Mace, F. Charles

THEORIES OF REACTIVITY IN SELF-MONITORING: A COMPARISON OF OPERANT AND COGNITIVE-BEHAVIORAL MODELS

The University of Arizona PH.D. 1983

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __✓__.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages _____
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Other Dissertation contains pages with print at a slant, filmed as received.

THEORIES OF REACTIVITY IN SELF-MONITORING:
A COMPARISON OF OPERANT AND COGNITIVE-BEHAVIORAL MODELS

by

F. Charles Mace

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 3

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by F. Charles Mace

entitled Theories of Reactivity in Self-Monitoring: A Comparison of Operant and Cognitive-Behavioral Models

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Thomas R Kratochwill	3/25/83 Date
John R. Bergan	3-25-83 Date
Anne J. Miller	3-25-83 Date
Tommy C. Cox	3/25/83 Date
Allen D. Nebelkopf	3/25/83 Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Thomas R Kratochwill	3/25/83
Dissertation Director	Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his/her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: 

ACKNOWLEDGMENTS

I want to express my deep gratitude to my mother, Laura, who provided continued support, encouragement and love throughout my education. Great appreciation is also extended to my brother and sister-in-law, Dick and Colleen, for providing much needed direction and support during critical stages of my development. I also want to thank Dr. Leon and Clara McGill and Eleanor Means whose unwavering encouragement and support will always be appreciated. My sincere appreciation to my committee members, Drs. John Bergan, Rosemary Rosser, Glen Nicholson and Anne Miller who provided constructive suggestions and criticisms. Special thanks to Dr. Thomas Kratochwill who directed this project and whose knowledge, patience and friendship will be a source of strength throughout my career. Thanks also to Dr. Kirby Brown who offered friendship, advice and assistance whenever needed. Finally, I will always be indebted to my wife, Sarah, and children, Nicole and Hilary, whose love, patience and steadfast support carried us though. Indeed, this work is a tribute to their efforts.

TABLE OF CONTENTS

PAGE

LIST OF TABLES vi

LIST OF ILLUSTRATIONS vii

ABSTRACT viii

1. INTRODUCTION 1

2. LITERATURE REVIEW 14

- Definition of Self-Monitoring 15
- The Relationship of SM to Self-Control 16
- The Role of SM in Assessment and Treatment 19
- Self-Recording Procedures 24
  - General Considerations 24
  - Direct Methods 26
  - Indirect Methods 40
  - Self-Recording Devices 45
  - Summary 49
- Training in Self-Monitoring 50
- The Relationship of Self-Monitoring to Other Assessment Procedures 54
- Methodological Issues 56
  - Reactivity of Measures 57
  - Experimental Controls for Reactivity 58
  - Evidence of Reactivity 60
  - Variables Affecting the Reactivity of Self-Monitoring 64
  - Validity of Self-Monitoring 71
  - Variables Affecting the Validity of Self-Monitored Data 73
- Summary and Implications 78

3. METHOD 80
- Participants 80
- Experimental Design 80
- Dependent Variable 81
- Procedures 83

4. RESULTS 88

5. DISCUSSION 99

APPENDIX A: Introductory Instructions for the Preself-Monitoring Phase and Question Series A 110

APPENDIX B: Introductory Instructions for the Self-Monitoring Phase and Question Series B 112

REFERENCES 114

LIST OF TABLES

TABLE		PAGE
1.	Repeated Measures ANOVA	89
2.	Group Pretest, Posttest and Adjusted Posttest Means and Standard Deviations	92
3.	Analysis of Covariance (Prestest Covariate)	93
4.	Differences Among Adjusted Posttest Means	96
5.	Statistical Tests and Significance Levels for Between Group Comparisons	97

LIST OF ILLUSTRATIONS

Figure | Page

1. Mean rate per minute of verbal nonfluencies by experimental group on the pretest and the adjusted posttest (pretest and covariate) 90

ABSTRACT

Three theoretical models explaining reactivity in self-monitoring were examined including one cognitive-behavioral and two operant views. Each theoretical account was represented by the following self-monitoring conditions: (1) self-monitoring (Rachlin-operant recording response model), (2) self-monitoring and goal setting (Kanfer-cognitive-mediational model), (3) self-monitoring, goal setting and self-reinforcement (Nelson and Hayes-multiple cueing stimuli model), (4) goal setting and self-reinforcement (Kanfer-cognitive-mediational model), and (5) training only.

The comparative effects of the five self-monitoring conditions on the dependent measure, verbal nonfluencies, were evaluated using a repeated measures analysis of covariance design with the pretest as covariate.

Results of the study indicated that self-monitoring conditions containing a self-reinforcement component (i.e., conditions 3 and 4) produced the greatest reactivity. Moreover, the presence of reinforcement appeared to positively influence whether subjects reached their individually set goals for reducing nonfluencies. The vast majority of self-reported cognitions associated with the occurrence of the target behavior were independently judged to be neutral rather than self-reinforcing or self-punishing.

The implications of this study for the role of external versus covert forms of reinforcement were discussed as well as the use of this technique in clinical practice.

CHAPTER 1

INTRODUCTION

Applied behavioral psychology has undergone a period of rapid growth over the past decade and a half. Only recently, however, has the development of assessment techniques kept pace with advances in treatment strategies. Direct observation, the one-time mainstay of behavioral assessment, is now complemented by a variety of assessment procedures targeting a broad range of human functioning. These procedures include psychophysiological assessment, self-report inventories, checklists and rating scales, behavioral interviews and criterion-referenced assessment.

A recent and important addition to this collection of assessment strategies is self-monitoring. The technique of self-monitoring is a two stage process (Nelson, 1977b). First, an individual observes his or her own behavior and discriminates the occurrence of a target response in the context of the current behavioral repertoire. Second, a self-recording response is made in which the individual records some aspect of the target behavior (e.g., frequency, duration or intensity). Accurate self-monitoring data are dependent on the reliable occurrence of both response discrimination and self-recording (Nelson, 1977a).

As an assessment procedure, self-monitoring serves a number of unique functions. It combines the objective characteristics of direct observation with the flexibility of self-report. Many behaviors, by nature or convention, are unavailable for direct observation by external observers (e.g., cognitive behaviors, marital quarrels, and sexual behavior). In such cases, self-monitoring can provide a more valid measure of client functioning than other available assessment procedures (Haynes and Wilson, 1979). From a clinical standpoint, having individuals record their own behavior is more convenient and cost-efficient than employing outside observers. In addition, self-monitoring may serve an educational function by improving clients' awareness of the effects their behavior has on others and, in turn, the effects of the environment on their own behavior (Haynes, 1978).

Another attractive facet of self-monitoring is its tendency, under certain conditions, to produce alterations in the target response. Such changes in behavior are referred to as the reactivity of self-monitoring. Following a number of early studies demonstrating reactive effects (e.g., Leitenberg et al., 1968; Rutner and Bugle, 1969, Thomas, 1971), investigators began employing self-monitoring as a therapeutic variable. Among the child behaviors successfully altered by self-monitoring are: Study rates (Broden, Hall and Mitts, 1971); out-of-seat behavior (Bornstein, Hamilton and Quevillon, 1977); nervous tics (Ollendick, 1981); and cross-gender behavior (Rekers, Amaro-Plotkin and Low, 1977). Similar modifications have been reported with the following adult behaviors: drug urges (Hay, Hay and Angle, 1977);

obsessive ruminations (Emmelkamp and Kwee, 1977); nail biting (Vargas and Adesso, 1976); and depressive mood and inactivity (Harmon, Nelson and Hayes, 1980).

In contrast to these generally supportive findings, other researchers have found marginal or nonexistent reactivity due to self-monitoring. For example, Fixen, Phillips and Wolf (1972) observed only a small, transient increase in room cleaning behaviors and Romanczyk (1974) found no differences in weight loss between subjects who self-recorded their daily weight and a no-monitoring control group. Such inconsistencies across studies have prompted examination of the variables controlling reactivity. The growing literature in this area, reviewed by McFall (1977) and Nelson (1977a, 1977b), suggests that reactivity may be a function of a specified set of conditions present during self-monitoring.

These recent advances in uncovering some of the parameters of self-monitoring underscore the need for continued systematic research. Future attempts would be greatly enhanced under the direction of a theoretical model explaining the processes underlying reactive effects (McFall, 1977). Such a framework is needed not only to integrate available evidence but to guide prospective research efforts. To date, three theories of reactivity appear in the self-monitoring literature. These include the cognitive-mediational model proposed by Kanfer (1970a, 1970b, 1971, 1975, 1977), the operant recording response model by Rachlin (1974), and the multiple cueing stimuli model presented by Nelson and Hayes (1981).

