account yield. And what I mean by that is, right now, our 6-to-10 year account is paying 7% and the yield on the account is 7.25. We would bill you at 4% above the account yield, or 11.25. But your IRA would begin earning interest right away. "

"Ok, your monthly payment would be $103.86."

"OK? You can go with whatever way would fit your budget the best."

"No, really, it's about $15 difference."

"Mm-mm. Right. Your account would be drawing interest while you're paying back the loan. Where going the other route, you wouldn't be earning any interest. If you can afford the extra $13 in your budget, generally you're going to make out better going that way."

"It would be eleven-and-a-quarter? Yeah."

"Ok. In the course of four years, that would be, the interest is, **you're looking at earnings of approximately $1293.53."**

"So you'd be earning a lot of interest during that time."

"After that, it'll still stay at 7% with 7.25 yield. Ok? It'll stay that way for the minimum term of six years."

"Yup. Yup. You, at that point, after the loan is repaid and if the six-year term is up, you can do with it what you'd like to do with it."

"Well, no, no, the way it works is, it has to be on deposit with [bank name] a minimum of six years. Ok, so, you know, after the six years is up, that includes your repayment time. Ok? You can do whatever you'd like with it."

"Ok. Um, for four years... on $4000, this is going to be just a tad low. You would have, over the course of four years, the current rate is 6.80, ok, um, **you would have $1242. 82 in earned interest, giving you a combined total of $5242.82."**

"That's because the interest rate's just a little bit higher."

"Right."

"Uh-huh, yes, yes. You end up, see, **with the tax savings**, if you talk to your tax man about it, **by paying the higher interest rate, too, next year, see, you're not losing all your deductions for consumer loans completely. You can deduct up to 65%. So, you know, taking that into consideration, as well as, you know, the tax savings immediately, and earning interest on the account, you're going to make out ahead going, paying, 4% above the account yield.**"

"Well, what I need from you is your name, address, social security number, birth dates, place of employment, address, how long you've been there, gross wage figure, and a phone number, and a complete list of all your debts, including your mortgage payment, your visa, mastercard - -"

"Yup, yup, exactly. Exactly. Did you want to apply, or did you want to think this over?"

"Ok. Ok. Ok."

"We are located downtown at [street], but you can call me up . . . the fastest way is for you to call me back with the information. Ok. And that way, what we do is, I can take the information over the phone, send it right down to our consumer lending department, and they will go ahead and review, and call me back with an answer."

"About a day."

"Um, actually we're located downtown on [street], and you've got our phone number, right?"

"My name is [D.]"

"Yup, uh-huh, I'll take the application. It takes about 10 minutes, you know, provided you have all the information and, um, what's your name?"

"Ok, [customer name]. I'll be here until five and I'll be here in the morning at nine."

"Ok. Yeah. Ok?"

"Ok, very good."

"Um, I won't need the numbers; the things that I need are just who your credit card is with, and um, the minimum monthly payment, and just an estimation of what your outstanding balance is. Ok?"

"All right. Thanks a lot. Bye-bye."

Account: checking

"This is [M.] May I help you?"

"Ok. Let's put is this way. We have two kinds of checking accounts. Ok, we have a basic checking account which can be opened with a minimum of $100. Ok., in order to earn interest and avoid services charges, you should maintain a $1000 average monthly balance."

"Ok. Should your balance fall below the $1000 you would be service charged $5.00."

"No, there is not. You pay for your checks, it's, um, it's, at this time it's eight-and-a quarter for your first 200 checks."

"Ok, and that would, um, with the $1000 average balance, **that would earn you also five-and-a-quarter percent interest."**

"Ok, the other checking account we have, to avoid service charges, you should maintain a $300 minimum average daily balance; ok, to earn interest, you would have to maintain an average balance of $10,000."

"Ok, if your balance falls below $300, it's a $2.00 fee per cycle, ok, along with a 15-cent per-check charge."

"Yes, yes it is. Well, statement cycle. It's not going to run definitely from the first to the thirty-first, but it'll be a 30-day cycle depending on your statement cut-off."

"Ok. Now the service charges can be waived on, let's see, the $300 account, ok, if you maintain the $300 minimum balance or you maintain $2500 in *a related savings account.* "

"Ok. *Or if you have a direct deposit feature.* "

"Ok. *We have a [product name] statement savings and a passbook savings.* And they both have a five-and-a-half percent rate."