The explanatory mechanism of Kanfer's theory is nested within a multistage model of self-regulation. According to this view, effective control over one's own behavior is a product of self-monitoring, self-evaluation, and self-reinforcement. The components of the self-regulatory process occur as a sequential chain of events, each triggered by the occurrence of the preceding component. The reactive chain begins when the individual observes and records some aspect of his or her own behavior (self-monitoring). Inevitably, the person reacts to the products of self-monitoring by comparing their behavior against some performance standard or goal (self-evaluation). Performance equal to or exceeding the criterion leads to self-delivery of reinforcers. These may take the form of covert positive statements or actual administration of external rewards. Alternatively, failure to meet expectations for performance may result in covert or overt self-punishment. These self-administered consequences acquire "motivational" properties which may in turn influence subsequent performance of the self-monitored response. Thus, the distinguishing characteristics of the Kanfer model are: 1) the role self-monitoring plays in initiating the reactive process; 2) self-recorded performance prompts self-evaluation against some personal performance standard; and 3) self-monitored behavior changes as a function of self-administered consequences (Nelson and Hayes, 1981).

Yet, all components of the cognitive-mediational model are not considered equal in their contribution to behavior change. Proponents of this model emphasize the importance of goal setting and/or

self-reinforcement and minimize the potential of self-monitoring alone to produce reactive effects. For example, Spates and Kanfer (1977, p. 10, 13) state, "self-monitoring may contribute to behavior change only in conjunction with other procedures...", "criterion setting (goal setting) represents a critical stage in the model and...training which did not include this component was ineffective." Other advocates of this model similarly de-emphasize the role of self-monitoring in favor of self-reinforcement. Bandura (1976, p. 151, 152) asserts,

> "As a rule, simply observing and recording one's own behavior has no consistent behavioral effects. When self-monitoring does produce change, it is likely under circumstances that activate covert goal setting and self-evaluative consequences... Those who self-monitor and receive feedback on goal attainments often do not perform any better than do baseline control groups."

At the present time, empirical evidence pertaining to Kanfer's theory of reactivity is mixed. Much of the literature has examined the individual components of the self-regulation model. As noted earlier, considerable support exists for the reactive effects of self-monitoring alone (Nelson, 1977a; Haynes, 1978). However, when self-evaluation and self-reinforcement are investigated separately, results are less clear. For instance, Santogrossi, et al. (1973) trained disruptive adolescents in a psychiatric school setting to self-evaluate their adherence to the rules of classroom conduct. In this case, self-evaluation was insufficient to alter client levels of disruptive behavior. Similar findings were reported by Turkewitz, O'Leary and Ironsmith (1975) when self-evaluation was applied to a range of academic and social problems. Yet as Roberts and Dick (in press) point

out, the effects of externally instructed self-evaluation cannot be assessed independent of self-monitoring since self-recorded data may function as a discriminative stimulus for self-evaluation. In contrast, the effects of self-administered consequences are more easily isolated. In a study representative of work in the area, Humphrey, Karoly and Kirschenbaum (1978) had second-grade students either self-impose response cost or self-administer rewards on the basis of their classroom behavior. Under both conditions, behavior change occurred in the direction predicted by the experimental paradigm.

Attempts to concurrently evaluate multiple components of Kanfer's self-regulation model have likewise generated inconclusive findings. Spates and Kanfer (1977) compared the arithmetic performance of first-graders under the following conditions: 1) self-monitoring, 2) criterion setting, 3) self-monitoring plus criterion setting, 4) self-evaluation plus self-reinforcement, and 5) a control group. Subtle distinctions were made between the experimental instructions provided to each group. For example, subjects in the self-monitoring group were told to verbalize aloud their addition steps; no self-recording occurred. The criterion-setting group was instructed to verbalize aloud the correct sequence of steps (e.g., "First I should add the two numbers on the right, then I should add the two numbers in the middle..."). The third group combined the self-monitoring and criterion-setting conditions. Finally, children in the self-evaluation plus self-reinforcement group verbalized aloud the correct sequence of steps and also stated out loud whether their answer was correct or not

following each problem. No significant differences were found between the control and self-monitoring groups, although both groups differed from the remaining experimental conditions. The authors concluded that the critical component responsible for change was criterion setting. The addition of self-monitoring to criterion setting produced no significant gains in performance.

Although the Spates and Kanfer (1977) findings appear impressive, interpretation of these findings is not altogether clear. First, the absence of self-recording from the self-monitoring condition omits an important component of self-monitoring. Failure to find differences between the control and self-monitoring groups could be understood as self-observation alone being insufficient to effect change (although the null hypothesis can never be proven). Second, it is difficult to conceive of the criterion-setting condition as establishing criteria for performance. The criterion was not to increase or decrease performance but to change the quality of their performance consistent with a rule for solving addition problems. Since the control and self-monitoring conditions were not given the benefit of a problem solving strategy, it is not surprising that their accuracy lagged behind the other groups. And third, a study by Sagotsky, Patterson and Lepper (1978) compared self-monitoring, goal setting, and both self-monitoring and goal setting, without the use of contingent reinforcement. In this case, experimental conditions were consistent with common definitions used in the literature. Using arithmetic performance and on-task behavior as the dependent variables,

the goal setting group was asked to "look ahead" and decide how many problems thay could complete if they worked hard throughout the period. Self-monitoring subjects periodically recorded whether or not they were at work on their math units. The results revealed that self-monitoring was the most effective treatment for increasing both on-task and academic performance. Goal setting did not affect either student measure nor did it contribute to the effectiveness of self-monitoring. Considered collectively, these discrepant findings point to the need for further research before the tenets of Kanfer's theory can be empirically validated.

The second of the three theories of reactivity is presented by Rachlin (1974). He contends that the search for the determinants of "self-effected behavior change" should focus on environmental rather than hypothetical events. According to Rachlin, cognitive/motivational theorists unnecessarily infer that covert states or events operate in the context of a reinforcement paradigm to control behavior. He states, "The behavior itself is all the evidence we need that self-control is going on. It is...behavior in relation to the contingencies imposed that comprise self-control" (Rachlin, 1974, p. 97). Within this framework, either the self-recording response or self-administered consequences serve as cues to remind the person of the ultimate environmental contingencies operating to control behavior (Nelson and Hayes, 1981). "Counting and timing of events with mechanical or written aids and the techniques of self-reinforcement (and self-punishment) are...ways to increase the salience of the relationship

between behavior and its consequences" (Rachlin, 1974, p. 105). Thus, Rachlin emphasizes the importance of the self-recording response in producing reactive effects. In contrast to the Kanfer model, however, it is the external rather than internal consequences associated with the self-monitored behavior that inevitably determine response strength.

As was the case with the Kanfer model, unequivocal evidence for the efficacy of Rachlin's theory is also hard to find. On one hand, proponents of the cognitive-mediational model led by Kanfer (1977) and Bandura (1976) cite a literature demonstrating the importance of self-evaluation and self-reinforcement. Bandura and Perloff (1967) found self-monitored performance on a motor task to be comparable to an external reinforcement condition in increasing productivity. However, when self-reinforcement and external reinforcement were compared holding self-monitoring constant, the self-reinforcement condition produced significantly superior performance on a motor task (Switzky and Haywood, 1974). Similarly, self-reward was found to effect greater weight loss than self-monitoring of daily weight alone (Mahoney, 1974). Alternatively, these findings are countered by a large number of studies in which self-monitoring alone did effect behavior change (see Haynes and Wilson, 1979; McFall, 1977; and Nelson, 1977a). In addition, Rachlin (1974) reports a number of animal experiments in which rats and pigeons demonstrated "self-control", presumably without the aid of self-evaluative and reinforcing cognitions.

The third theoretical account of the reactive effects of self-monitoring, the multiple cueing stimuli model, was recently

advanced by Nelson and Hayes (1981). Their theory shares with the others the assumption that the antecedent self-recording response initiates the reactive chain. However, the Rachlin view that reactivity can be triggered by the self-monitoring response alone is extended to implicate the entire self-recording process. The stimuli that cue the environmental consequences responsible for behavior change include: therapist instructions, training in self-monitoring, the self-monitoring device, feedback from others, and the self-recording response itself. Also acknowledged within this view is the potential contribution of self-evaluation and self-consequation. Rather than ascribing causal properties to these private events, Nelson and Hayes contend that they too serve to cue the individual to the environmental contingencies controlling behavior.

Although direct examination of the multiple cueing stimuli model remains to be done, Nelson and Hayes (1981) cite a number of studies whose findings are consistent with their theory. Several experiments have demonstrated that accurate self-monitoring is not necessary in order for reactivity to occur (Broden et al., 1971; Fixen et al., 1972; Hayes and Cavior, 1977, 1980; Lipinski and Nelson, 1974). Hayes and Cavior (1977) made this point by correlating subject accuracy scores with their ratio change scores. The correlations were .01 for face touching, -.02 for nonfluencies, and .26 for value judgments, indicating that the magnitude of reactivity was unrelated to accurate self-monitoring. Nelson and Hayes argue that the above data are evidence that reactivity is not dependent on the self-recording

response. That is, accurate self-observation without self-recording may be sufficient to produce reactive effects. Further support for the contention comes from studies in which reactivity was maintained even after the frequency of the target behavior dropped to zero (e.g., Lipinski, et al., 1975; Maletzky, 1974). This suggests that environmental variables other than the occurrence of the target behavior function to cue behavior change.

One such variable which appears to contribute to reactivity is the obtrusiveness of the self-recording device. Maletzky (1974) had five psychiatric patients self-record maladaptive motor responses (e.g., repetitive scratching and facial tics) with a wrist-worn counter. In each case, removal of the wrist counter resulted in a resurgence of the maladaptive behavior, necessitating a fading procedure to maintain the reactive effects. Similarly, Broden et al. (1971) found that removing paper slips used for self-recording study behavior was associated with a drop in study rates. However, when Nelson, Lipinski and Boykin (1978) compared the impact of belt-worn versus hand-held counters, no significant differences were found in the degree of reactivity obtained. In this instance, the authors appropriately emphasized the need for further research to clarify these findings.

In summary, three theories explaining the reactivity of self-monitoring have been reported in the behavioral literature. The cognitive-mediational model proposed by Kanfer (1970a) considers reactivity in the context of a three stage model of self-regulation.

Self-monitoring initiates the reactive chain by generating performance data which are compared to some personal performance standard. This comparison of self-evaluation leads to the self-administration of either covert or overt consequences contingent on the individual's rate of behavior in relation to the performance criteria. Behavior change is presumed to occur as a result of these self-administered consequences. However, the role of self-monitoring is considered subordinate to goal setting and self-reinforcement. In contrast, Rachlin's (1974) operant recording response model considers the self-recording response to be the catalyst for reactivity. The act of self-recording (or self-delivery of consequences) serves as a cue to remind the person of the environmental contingencies which ultimately control behavior. Finally, the multiple cueing stimuli model advanced by Nelson and Hayes (1981) extends the range of environmental stimuli which may prompt a change in behavior. The entire self-monitoring process is hypothesized to affect reactivity including therapist instructions, training in self-monitoring, the self-recording device, feedback from others, the self-recording act, goal setting and self-reinforcement.

In reviewing the literature relevant to these theories, there is a striking lack of consistent evidence to support any single model. Research efforts thus far have limited their investigation to a single theoretical model, to components within a model, and to retrospective surveys of past research to identify findings consistent with a particular theoretical view. Since some of the discrepancies cited

earlier may be due to divergent subjects, settings and experimental conditions employed across studies, cross-study comparisons are ambiguous. The present study attempts to extend this literature by providing a comparison of the three theories of reactivity under uniform conditions. Four experimental conditions and a training only group were contrasted: 1) self-monitoring (SM); 2) self-monitoring and goal setting (SM + GS); 3) self-monitoring, goal setting and self-reinforcement (SM + GS + SR); 4) goal setting and self-reinforcement (GS + SR); and 5) TRAINING ONLY group. The following conditions were used to represent the various theoretical views: SM (Rachlin); SM + GS (Kanfer): SM + GS + SR (Nelson and Hayes): and GS + SR (Kanfer). It was predicted that if goal setting is the central component in the reactive process as stated by Kanfer, then reactive effects due to SM + GS, SM + GS + SR and GS + SR would exceed SM. If, as Bandura contends, self-reinforcement is a critical factor in reactivity, then SM + GS + SR and GS + SR were expected to be greater than SM and SM + GS. Finally, if the three experimental components operate additively to produce reactivity, then SM + GS + SR was expected to be greater than SM, SM + GS and GS + SR.

CHAPTER 2

LITERATURE REVIEW

Most behavioral assessment procedures involve the use of persons other than the target individual to measure student/client behavior. Self-monitoring (SM) differs from these procedures in some important respects. A major difference is the active role the student/client plays in assessing his or her own behavior. Data generated from SM represents self-observation, self-measurement, or self-ratings. From this perspective, SM is consistent with a trend in behavior therapy away from external control of behavior toward greater participation of the client in managing their own therapeutic programs (Bandura, 1969; Ciminero, Nelson and Lipinski, 1977; Mahoney, 1977; Thoresen and Mahoney, 1974). Another distinctive feature of SM is the scope of behavior available for assessment. Numerous child behaviors are inaccessible to external observers, particularly school personnel. Cognitions, toileting, sexual activity, delinquent behavior, and academic strategies are private events by nature or convention. Moreover, self-observers are exposed to the entire population of target responses rather than limited samples under limited conditions (Kazdin, 1974). Therefore, the potential exists for self-observers to obtain more complete data and data inaccessible to external observations (Nelson, 1977a).

The applications of self-monitoring to school psychology are as diverse as the subspecialty itself. Its utility has been demonstrated with disruptive behavior (Bornstein et al., 1977) academic performance (Hundert and Batstone, 1978), study behavior (Broden et al., 1971), cheating (Flowers, 1972), nail biting (McNamara, 1971), teacher behavior (Knapczyk and Livingston, 1973), tics (Thomas, Abrams and Johnson, 1971), and severe behavior problems (Santogrossi et al., 1973), to sample a few. The value of SM as an assessment tool merits considerable attention as do the potential problems associated with its use. This review will examine issues related to definitions, the relationship of SM to self-control, the role of SM in assessment and treatment, SM procedures and methods, methodological concerns, and the relation of SM to other assessment procedures.

Definition of Self-Monitoring

The term self-monitoring refers to a multistage process. First, the child must become aware of or discriminate the presence or absence of a target response. The discrimination response itself is controlled by the occurrence of covert or overt cues (Thoresen and Mahoney, 1974). As the saliency of the cues increases, so does the likelihood of making the appropriate discrimination. Second, the occurrence of the target response is systematically recorded. In contrast to discriminating, self-recording requires overt action on the part of the child to make a permanent record of the event. Both these stages are essential to SM and are considered by some authors to be sufficient to describe the process (Nelson, 1977a; Nelson and Hayes,

1981; Rachlin, 1974, Simkins, 1971). Other authors, however, consider SM to involve a third stage, namely self-evaluation. During this stage, the individual evaluates his or her behavior against some standard (Kanfer, 1977; Mahoney and Arnkoff, 1978; Thoresen and Mahoney, 1974). These standards may represent criteria established by teachers, parents or an agency or may be self-established goals. Usually, meeting the performance criteria is linked with some form of reinforcement, while behavior falling short of the criteria is sometimes associated with self-administered aversive events. Although controversy continues over the stages involved in SM, this review will refer to self-monitoring as the two-stage process of response discrimination and self-recording.

The Relationship of SM to Self-Control

Much of the self-monitoring literature treats SM as a component of self-control procedures (also referred to as self-management and self-regulation). In addition to functioning as a method of assessment, SM is viewed as a key feature of a well integrated therapeutic strategy. Current behavioral perspectives of self-control include the cognitive-behavioral and operant models (Karoly, 1977).

The cognitive-behavioral perspective considers self-control responses to be a product of self-management of environmental contingencies as well as covert contingencies which are assumed to mediate overt behavior (Kanfer, 1977; Mahoney and Arnkoff, 1978; Meichenbaum, 1977). Kanfer's (1970a, 1975, 1977) three stage model of self-regulation is representative of this mediational approach,

although numerous variations have developed (see Mahoney and Arnkoff, 1978). The first stage of the model is self-monitoring. When performance criteria are not specified, SM involves response discrimination and self-recording. Once performance criteria are established, either internally or externally, conditions are set for the second stage, self-evaluation. During this stage, the individual compares his or her own performance against an a priori set of performance criteria. Persons trained to make this comparison can be expected to direct their behavior toward this standard. Meeting or exceeding the pre-established criterion triggers self-reinforcement. This final stage involves self-delivery of reinforcers contingent on successful performance. Reinforcers may include those available in the external environment, self-generated thoughts or statements, or some combinaton of the two.

The operant model of self-control (Skinner, 1953; Rachlin, 1974) interfaces with the cognitive-behavioral perspective along several dimensions. First, both consider consequences to ultimately control response frequency (Nelson and Hays, 1981). This implies that the principles of learning operate to determine the self-control response. Second, the individual participates in the management and manipulation of stimulus conditions which influence behavior. That is, the student/client can arrange the environment to maximize the chances that desired responses will occur. For example, a student may position a picture, symbol or statement on his/her desk which reminds him/her to raise a hand before speaking out. A third dimension common to both

models is the self-management of environmental contingencies to increase or decrease target responses. This refers to dispensing reinforcers or punishers to alter behavior in the direction of the performance criteria. Despite these shared characteristics, a fundamental distinction between these two models exists. According to the cognitive-behavioral view, self-evaluation involves cognitive processes which mediate the influence of the external environment on behavior (Kanfer and Karoly, 1972). The assumption is that covert self-statements or attitudes (Kanfer, 1977) can be functionally related to the overt self-control response. In contrast, the explanatory power of the operant model rests on the control that environmental contingencies have on behavior. Little recognition is given to mediational processes assumed to contribute to overt behavior.