"With the passbook savings you obviously receive a passbook, and the passbook is necessary for all withdrawals. With the statement savings account, you receive a bank card. You can do transactions at **Metroteller, you can do transfers from checking into savings, you have more options available** as far as the statement account goes, than you would have on the passbook account."

"Ok, do you think maybe it would help if I sent some brochures that showed some of - *we have certificate accounts, we have money management accounts;* I could send some brochures to you so you'd have it to look over."

"Oh, they'll be there tomorrow."

"Ok. You're welcome. Bye-bye."

Appendix B

Definitions of Bank Accounts

Listed below are several bank accounts which were considered suitable for the use of sales skills for purposes of this study. Other accounts, such as mortgage inquiries or the Home Equity Line of Credit, were not considered suitable because they were too complex or beyond the expertise of the telemarketer. Definitions of these accounts were obtained from Rachlin (1984).

Certificate of Deposit: Official receipts issued by a bank that state that a specified sum has been deposited for a specified period to time and at a specified rate of interest... The terms of certificates of deposit vary widely and it is important to remember that they can be negotiable or nonnegotiable.

Check: A check, in its most basic form, is nothing more than written instructions to a bank to transfer funds from an account to a specific, named party. All checks have the command "Pay to the order of . . . " written on them, as well as the name of the bank and the name of the writer of the check.

Direct Deposit: A service which authorizes an employer to deposit employees' paychecks directly into accounts specified by the employee.

Individual Retirement Account (IRA): An individual retirement account is a tax-deferred investment plan that allows any wage earner (and spouse) to save a portion of income for retirement or to legally shelter income from taxation. In other words, an investor can accumulate funds for retirement by making tax-deductible payments to the plan. The earnings (interest) of the IRA are not taxed until they are distributed.

IRA Loan: A consumer loan in which a customer borrows a specified sum of money at a low interest rate in order to open a long-term IRA account with the lending institution.

Installment Loans (includes personal, auto, home improvement): Given a good credit history, consumers may borrow specified sums of money and repay them at either fixed or variable interest rates. Installment loans are of two main types: open-end and close-end. Open-end credit is also called "revolving credit." This means that there is a line of credit that a consumer can use at his or her discretion up to a prescribed limit. Repayments are also made at the borrower's discretion, as long as an agreed-upon minimum payment is made when due. The most common examples of open-end or revolving credit are bank credit cards and retailers' credit cards. Closed-end loans are for a specific amount for a definite period of time, with a specified number of payments to be made. Most large extensions of credit, such as for an automobile purchase, are closed-end loans.

Loans can also be categorized as secured or unsecured. A secured loan is backed by collateral - that is, an item of value that serves as security for the loan and is claimed by the creditor if the borrower does not repay the loan. An unsecured loan, on the other hand, does not have collateral to back up the loan in case of default. Unsecured loans generally have higher interest rates.

Savings: A regular "passbook account" is an "in-and-out" account that allows a saver to add to or withdraw from the account any amount at any time without a penalty. "Statement accounts" provide a periodic statement giving a concise picture of all account transactions during the period.

Visa Debit Card: A debit card works in conjunction with a checking account and may be used instead of a check. The card is accepted like any other credit card; however, funds are deducted directly from the consumer's checking account. A line of credit on the checking account is required to cover those instances in which there is not a large enough balance to pay for the purchase, and an interest rate is charged if the line of credit is used.

Appendix C

TELEMARKETER SALES SCORE FORM

Telemarketer name _____ **Date of call** _____

CALL NUMBER

	1		2		3		4		
	OPP.	OCC.	OPP.	OCC.	OPP.	OCC.	OPP.	OCC.	
USED CUSTOMER'S NAME	Y N		Y N		Y N		Y N		
ATTEMPTED CROSS-SELL (WHAT?)	Y N		Y N		Y N		Y N		
DESCRIBED ACCOUNT/ SERVICE BENEFITS	Y N		Y N		Y N		Y N		
ASKED FOR SALE	Y N		Y N		Y N		Y N		
WHAT TYPE OF ACCOUNT?									
Not a saleable call because: referal to another t.m. or dep't									
follow-up call on current account or problem (what?)									
returned message on maturing CD with script									
Home Equity Line of Credit, student loan, or mortgage rates inquiries									