The role SM plays in self-control strategies will be influenced by the theoretical model adopted. This influence will permeate both research and application considerations. From the standpoint of research, issues concerning the accuracy, reliability, and reactivity of SM are affected by theory. Investigators following the congnitive-behavioral model have explained these issues in terms of cognitive variables interacting with environmental events (Kanfer, 1970a, 1975, 1977; Thoresen and Mahoney, 1974). Operant researchers, on the other hand, prefer to restrict their investigation to identifying functional relationships between observable phenomena (Nelson, 1977a; Nelson and Hayes, 1981). School psychologists in educational or clinical settings who avail themselves of SM procedures will be similarly influenced.

Guided by the mediational model, training youngsters in the use of SM can focus on developing "attitudes" consistent with behavioral goals, self-evaluative statements comparing performance with performance criteria and the use of covert rewarding or punishing statements contingent on performance. From the operant perspective, training efforts would emphasize response discrimination and development of the self-recording response. Efforts at behavior change would rely heavily on the self-management of environmental contingencies.

The above discussion underscores the close link between self-monitoring and self-control. Many of the uses of SM with children will be in the context of self-management programs. A large literature exists in this area to guide the practitioner in the devlopment and application of self-control strategies in a number of settings (see Karoly, 1977; McLaughlin, 1976; Mahoney and Arnkoff, 1978; Meichenbaum, 1977; Ross, 1981). However, in addition to its application in self-regulation, SM has evolved to become an important vehicle for assessment independent of the therapeutic package to which it is applied. Moreover, numerous researchers have demonstrated that SM can be a significant therapeutic variable exclusive of other interventions (see Nelson, 1977b). The following sections explicate the function of SM in assessment and treatment.

The Role of SM in Assessment and Treatment

The practitioner selecting self-monitoring as an assessment procedure will find it useful for a number of purposes. First, SM is a convenient way to obtain global information regarding child, teacher,

or parent functioning at the preintervention stage. Data can take the form of narrations, self-ratings, responses to checklists, or audio recordings of events as they relate to the target problem. Global measures of this nature are useful in identifying potential target behaviors and generating preliminary hypotheses regarding the functional relationship between problematic behavior and antecedent and consequent events (Ciminero et al., 1977). As relationships begin to emerge, global measures can be narrowed to focus on the target behavior and a description of the antecedent and consequent conditions surrounding its occurrence. These data can then function as baseline measures in addition to providing valuable information toward the development of intervention strategies. Mahoney (1977) points out that beginning with global measures, particularly in the form of unstructured narrations, can also be helpful in managing resistance to "structured behavioral programs". Frequently, school psychologists encounter such resistance when dealing with teachers and parents. When it becomes desirable to alter teacher or parent behavior as a means of achieving results with children, Mahoney suggests gradual introduction of structured assessment procedures to maintain vital cooperation.

At this stage of intervention, target behaviors have been defined, baseline measures established, and a therapeutic plan has been developed. The second assessment function of SM can be to obtain measures useful in evaluating the effectiveness of treatment (McFall, 1977). Toward this end, SM measures should be taken repeatedly over the course of treatment and preferably in the context of a time-series design (Kratochwill, 1978).

Third, self-monitoring may be the procedure of choice when covert behaviors are the target of assessment (Haynes and Wilson, 1979). Behaviors such as headaches, muscle tension, nightmares, depressive thoughts, obsessive thoughts, pain, attentiveness, emotions, covert reinforcers and covert punishers are not readily amenable to external observation. In these cases, SM may provide more valid estimates of behavior than indirect methods such as interviews and questionnaires, although empirical support for this assumption remains to be generated (McFall, 1977).

A fourth potential use of SM in assessment is in the area of social validation. Social validation grew out of the movement to develop treatments which are clinically relevant to consumers (Kazdin, 1977; Wolf, 1978). Those receiving services as well as those impacted by them are typically queried concerning their perception of various facets of treatment (i.e., did therapy achieve what it set out to). Unfortunately, this form of assessment has received little attention in the school psychology literature. Input from teachers, parents and especially students should be solicited in order to better assess the value of psychological services. In this regard, SM measures could target consumer perceptions of the severity of the problem, collateral changes in other areas of functioning, generalization across settings, and satisfaction with treatment progress. When taken periodically over the course of treatment, SM measures of social validity can be a valuable tool for both providers and recipients of therapeutic services in their formative evaluation of intervention.

An equally important role of SM is in the area of treatment. Numerous authors have reported changes in client target behaviors as a function of self-monitoring alone (see Ciminero et al., 1977; McFall, 1977; Nelson, 1977a, 1977b). This therapeutic aspect of self-monitoring is commonly referred to as reactivity of measurement. Alternatively, a number of researchers have failed to obtain reactive effects, suggesting that the pheonomenon is related to a finite set of variables that influence its occurrence. School psychologists intending to employ self-monitoring for its therapeutic value are advised to include many of the variables associated with reactivity (Nelson, 1977a) and couple self-monitoring with other treatment strategies (Mahoney, 1977). It should be briefly mentioned at this point that an inherent consequence of reactivity is diminished utility of self-monitoring for assessment purposes. The problem is particularly troublesome in research efforts aimed at evaluating discrete treatment variables. Reactivity can confound the data unless its effects are controlled and alternative assessment methods are employed (see section on reactivity in this chapter for further discussion of this topic).

Other potential contributions of self-monitoring to treatment outcome have recently received attention. Haynes and Wilson (1979) speculate that self-monitoring may serve to increase an individual's ability to discriminate the occurrence of target responses. The act of discrimination and subsequent self-recording may interrupt a well-established behavioral chain, thus interfering with maladaptive

behaviors and setting the stage for learning adaptive responses. For example, consider a pupil whose out-of-seat behavior frequently leads to annoying neighboring classmates and eventually culminates in disruptive arguments. Having the child self-monitor out-of-seat responses as they occur engages him/her in an activity incompatible with classroom disruption. Moreover, if the child must return to his/her seat to make the self-recording response, the behavioral chain leading to disruptive arguments is obstructed and academically related behaviors can then be strengthened in the absence of these competing responses.

Recording consequent conditions may reveal that students consistently respond with inattention to a given set of instructional materials. This information may function as feedback to the teacher which in turn may affect his/her future use of the materials. A final consideration is the role self-monitoring may play in enhancing an individual's understanding of the relationship between environmental events and behavior. Learning how the environment controls actions may prompt efforts at self-management of antecedent and consequent events (Karoly, 1977). Similarly, individuals who self-monitor may also develop a greater awareness of the effects their behavior has on others. At the present time, only the effects of reactivity have been empirically verified. The ideas of Haynes and Wilson (1979) elaborated here await similar scrutiny before their contribution to treatment can be determined.

Self-Recording Procedures

Nelson (1977a) noted that almost no research has been conducted examining the critical features of self-recording procedures. Little progress has been made since that time on identifying the relative effectiveness of procedures, examining components within procedures and client preference among procedures, although some efforts are currently under way (Nelson and Hayes, 1981). A critical issue of importance to school psychology is whether self-recording procedures used with adults are applicable to children. Although Kazdin (1974) suggested that self-monitoring may be less appropriate for children than adults, self-monitoring has been successfully employed with youth under a variety of circumstances (McLaughlin, 1976). Yet, the parameters of self-recording with children remain to be systematically investigated. In light of the dearth of literature directed at examining self-recording procedures, the following discussion is based on available research, suggested practices, and logical continuity with behavioral assessment methods.

General Considerations

A number of general considerations arise for selecting a self-recording procedure. Foremost among these concerns is that the self-recording procedure be compatible with the target behavior (Mahoney, 1977). Attention should be given to matching the target response with the measurement procedure (e.g., using frequency counts for discrete behaviors). This match should also take into account the

frequency of the behavior. When the behavior under analysis occurs at a high frequency, it may not be feasible to record the response continuously. In such cases, a time sampling procedure may be indicated, in which case concern is directed at obtaining representative samples of the target behavior. A further issue affecting compatibility is the way in which treatment is expected to affect the target response. Measures should be selected which will be sensitive to the effects of intervention. For example, Johnson and White (1971) assessed the effects of self-monitoring on student study behavior. These researchers considered using the duration of study time as the dependent variable, but decided against it when they found study output more amenable to change than study time.