Appendix D

FEEDBACK FORM

Telemarketer D./ T./ M. For Days:_____

Using the customer's name

Goal is at least once per opportunity

You had the following number of opportunities
to use the customer's name in a sales call .....____

You used the skill the following number of times.....____

____ You achieved ____% of your goal OR

____ You exceeded your goal by ____%.

Asking for the sale

Goal is at least once per opportunity

You had the following number of opportunities
to ask for the sale in a sales call .....____

You used the skill the following number of times .....____

____ You achieved ____% of your goal OR

____ You exceeded your goal by ____%.

Cross-selling

Goal is at least once per opportunity

You had the following number of opportunities
to cross-sell in a sales call .....____

You used the skill the following number of times .....____

____ You achieved ____% of your goal OR

____ You exceeded your goal by ____%.

Appendix E

INCENTIVE GOALS

Telemarketer: D. / T. / M.

NEW GOALS ARE FOR WEEK OF: _____

USING THE CUSTOMER'S NAME

The old goal for the week of _____ was _____ and you achieved _____

You { earned / did not earn } 25 points for meeting your goal on this skill.

The new goal is _____. Good Luck!

CROSS-SELLING

The old goal for the week of _____ was _____ and you achieved _____

You { earned / did not earn } 25 points for meeting your goal on this skill.

The new goal is _____. Good Luck!

_____ YOU EXCEEDED YOUR GOAL OR MAINTAINED THE CEILING LEVEL ON ONE OR BOTH SKILLS, SO YOU EARN AN ADDITIONAL 25 POINTS. CONGRATULATIONS!

TOTAL POINTS THIS WEEK: _____

References

American Banker, Inc. (1988). American Banker 1988 Yearbook. New York, NY: Author.

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formation. Journal of General Psychology, 54, 279-299.

Anderson, D. C., Crowell, C. R., Sucec, J., Gilligen, K. D., & Wikoff, M. (1982). Behavioral management of client contacts in real estate brokerage: Getting agents to sell more. Journal of Organizational Behavior Management, 4, 67-96.

Annett, J. (1969). Feedback and Human Behavior. Baltimore: Penguin Books.

Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. Organizational Behavior and Human Decision Processes, 39, 133-144.

Balcazar, F., Hopkins, B., & Suarez, Y. (1985-86). A critical, objective review of performance feedback. Journal of Organizational Behavior Management, 7(3-4), 65-89.

Becker, L. J. (1978). Joint effect of feedback and goal-setting on performance: A field study of residential energy conservation. Journal of Applied Psychology, 63, 428-433.

Bennett, D., & Simmons, R. W. (1984). Effects of precision of knowledge of results on acquisition and retention of a simple motor skill. Perceptual and Motor Skills, 58, 785-786.

Bilodeau, I. M. (1966). Information feedback. In E. A. Bilodeau (Ed.), Acquisition of skill (pp. 255-285). New York: Academic Press.

Brethower, D. M. (1972). Behavior analysis in business and industry. Kalamazoo: Behaviordelia.

Brewer, A. & Gipson, M. (1989). Formative feedback in the training of psychology clinic interviewing skills. Unpublished manuscript, University of the Pacific, Stockton, CA.

Broden, M., Hall, R. V., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eighth-grade students. Journal of Applied Behavior Analysis, 4, 191-199.

Brown, K. M., Willis, B. S., & Reid, D. H., (1981). Differential effects of supervisor verbal feedback and feedback plus approval on institutional staff performance. Journal of Organizational Behavior Management, 3, 57-68.

Brown, M. G., Malott, R. W., Dillon, M. J., & Keeps, E. J. (1980). Improving customer service in a large department store through training and feedback. Journal of Organizational Behavior Management, 2, 251-265.

Calpin, J. P., Edelstein, B., & Redmon, W. K. (1988). Performance feedback and goal setting to improve mental health center staff productivity. Journal of Organizational Behavior Management, 9 (2), 35-58.

Catano, V. M. (1976). Improvement in workers' performance through feedback of information of system performance. Perceptual and Motor Skills, 42, 487-490.

Chhokar, J. S. & Wallin, J. A. (1984). A field study of the effect of feedback frequency on performance. Journal of Applied Psychology, 69, 524-530.

Connellan, T. K. (1978). How to improve human performance: Behaviorism in business and industry. NY: Harper and Row.

Cook, T. D. & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Dick, H. W. (1978). Increasing the productivity of the day relief textile machine operator. Journal of Organizational Behavior Management, 2, 45-57.

Duncan, P. K., & Bruwelheide, L. R. (1985-86). Feedback: Use and possible behavioral functions. Journal of Organizational Behavior Management, 7, 91-114.

Eldridge, L., Lemasters, S., & Szypot, B. (1978). A performance feedback intervention to reduce waste: Performance data and participants' responses. Journal of Organizational Behavior Management, 1, 258-268.

Fairbank, J. A. & Prue, D. M. (1982). Developing performance feedback systems. In L. W. Frederiksen (Ed.), Handbook of organizational behavior management (pp. 281-299). NY: John Wiley and Sons.

Fellner, D. J. & Sulzer-Azaroff, B. (1984a). Increasing industrial safety practices and conditions through posted feedback. Journal of Safety and Research, 15, 7-21.

Fellner, D. J. & Sulzer-Azaroff, B. (1984b). A behavioral analysis of goal-setting. Journal of Organizational Behavior Management, 6, 23-34.

Fixsen, D. L., Phillips, E. L., & Wolf, M. M. (1972). Achievement place: The reliability of self-reporting and peer-reporting and their effects of behavior. Journal of Applied Behavior Analysis, 5, 19-30.

Ford, J. E. (1980). A classification system for feedback procedures. Journal of Organizational Behavior Management, 2, 183-191.

Ford, J. E. (1984). A comparison of three feedback procedures for improving teaching skills. Journal of Organizational Behavior Management, 6, 65-98.

Frederiksen, L. W., Richter, W. T., Johnson, R. P., & Solomon, L. J. (1981). Specificity of performance feedback in a professional service delivery setting. Journal of Organizational Behavior Management, 3, 41-55.

Frost, J. M., Hopkins, B. L., & Conard, R. J. (1981). An analysis of the effects of feedback and reinforcement on machine-paced production. Journal of Organizational Behavior Management, 3(2), 5-17.

Geller, E. S., Eason, S. L., Phillips, J. A., & Pierson, M. D. (1980). Interventions to improve sanitation during food preparation. Journal of Organizational Behavior Management, 2(3), 229-240.

Gottman, J. M. (1981). Time series analysis: A comprehensive introduction for social scientists. Cambridge: Cambridge University Press.

Harris, V. W., Bushell, D., Sherman, J. A., & Kane, J. F. (1975). Instructions, feedback, praise, bonus payments, and teacher behavior. Journal of Applied Behavior Analysis, 8, 462.

Hayes, S. C. & Cone, J. D. (1981). Reduction of residential consumption of electricity through simple monthly feedback. Journal of Applied Behavior Analysis, 14, 81-88.

Herold, D. M., & Greller, M. M. (1977). Feedback: The definition of a construct. Academy of Management Journal, 20, 142-147.

Huber, V. L. (1985-86). The interplay of goal and promises of pay-for-performance on individual and group performance: An operant interpretation. Journal of Organizational Behavior Management, 7, 45-64.

Ilgen, D. R., Fisher, C. D. & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 64, 349-391.

Ilgen, D. R., Hobson, C. J., & Dugoni, B. L. (1981). Performance feedback in organizations: The development of a measure. Catalog of selected documents in psychology, 11, 62, MS: 2319.

Jens, K., & Shores, R. (1969). Behavioral graphs as reinforcers for work behavior of mentally retarded adolescents. Education and Training of the Mentally Retarded, 4, 21-26.

Johnson, R. P., & Frederiksen, L. W. (1983). Process versus outcome feedback and goal setting in a human service organization. Journal of Organizational Behavior Management, 5 (3/4), 37-52.

Karan, B. S., & Kopelman, R. E. (1986). The effects of objective feedback on vehicular and industrial accidents: A field experiment using outcome feedback. Journal of Organizational Behavior Management, 8, 45-56.

Kazdin, A. E. (1974). Reactive self-monitoring: The effects of response desirability, goal-setting, and feedback. Journal of Consulting and Clinical Psychology, 42, 704-716.

Kazdin, A. E. (1982). Single case research designs. NY: Oxford.

Komaki, J. (1981). A behavioral view of paradigm debates: Let the data speak. Journal of Applied Psychology, 66, 111-112.