Other considerations in the selection of a self-recording procedure include ease and acceptability. Recording procedures should be easy to use for several reasons. First, children may have learned to discriminate the occurrence of the target behavior, but may be unable to accurately self-record because of the complexity or awkwardness of the procedure (Nelson, 1977a). Second, difficult recording procedures may interfere with learning adaptive behavior. A student who is asked to record the number of pages read, the content and the number of reading assignments completed, may do so at the expense of other academic learning time. And third, children may become discouraged if the demands of the recording procedure exceed their ability or willingness to comply. This can result in highly inaccurate data or abandoning self-monitoring altogether. A related

factor is the accessibility of the procedure. Frederickson, Epstein and Kosevsky (1975) found that accuracy increased when subjects self-recorded smoking behavior immediately after it occurred compared to subjects recording at daily or weekly intervals. Having the recording materials immediately available should facilitate accurate self-monitoring, particularly for high frequency and/or non-discrete behaviors (Watson and Tharp, 1972).

Direct Methods

Many of the recording procedures available to external observers are applicable for self-recording. Procedures which call on the child to observe the target response as it occurs are termed direct methods. These include narrations, frequency counts, duration measures and time sampling.

Narrations. During the initial stages of assessment, information regarding the nature of the target behavior and the conditions associated with its occurrence is needed for the selection of formal self-recording procedures and the development of an intervention plan (Nelson, 1977a). A time-honored method of collecting such information is the behavioral diary. Typically, the client is asked to provide a narrative account of the target behavior along with a description of events immediately preceding and following it. Of interest is a description of the target response in objective and, where possible, observable terms. Also important in many cases are estimates of the strength, duration and degree of unacceptability to

the client. Antecedent events of concern are the time, setting, persons present, academic subject and materials, emotional state, and client-other interactions. Significant changes in these areas following the occurrence of the target behavior are the focus of narrations about consequent conditions. Changes conceived of as reinforcers or punishers form the basis of hypotheses regarding the variables maintaining the target response. Armed with the above information, the psychologist can then select a formal self-recording procedure which fits the target problem.

Most of the self-monitoring literature employing narrations or diaries has been conducted with adults. The nature of narrative accounts poses obvious constraints on their use with young children; however, many of these problems may not hold for adolescents. Narrative reports have been used to assess anger (Novaco, 1977), eating behaviors (Stuart and Davis, 1972), sleep disorders (Ribordy and Denny, 1977), marital interactions (Knox, 1972), hair pulling (Horne, 1977), anxiety (Mathews and Shaw, 1977), tension headaches (Feuerstein and Adams, 1977) and migraine headaches (Mitchel and White, 1977). The broad range of these behaviors suggests that applications are possible for a large range of behavior problems. Since most of the above disorders have been reported in adolescents, members of this population would appear to be good candidates for future applications and research.

The formats used in recording narrations vary in terms of their degree of structure. In general, the more structure imposed on narrative accounts, the more readily these data may be used to develop

and evaluate interventions. An example of a highly unstructured format is a diary calling for a designation of the date, time and an anecdotal description of a problematic event. While this strategy offers the client considerable flexibility, it often yields little useful information toward a behavior analysis. A more fruitful strategy that succeeds in procuring basic and essential data is the ABC format. Bell and Low (1977) have adapted this format for use in self-recording or direct observational recording. In addition to obtaining a description of the target behavior and its antecedent and consequent events, the format provides for an operational definition of the target behaviors, a description of the setting, and observation dates and times. The major advantage of this form is its general purpose application. Numerous problematic behaviors (both excesses and deficits) can be conceptualized and assessed using this structure. Moreover, it can be readily adapted to collect pertinent information on specific disorders. For example, Feuerstein and Adams (1977) had headache sufferers keep a diary of the frequency, duration and intensity of headaches, and Mitchel and White (1977) asked clients to include a description of symptoms. Mahoney (1977) suggests that structured narrative formats include an open-ended column to allow individuals to self-record significant thoughts, feelings or perceptions which may shed light on the problem.

Written narratives discussed thus far have applications limited to preadolescents and older; however, this form of assessment need not preclude younger children. Kunzelman (1970) has developed a structured

pictorial diary designed especially for elementary school populations. "Countoons" depict in pictures the sequence of events surrounding a particular target behavior. A pictorial sequence might include, for example, the events leading up to hitting another child, making him or her cry and consequently the child being scolded by the teacher. Each time the child observes the sequence, he/she circles consecutive numbers under the "my count" column. This procedure has several advantages worth noting. First, it provides a vehicle for young children to participate in self-monitoring and its associated benefits. Second, it may serve a therapeutic function by helping children learn the relationship between environmental events and their behavior, thus setting the stage for self-management. Moreover, countoons made readily visible to the target child may serve to cue him/her to behave in a positive manner to obtain reinforcers.

To summarize, behavioral diaries serve to gather information about the target behavior and the conditions under which it occurs. In the beginning stages of intervention, narrations help to narrow and define client behaviors and their controlling variables. Although available research on behavioral diaries is limited to adults, this procedure holds promise for adolescents and, when modified, for younger children (e.g., countoons).

Frequency Counts. Frequency counts are an appropriate method of recording behaviors that are discrete (i.e., have an identifiable start and finish) and of relatively consistent duration. Self-monitoring frequency involves recording each occurrence of the target

behavior within a specific period of time (Ciminero et al., 1977). Records may be continuous, that is, reflect every occurrence of the behavior, or sampled from various times and settings throughout the day. Data are often presented as response rates (i.e., the ratio of response frequency to a time interval), permitting comparisons across measurement occasions of varying duration.

This procedure can also be adapted to suit specific assessment needs. When a finite response class can be identified, such as the number of test items or assignments completed, data can be computed as a percentage of the response class (Skinner, 1953). For example, of the semester's 32 arithmetic assignments, Arthur completed 24 or 75 percent of the work assigned. Another application of percentages is under conditions where a specific discriminative stimulus can be identified as controlling a target response (Ciminero et al., 1977). For instance, an obese child may find that weight reduction is possible when between-meal snacks are omitted. The denominator is comprised of the number of times he is offered between meal snacks and the numerator consists of the number of offers accepted.

Frequency counts represent the most prevalent self-recording procedure in the self-monitoring literature. Applications span the gamut of behavior problems encountered in applied psychology. The following examples illustrate applications particularly relevant to school psychology.

Many academic behaviors appear well-suited for self-recording frequency counts. Broden et al. (1971) trained an eighth-grade girl to

record her study behavior at intermittent and self-determined points in time during class periods. Using an a priori definition of studying, the subject recorded a plus or a minus on the basis of her immediately preceding study behavior. Along similar lines, Johnson and White (1971) had college students self-observe and record study output measured by pages read and summarization of readings. In addition, these students calculated points earned for study output and graphed their progress over time. Hundert and Batstone (1978) used a self-scoring procedure with five 9 and 10-year-old boys enrolled in a special education class in a large psychiatric facility. These subjects scored and recorded the number of workbook problems answered correctly and submitted a "report slip" to the teacher on each occasion. Two of four students also self-recorded academic performance in a study by Piersel and Kratochwill (1979). A second-grade girl monitored worksheet scores on a card taped inside her desk and a ninth-grade boy was able to self-chart the number of language arts units completed. A number of studies have employed self-recording in the context of a programmed instruction paradigm. Piersel, Brody and Kratochwill (1977) gave 63 children points for correct responses to WISC-R items. During the administration, subjects self-recorded the accuracy of their responses and the number of points earned. Similarly, mechanical counters were used by college students to record accurate responses to test items in separate studies by Mahoney, et al. (1973), and Kirschenbaum and Karoly (1977). Finally, in conjunction with a token program to increase reading performance, Knapczyk and

Livingston (1973) had 13 students, grades 7-9, self-record the percent of correct responses on their daily reading assignments. A conversion chart was provided to assist subjects in determining percentage values from the number of correct responses.

The frequency of disruptive and maladaptive behaviors has also been the target of self-recording procedures. Maletzky (1974) had an 11-year-old girl self-monitor out-of-seat behaviors using a wrist counter. A similar device was later used with a 9-year-old boy to self-record incidents of hand raising following teacher questions for which he did not know the answer. Broden et al. (1971) employed a self-recording procedure with an eighth-grade boy referred for excessive and disruptive verbalizations. The subject recorded each time he "talked out" without teacher permission on a sheet of paper on which instructions for self-recording were typed. A more complex procedure was used by Bolstad and Johnson (1972). Twenty-four "predelinquent" males, grades 3-6, were instructed in recording their own behavior in three disruptive behavior categories. Each occurrence of talking out, hitting, or leaving their work area without permission was recorded on observation cards. Similar strategies were used for loud talking (Piersel and Kratochwill, 1979), out-of-seat behavior (Bornstein et al. 1977) and verbal interruptions (Piersel and Kratochwill, 1979).

Self-recording the frequency of more serious behavior problems in school age children has also been reported. Rekers and Varni (1977) taught a six-year-old "pretransexual" boy to self-record the occurrence

of play with "masculine" toys. In a similar case, Rekers et al. (1977) used self-observation in the treatment of an eight-year-old male having a gender role problem. Videotape was employed as a teaching strategy to increase the child's discrimination of the occurrence of feminine voice inflections and behavioral mannerisms. Prior to treatment in a case of trichotillomania, McLaughlin and Nay (1977) asked a 17-year-old female to self-monitor the frequency of eyelash and hair pulling incidents. And in 1977, Lubor and Shouse successfully trained a group of adolescents to self-record measures of seizure activity.

A final area of interest to school psychology is the use of frequency counts in the self-recording of social behavior and activities. Nelson et al. (1978) provided a group of four mentally retarded adolescents training in self-monitoring appropriate classroom verbalizations. These researchers found that retarded youths could successfully use both hand-held counters and belt-worn counters to self-record verbalizations. In a similar vein, Gottman and McFall (1972) applied self-monitoring procedures in an intervention to increase classroom participation in a group of inner city high school sophomores characterized as "predelinquent". Half of the subjects were instructed to self-record the frequency of verbal comments while the remainder were asked to mark the number of times they "would like to talk but, for any reason, do not". Prosocial teacher behavior has also been the target of self-monitoring procedures. Nelson, et al. (1977) also asked two teachers to self-record positive verbal comments during their teaching activities. Each teacher was provided a counter to wear

around her neck and was instructed to click the counter each time the target response occurred. In connection with a timeout from teacher interaction program, Plummer, Baer and LeBlanc (1977) taught teachers to discriminate and pace their delivery of instructions and "approval" behaviors to disruptive children. Finally, McKenzie and Rushall (1974) applied self-recording procedures to members of a swimming team. Youngsters ranging in age from 9-16 were instructed to self-record attendance at team practice sessions and laps swum. Recordings were made with a grease marker on a large waterproof display board posted near the swimming pool.

The above literature points to the general utility and wide range of behaviors and populations to which self-recording frequency counts are applicable. Frequency counts have the advantage of being easy to use relative to narrations, duration and time sampling procedures. Their primary limitation is an insensitivity to changes in intensity and duration of the target response. However, when target behaviors are discrete and the focus of intervention is not to alter the intensity or duration of behavior, frequency counts represent a useful method of self-recording.

Duration Measures. For many behaviors, concern about the frequency or intensity of the response may be secondary to the amount of time spent engaged in the activity. Examples of these behaviors include studying, watching television, exercising, headaches, contact with a feared stimulus, and tantrums. In these cases, duration measures indicate the amount of time consumed by the target behavior

(Haynes and Wilson, 1979). Such measures are particularly useful when intervention is aimed at altering the time engaged in the activity. Although frequency and intensity measures may covary with the duration of the response, they are an indirect and potentially invalid index of duration and much less sensitive to change (Nelson, 1977a). Another useful application of duration measures is in combination with frequency or intensity measures (Feuerstein and Adams, 1977; Emmelkamp and Kraanen, 1977). For example, disruptive tantrums may vary along three dimensions: how often, how intense, and how long. Any single measure would not reflect levels in the other two and thus may be inadequate to evaluate the effects of treatment.

Although duration measures may be highly desirable in many instances, few studies employing self-monitoring procedures have used them with school-aged populations. A notable exception is a study by Schwartz (1977) in which 260 seventh-grade students were involved in a summer remedial reading program. A component of the program required students to practice reading each night in order to build their stamina. Each student self-recorded the number of minutes spent reading and number of pages read. Students were able to earn points for their performance which were accumulated to form the basis of their final course grade. Other examples of self-monitoring duration measures appear with adult subjects. Johnson and White (1971) asked a large group of college students to self-record time spent in dating activities (i.e., any recreational activity involving the opposite sex). As part of a treatment package for muscle-contraction and

migraine headaches, Feuerstein and Adams (1977) had clients record the duration of their headaches along with frequency and intensity measures. Turner, Hersen and Bellack (1977) employed a self-observation procedure to assess the duration and frequency of hallucinations. Their female subject was instructed to raise her right index finger at the onset of hearing "voices" and to lower it once the hallucinations ceased. Finally, Leitenberg et al. (1968) had a claustrophobic woman use a stop watch to self-monitor the number of seconds spent in a closed room as a measure of fear reduction.

A number of factors can be identified which may explain the scarcity of self-monitoring literature using duration measures, especially with children. First, compared to frequency counts and intensity ratings, obtaining measures of time is more complex (Watson and Tharp, 1972). For behaviors varying in minutes or seconds, timers or clocks are required to obtain accurate readings. In order for children to effectively use these devices, they must be able to coordinate discrimination of the onset and cessation of the behavior with operating the timer. For many children, particularly special populations, the complexity of the task may be prohibitive. Second, the complexity of the procedure may decrease the reliability of the measures. This may frustrate attempts to achieve good interobserver agreement and diminish the experimenter's power to detect differences. A third factor relates to the high potential duration measures may have for reactivity. According to one theory, the act of self-monitoring is responsible for reactive changes (Nelson and Hayes, 1981). Since

duration measures involve continuous self-observation, reactivity would tend to be enhanced at the expense of the utility of the measure for assessment purposes; however, empirical support for this hypothesis is needed. The limitations cited above should not militate against the use of duration measures, but rather serve as guidelines for minimizing potential problems. Methods for measuring and recording time should be selected for their simplicity (Mahoney, 1977), and adequate preassessment training in the use of devices will enhance reliability (Nelson, 1977b).

Time Sampling Procedures. Time sampling is a convenient strategy for obtaining a representative estimate of the strength of some behavior. These procedures are very useful when the target response occurs at high frequencies. Many times individuals will be unwilling to self-record occurrences of a high rate behavior. Moreover, repetitive self-recording may actually interfere with learning adaptive responses. In these cases, samples of behavior can serve as adequate measures and minimize the aversive side effects of continuous recording.

Ciminero et al. (1977) have identified four prevalent time sampling procedures used in self-monitoring. The first of these, the "all or none" method (Thoresen and Mahoney, 1974), involves dividing an observation period into equal intervals. During each interval, the client merely records whether the target response occurred. Intervals are coded as 1 - occurred or 0 - did not occur, and are thus insensitive to variable frequency rates. The percentage of intervals

in which the target behavior occurred is used to assess response strength. Therefore, the length of each interval must be short enough to detect nonoccurrences, yet long enough to avoid tedious recording.

A second procedure is a variation of the "all or none" method. The dichotomous coding system (occurrence/nonoccurrence) is replaced by an ordinal scale which rates the frequency of occurrence. An example reported by Ciminero et al. (1977) is a 4-point self-rating scale developed by Stumphauzer (1974). At the end of each interval, the target behavior is rated as 0, 1, 2, or 3 on the basis of whether the behavior occurred "never", "occasionally", "often", or "very frequently". This procedure has the advantage of providing a more fine-grained measure which retains information lost in a dichotomous system. However, as the length of the interval increases, measurement errors become more common (Tasto, 1977).

Another procedure used to sample high frequency behaviors involves self-recording target responses in a limited number of observation periods. Ideally, blocks of time are selected randomly from the time available for self-monitoring. For example, class time would be broken down into 10 half-hour periods. Each day, two of the ten periods are chosen via a random procedure. Target responses would only be self-recorded during the selected periods. Because the time of the half-hour periods would vary from day to day, a representative sample of time intervals is assured. Moreover, using random rather than arbitrary methods of interval selection guards against obtaining biased samples resulting from behaviors that may occur at higher frequencies during certain times of the day.

The fourth time sampling procedure useful for high rate behaviors is known as spot checking. Typically a timing device is set on a variable interval schedule. Each time the device signals the end of the interval, the subject self-records whether she/he was engaged in the target response at the time of the signal. Data are computed as the percentage of spot checks in which the behavior occurred. Since assessment occurs at points in time rather than within time intervals, it is important that the target behavior consume a period of time rather than occur at an instant. For example, spot checking is an appropriate and common strategy for assessing on-task behavior because random "checks" will likely "catch" the student engaging in that behavior if it occurs. In contrast, it is unlikely that a behavior occurring for a brief instant (e.g., a facial tic) will coincide with spot checks. For behaviors such as these, an interval sampling procedure will be more sensitive to the behavior's occurrence.

Some examples from literature pertinent to school psychology will help illustrate the applicaton of time sampling procedures. In conjunction with a self-control project, Glynn, Thomas and Shee (1973) instructed a class of second-grade pupils in self-recording on-task behavior. A tape recorded series of intermittent "beeps" varied randomly between one and five minutes so that children would be unable to predict when the tone would occur. Each time the "beep" sounded, the students placed a mark in a box on a self-recording sheet if they were "on-task". Behavioral definitions of on-task behavior served as the criteria for judging its occurrence. This procedure was later

replicated by Glynn and Thomas (1974) with a class of third graders. However, these researchers found it beneficial to operationalize on-task behavior specific to the academic task. Two definitions were printed on cue cards, one relating to teacher lectures and the other to independent seat work. One of the two cards was displayed in the front of the classroom, indicating the target criteria at any given time. In another study, Seymour and Stokes (1976) trained four delinquent girls in a detention center for adolescents to self-record work behavior. Target behaviors were defined and written on index cards and included the following: hands working continuously, looking at work, doing the work set and not something else, not fiddling when getting something, not dawdling, and not talking out of turn. Work periods were divided into 3-minute intervals. Girls were instructed to mark their cards during each interval in which they judged their behavior to be consistent with the work criteria.