Komaki, J., Collins, R. L., & Penn, P. (1982). The role of performance antecedents and consequences in work motivation. Journal of Applied Psychology, 67, 334-340.

Komaki, J., Heinzmann, A. T., & Lawson, L. (1980). Effect of training and feedback: Component analysis of a behavioral safety program. Journal of Applied Psychology, 65, 261-270.

Kopelman, R. E. (1982). Improving productivity through objective feedback: A review of the evidence. National Productivity Review, 2, 43-55.

Kreitner, R., Reif, W. E., & Morris, M. (1977). Measuring the impact of feedback on the performance of mental health technicians. Journal of Organizational Behavior Management, 1, 105-109.

Krumhus, K. M., & Malott, R. W. (1980). The effects of modeling and immediate and delayed feedback in staff training. Journal of Organizational Behavior Management, 2, 279-294.

Larson, J. R. (1984). The performance feedback process: A preliminary model. Organizational Behavior and Human Performance, 33, 42-76.

Latham, G. P., & Baldes, J. J. (1975). The "practical significance" of Locke's theory of goal-setting. Journal of Applied Psychology, 60, 122-124.

Latham, G. P., & Yukl, G. A. (1976). Effects of assigned and participative goal-setting on performance and jobsatisfaction. Journal of Applied Psychology, 61, 166-171.

Leitenberg, H., Agras, W. S., Thompson, L. E., & Wright, D. E. (1968). Feedback in behavior modification: An experimental analysis in two phobic cases. Journal of Applied Behavior Analysis, 1, 131-137.

Locke, E. A. (1967). Motivational effects of knowledge of results: knowledge or goal-setting? Journal of Applied Psychology, 51, 324-329.

Locke, E. A. (1980). Latham vs. Komaki: A tale of two paradigms. Journal of Applied Psychology, 65, 16-23.

Locke, E. A., Cartledge, N., & Koeppel, J. (1968). Motivational effects of knowledge of results: A goal-setting phenomenon. Psychological Bulletin, 70, 474-485.

Locke, E. A., Feren, D. B., McCaleb, V. M., Shaw, K. N., & Denny, A. T. (1980). The relative effectiveness of four methods of motivating employee performance. In K. D. Duncan, M. M. Gruenberg and P. Wallis (Eds.), Change in work life. NY: John Wiley & Sons, Ltd.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. Psychological Bulletin, 90, 125-152.

Luyben, P. (1984). Drop and tilt: A comparison of two procedures to increase the use of venetian blinds to conserve energy. Journal of Community Psychology, 12, 149-154.

Maher, C. (1981-82). Performance feedback to improve the planning and evaluation of instructional programs. Journal of Organizational Behavior Management, 3, 33-40.

Mahoney, M. J. (1974). Self-reward and self-monitoring techniques for weight control. Behavior Therapy, 5, 48-57.

McCarthy, M. (1978). Decreasing the incidence of "high bobbins" in a textile spinning department through a group feedback procedure. Journal of Organizational Behavior Management, 1, 150-159.

Nadler, D. A. (1979). The effects of feedback on task group behavior: A review of the experimental literature. Organizational Behavior and Human Performance, 23, 309-338.

Newby, T. J., & Robinson, P. W. (1983). Effects of grouped and individual feedback and reinforcement on retail employee performance. Journal of Organizational Behavior Management, 5, 51-68.

O'Brien, R. M. (1988). "Goals versus feedback; Explaining employee performance improvement effects." Paper presented at the Association for Behavior Analysis, Florida Association for Behavior Analysis, Winter meeting; St. Petersburg, Fla.

Panyon, M., Boozer, H., & Morris, N. (1970). Feedback to attendents as a reinforcer for applying operant techniques. Journal of Applied Behavior Analysis, 3, 1-4.

Peterson, G. B., Wheeler, R. L., & Trapold, M. A. (1980). Enhancement of pigeons' conditional descriminaition performance by expectancies of reinforcement and nonreinforcement. Animal Learning and Behavior, 8, 22-30.

Peterson, N. (1982). Feedback is not a new principle of behavior. Behavior Analyst, 5, 101-102.

Powers, W. (1973). Feedback: Beyond behaviorism. Science, 179, 351-356.

Prue, D. M., & Fairbank, J. A. (1981). Performance feedback in organizational behavior management: A review. Journal of Organizational Behavior Management, 3, 1-16.