Indirect Methods

The self-monitoring procedures discussed thus far involve self-observation and recording of the target response as it occurs. However, in many cases it is not feasible or desirable to obtain direct measures for a variety of reasons. Teachers, students and parents may find that self-recording narrations, frequency counts or duration measures may interfere with their primary task. When this occurs, indirect methods are a viable alternative because they require less or no effort during the time in which the target behavior occurs. Another consideration is the potential for reactivity in direct versus indirect

methods. Because reactivity is closely associated with the self-recording act (Nelson and Hayes, 1981), recording procedures that are further removed from the behavior may be expected to be less reactive (Mahoney et al., 1973; Romanczyk, 1974). These concerns may lead the psychologist to select a self-rating procedure or self-recording of behavioral traces or by-products.

Self-Ratings. A subtle distinction exists between self-report and self-ratings as part of a self-monitoring procedure. The distinction is one of the degree of removal from the observed behavior. Self-report methods require the subject to engage in retrospective observations of the target behavior and recall it accurately in response to a question or rating scale (Bellack and Hersen, 1977). The time between the occurrence of the behavior and his or her self-report can range anywhere between hours to several years (Tasto, 1977). On the other hand, self-ratings in the context of a self-monitoring procedure are generally closely linked to the actual behavior (Stumphauzer, 1974). Typically, subjects are asked to rate their behavior along some dimension during time intervals or judge the degree to which their behavior approximates a predetermined criterion. Self-rating requires considerably less effort than direct methods calling for a one-to-one correspondence between behavior and self-recording. This characteristic of self-ratings make them particularly attractive for use in the schools where teacher and student time is at a premium.

Several studies demonstrate the application of self-ratings with children and youths. As part of a program to teach children to print, Jones, Trap and Cooper (1977) trained a group of first grade students to self-rate and record the accuracy of their letter strokes. Children placed overlays of model letters on top of their written letters. Students self-evaluated their performance by tracing over letter strokes that "matched" the model overlay. Strokes rated as "matching" were then counted by the students and the number was written at the right edge of the self-recording sheets. Flowers (1972) made use of a self-grading procedure as part of an intervention to eliminate cheating in a school-aged girl. The student was allowed to grade her own daily assignments using the teacher's manual. Accuracy of her self-grading was improved by making her exam grades contingent on agreement between peformance on daily assignments and weekly exams (comprised of items from the daily assignments). Hundert and Bucher (1978) found self-rating of arithmetic assignments to be a useful strategy for reducing teacher time spent in grading. As with the Flower (1972) case, accuracy of student self-rating was enhanced through the application of positive contingencies.

Institutional settings have also been the scene of interventions incorporating self-ratings in assessment. Prior to implementing a teacher controlled token system, Santogrossi et al. (1973) instructed subjects in the self-evaluation of their classroom behavior. Nine adolescent males enrolled in a remedial reading class in a psychiatric hospital self-rated their compliance with classroom

rules. At the end of consecutive 15-minute intervals, each youth rated their degree of adherence to rules on a scale from 0-2 for each of five rules. Subjects publicly announced their self-ratings in an attempt to stem overevaluations. During later phases, self-rating was combined with a token system in which students determined the number of tokens they received. Following several phases of a token system, Kaufman and O'Leary (1972) employed a self-evaluation phase to transfer the responsibility for evaluation of pupil behavior from the teacher to the students. Adolescents in a psychiatric hospital school self-rated their disruptive classroom behavior according to 11 behavioral criteria and self-determined the amount of tokens earned for each rating. Finally, Fixen et al. (1972) investigated the effects of self-ratings with a group of "predelinquent" boys placed in a family-style rehabilitation program. Each youth was provided with a list of 21 room cleaning definitions. The boys were instructed to rate the condition of their bedrooms each morning according to each of the 21 criteria.

Behavioral Traces and Archival Records. Another indirect means of self-assessment is by observing the products of one's own behavior. Products of behavior refer to physical traces or records of performance which exist independent of formal assessment efforts (Kazdin, 1979). Behavioral traces or archival records may either serve as the primary dependent measure or as supplements to other direct methods of assessment (Nelson, 1981). For example, worksheet scores can provide an indirect measure of on-task behavior, home study, learning, and motivation. Other applications include a measure of weight as a

by-product of eating (Monti, McCrady and Barlow, 1977), nurse records as a measure of stomach complaints (Miller and Kratochwill, 1979), hair length as an indicator of hair pulling (Anthony, 1978), fingernail length as a measure of nailbiting (McNamara, 1972), and the number of pills in a prescription bottle to assess compliance with a medical regimen (Dapcich-Miura and Hovell, 1979).

There are a number of advantages in using behavioral traces and archival records for self-monitoring purposes. First, a modicum of effort is required on the part of the client, making it a realistic dependent measure for use in applied settings (Nelson, 1981). Piersel (in press) suggests that psychologists' efforts to trim demands on teachers' time is an effective strategy to bolster support for classroom interventions. Second, as supplemental measures, behavioral traces and records contribute to the validity and reliability of assessment by providing multiple indices of behavior (Nelson, 1981). Synchronous measures enhance one's confidence in the observed effects, while desynchrony among measures may provide greater conceptual understanding of the problem (Barlow, 1981). A third advantage is that self-monitoring traces and records of behavior may be less reactive than direct measures (Nelson, 1977a). Because subjects are not required to self-observe and record the actual target behavior, the self-recording response is less likely to influence subsequent behavior (Nelson and Hayes, 1981).

Kazdin (1979) has observed that the apparent advantages of behavioral traces and archival records should be tempered by their

potential shortcomings. Foremost among these is the problem of deciding what is measured. Varying degrees of inference are required to pair the behavioral by-product or record with the behavior that caused it. For instance, using weight gain to make inferences to eating behavior in anorexic children may be erroneous when eating is followed by self-induced vomiting. Similarly, increases in worksheet scores may be a function of teaching strategies, increased study time, or cheating. Such confusion may be reduced by using direct assessment methods initially to identify functionally related variables. Then, a reliably occurring behavioral trace or record may be used to infer the occurrence of the target response. Occasional probes using the original direct methods may be used to ensure that the variables continue to be functionally related (e.g., worksheet scores continue to be a result of teaching methods rather than cheating).

Self-Recording Devices

Several self-recording devices are available to assist clients in obtaining accurate records of their behavior. Ciminero et al. (1977) provide a number of considerations for selecting a recording device. The availability of the device at the time the target behavior occurs is a major concern. Particularly with children, errors can be expected to increase the longer the interval is between the occurrence of a behavior and the act of self-recording. Another consideration is that the self-recording device be easy to use. Elaborate devices may confuse younger or retarded clients rendering the data collected unreliable or invalid. Finally, the obtrusiveness of the self-

recording device can affect the quality of the data obtained. A moderately conspicuous device may improve accuracy by serving as a discriminative stimulus for self-recording. However, classmates and teachers are more likely to attend to salient devices and alter the child's use of them. With these considerations in mind, a sampling of devices will be described. For a more detailed discussion the reader is referred to Ciminero et al. (1977).

Mechanical Devices. Mechanical devices available for self-monitoring fall into three categories: counters, timers and tape recorders. Counters are used to record the frequency of some behavior and take a variety of forms. For example, Lindsley (1968) reported the use of a wrist-worn golf counter; Hannum, Thoresen and Hubbard (1974) used a multiple event wrist counter and Thoresen and Mahoney (1974) described a pipe cleaner wrist counter composed of multiple pipe cleaners, each with nine beads, mounted on a wrist band. A potentially useful counter for children is the "Knit Talley" described by Sheehan and Casey (1974). When mounted on a child's pencil or desk top, this device is a relatively unobtrusive and portable method of event recording. Another innovative device is the pressure sensitive pad and frequency counter. This device may be used for a variety of purposes including out-of-seat behaviors, sleepwalking and exercising programs. The directional switch can be set to be activated each time the pad is depressed or conversely each time pressure is removed. Timing devices also take on a variety of shapes and sizes and may be used to self-record duration measures or signal recording times using a time

sampling procedure. Among the timing instruments available are stopwatches (Ciminero et al., 1977), wrist watches having elapsed time indicators (Katz, 1973), kitchen timers (Rainwater and Ayllon, 1976) and electric clocks with an in-cord/on-off switch (Thoresen and Mahoney, 1974). And finally, with the advent of highly portable video and audiotape recording devices, self-monitoring may occur outside the target setting or time. Thomas (1971) used video recordings of teachers' classroom behaviors to facilitate self-monitoring and subsequent changes in teaching style. Also, Litrownik, Freitas and Franzini (1978) combined videotape and live models in a procedure to train retarded children in self-monitoring.