Prue, D. M., Krapfl, J. E., Noah, J. C., Cannon, S., & Maley, R. F. (1980). Managing the treatment activities of state hospital staff. Journal of Organizational Management, 2, 165-181.

Quaglieri, P., & Carnazza, J. P. (1985). Critical inferences and the multidimensionality of feedback. Canadian Journal of Behavioral Science, 17, 285-293.

Rachlin, H. (1984). The money encyclopedia. NY: Harper & Row.

Ralis, M. T., & O'Brien, R. M. (1986). Prompts, goal-setting, and feedback to increase suggestive selling. Journal of Organizational Behavior Management, 8, 5-18.

Rapp, S. R. Carstensen, L. L., & Prue, D. M. (1983). Organizational behavior management 1978-1982: An annotated bibliography. Journal of Organizational Behavior Management, 5, 5-50.

Reber, R. A., & Wallin, J. A. (1984). The effects of training, goal setting, and knowledge of results on safe behavior: A component analysis. Academy of Management Journal, 27, 544-560.

Rothstein, R. N. (1980). Television feedback used to modify gasoline consumption. Behavior Therapy, 11, 683-688.

Rowe, B. J. (1981). Use of feedback and reinforcement to increase the telephone reporting of independent automobile appraisers. Journal of Organizational Behavior Management, 2, 35-41.

Runnion, A., Johnson, T., & McWhorter, J. (1978). The effects of feedback and reinforcement on truck turnaround time in materials transportation. Journal of Organizational Behavior Management, 1, 110-117

Runnion, A., Watson, J. O., & McWhorter, J. (1978). Energy savings in interstate transportation through feedback and reinforcement. Journal of Organizational Behavior Management, 1, 180-192.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. Psychological Bulletin, 95, 355-386.

Seligman, C., & Darley, J. M. (1977). Feedback as a means of decreasing residential energy consumption. Journal of Applied Psychology, 62, 363-368.

Stoerzinger, A., Johnston, J. M., Pisor, K., & Monroe, C. (1978). Implementation and evaluation of a feedback system for employees in a salvage operation. Journal of Organizational Behavior Management, 1, 268-282.

Stuart, R. B. (1971). A three-dimensional program for the treatment of obesity. Behaviour Research and Therapy, 9, 177-186.

Sulzer-Azaroff, B. (1978). Behavioral ecology and accident prevention. Journal of Organizational Management, 2, 11-44.

Sulzer-Azaroff, B., & de Santamaria, M. C., (1980). Industrial safety hazard reduction through performance feedback. Journal of Applied Behavior Analysis, 13, 287-295.

Tosti, D. & Jackson, S. (1981, May). Formative and summative feedback. Paper presented at the annual meeting of the Association for Behavior Analysis, Milwaukee, WI.

Van Houten, R., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effect of timing and feedback, public posting and praise upon academic performance and peer interaction. Journal of Applied Behavior Analysis, 8, 449-457.

Van Houten, R., & Nau, P. (1981). A comparison of the effects of posted feedback and increase police surveillance on highway speeding. Journal of Applied Behavior Analysis, 14, 261-271.

Van Houten, R., & Nau, P. (1983). Feedback interventions and driving speed: A paradigmatic and comparative analysis. Journal of Applied Behavior Analysis, 16, 253-281.

Weiner, N. (1954). The human use of human beings. Boston: Houghton Mifflin.

Welsch, W. V., Ludwig, C., Radiker, J. E., & Krapfl, J. E. (1973). Effects of feedback on daily completion of behavior modification projects. Mental Retardation, 11 (4), 24-27.

Williams, E. A., & Gottman, J. M. (1982). A user's guide to the Gottman-Williams time-series analysis computer programs for social scientists. Cambridge: Cambridge University Press.

Winett, R. A., Kagel, J., Battalio, R. C., & Winkler, R. C. (1978). The effects of rebates, feedback, and information on electricity conservation. Journal of Applied Psychology, 63, 73-80.

Winett, R. A., Neale, M. S., & Grier, H. C. (1979). The effects of self-monitoring and feedback on residential electricity consumption: Winter. Journal of Applied Behavior Analysis, 12, 173-184.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. Journal of Applied Behavior Analysis, 11, 203-214.

Zohar, D., Cohen, A., & Azar, N. (1980). Promoting increased use of ear protectors in noise through information feedback. Human Factors, 22, 69-79.