Graphing Performance. Once self-monitoring data have been collected, it is often useful to transfer the information to a behavior graph or chart (Thoresen and Mahoney, 1974). The client can be instructed in the essentials of graphing and be encouraged to assume responsibility for its maintenance. Graphing performance serves a number of functions. First, recalling the Chinese proverb, "One picture is worth a thousand words", aptly reflects the value of the graphic display of data (Parsonson and Baer, 1978). The graph summarizes performance in a manner that is usually more accurate and easier to comprehend than similar information provided verbally (Bergan, 1977). Second, data points connected by straight lines indicate the level and possible trends in performance. This information is crucial to the task of evaluating the effects of intervention. Moreover, data displayed across time can alert the

psychologist and client to possible variables influencing performance and lead to subsequent changes in the treatment strategy. Third, graphs used in conjunction with a self-monitoring procedure can serve a dual purpose: 1) to summarize information, and 2) provide a means of recording the target behavior as it occurs. For example, Plachetta (1976) used "self-charting" as part of a treatment package for encopresis in a 6-year-old boy. The graph doubled as a self-recording form plus an illustration of improvement in the child's condition. A fourth function of graphing is as a potential element of treatment. Public display of data may provoke feedback from others or individuals themselves that may have reinforcing or punishing effects (Ciminero et al., 1977). To the extent that this occurs, subsequent performance may be altered.

Bergan (1977) has outlined the essentials of graphing in terminology consistent with his consultation model of service delivery. The first characteristic of a readable graph is a title that describes what the graph represents. Included in the title should be a designation of who is being observed and the behavior that is being recorded. Another important graph component is the appropriate labeling of the vertical (ordinate) and horizontal (abscissa) axes. Ordinate labels should identify the behavior being recorded, the type of strength measure being used (e.g., frequency or duration), and the time period each datum point represents. The abscissa should specify the dates when observations or self-monitoring occurs. Bergan (1977) suggests using dates rather than days or sessions to avoid losing

information regarding interruptions in the monitoring of behavior. A final essential of graphing is to indicate the phase of the intervention in which data were collected. For example, the initial phase may be labelled baseline and may represent self-monitored performance or data collected by an independent observer. Subsequent phases may include various treatment components or follow-up periods. When using behavior graphs with younger children, some important steps may be taken to facilitate their use. All graph labels should be completed by the psychologist or person responsible for intervention in order to minimize demands on the child and avert possible confusion.

Summary

A broad range of self-monitoring procedures are available to the clinician and researcher. Direct methods in which the client self-records target responses as they occur include narrations, frequency counts, duration measures and most uses of time-sampling. Narrations or behavioral diaries are useful during the problem identification stage of intervention to identify target behaviors, how they will be measured, and the environmental variables which control them. Frequency counts are most applicable for discrete target behaviors of relatively consistent duration and when treatment is expected to impact on the number of times a behavior occurs. Duration measures, on the other hand, are appropriate when the time engaged in the behavior is variable or when concern for frequency is secondary to how much time the response consumes (e.g., contact time with a feared object). For high rate or highly variable behaviors, time sampling

procedures may provide adequate estimates of response strength and reduce the time required for self-monitoring. Indirect methods discussed in this section include self-ratings, behavioral traces and archival records. These methods are appropriate when direct assessment is impractical or as supplemental measures.

Several self-recording devices are available to assist children and their caretakers in collecting self-monitoring data. Their value lies in the simplification and organization of complex information, allowing for evaluation of treatment effects. Portable mechanical devices offer a convenient method of recording events as they occur, while behavior graphs serve to summarize data across treatment phases. In addition, both kinds of recording devices may influence the reactivity of self-monitored data and consideration should be given to this factor when designing self-assessment strategies (see section on reactivity in this chapter).

Training in Self-Monitoring

An important issue related to any assessment technique is that of training assessors in the proper use of the data collection procedure. Until recently, relatively little attention has been given to training clients, particularly children, in the use of self-monitoring procedures. In the child behavior therapy literature, training has typically consisted of verbal instructions describing the self-recording methods and providing clients with recording forms (e.g., Bornstein et al., 1977; Sagotsky et al., 1978; and Piersel and Kratochwill, 1979). Others have indicated praise for appropriate

monitoring behavior (Drabman, Spitalnik and O'Leary, 1973) or modeled the actual recording methods (Anthony, 1978; Hallahan, et al., 1979). In addition to the above training strategies, Litrownik and Freitas (1980) established performance criteria for each phase of training. Those clients meeting or exceeding the criterion for each phase progressed through subsequent phases until training was completed. Further training was provided to clients whose performance did not reach the performance standard. Jones et al. (1977) developed an innovative training procedure to teach children to self-record letter strokes. The experimenter employed verbal instructions, demonstrations and verbal performance feedback to train subjects to: 1) align the overlay of model letters above practice work samples; 2) identify letter stroke errors on sample training sheets; 3) trace the overlay on practice sheets of target letters; 4) count the number of correct letter strokes for each practice letter and self-record the number on the practice sheet; and 5) apply this procedure to self-monitor the accuracy of their own work samples.

Another issue parallel to the development of training procedures is the effects of training on the accuracy of self-monitored data. Addressing this question, Litrownik et al. (1978) assessed the self-monitoring performance of retarded children under three conditions. The group receiving training (i.e., verbal instructions, live and video models, and practice trials with performance criteria) acquired accurate self-monitoring skills compared to subjects in either of the attention or no-contact control groups. In a similar study,

Nelson et al. (1978) found that having retarded adolescents practice self-recording from videotapes and in the criterion classroom setting enhanced the accuracy of self-monitoring relative to no-training controls. Interestingly, the effects of training were not found to increase the reactivity of the self-recorded measures.

Although recent advances in training strategies have appeared in the applied literature, few authors have proposed a general training sequence of relevance to both practitioners and researchers alike. The training sequence proposed here represents an extension of the procedure described by Mahoney (1977). It is based primarily on procedures detailed in the self-monitoring literature, sound behavioral practice, and experience training youth in the use of self-monitoring. Conceptually, self-monitoring, like any other skill, is viewed as a coordinated set of behaviors, each of which can be targeted for acquisition. Usually, this involves some sort of task analysis, breaking the entire task into discrete units and their sequential ordering (Sulzer-Azaroff and Mayer, 1977). Because each application of self-monitoring will vary due to differing subjects, target behaviors and assessment goals, training should be individualized for each client or group of clients and their specific needs. Yet despite the unique features of each case, some general guidelines can be identified which have broad applications. The proposed training sequence consists of the following:

1) Provide clear and simple definitions of the target events to be self-recorded. Definitions can be written, represented

pictorially or both depending on the client's level of functioning. Post the definition on the recording form to reduce ambiguity in self-assessment.

2) Provide clear and simple instructions on how to self-monitor. Identify each step and the behavior to be performed (e.g., First, you set the timer for 25 minutes like this (demonstrate). Second, do as many math problems correctly on the worksheet as you can until the timer rings. Third, stop working when the timer rings and mark your answers using the answer key. Fourth, count the number of problems you completed correctly. And fifth, write that number in the box next to today's date on the record form.

3) Demonstrate or model the self-monitoring procedure using actual record forms or recording devices. Label each step as it is completed.

4) Provide the client with written or pictorial instructions describing the self-monitoring procedure.

5) Ask the client to repeat the target definitions and self-monitoring instructions. Ask the client if he or she has any questions about what is to be done.

6) Test client acquisition of self-monitoring skills by conducting several assessment trials in an analogue or if possible the target setting.

7) Specify performance criteria for self-monitoring accuracy. Use an upwardly changing criterion for clients slow to acquire the assessment skills.

8) Provide social and/or tangible reinforcement, as needed, to promote the development and maintenance of self-monitoring.

This sequence can be used to train children, parents and school personnel to self-monitor a broad spectrum of target behaviors. Although elements were derived from empirical work in the area, research is needed to evaluate the combined effects on both the acquisition and occurrence of self-monitoring skills. Also of interest is the contribution of the individual components of the training package to its overall efficacy. Research in this area may help streamline the training sequence for easier use in applied settings.

The Relationship of Self-Monitoring to Other Assessment Procedures

The utility of self-monitoring in child assessment is perhaps best understood relative to other assessment devices. In selecting an assessment procedure, the psychologist is concerned with its objective characteristics as well as its suitability for the specific case. In this regard, self-monitoring has a number of characteristics which merit consideration when choosing among assessment alternatives.

First, self-monitoring can be viewed as a humanistic form of measuring client performance. Compared to other modes of assessment, self-monitoring requires active participation of the client in his/her own evaluation. This essentially eliminates second-party assessor bias from the data. As a result, subsequent diagnosis and/or classification based on self-collected data reflect the client's input and evaluation of his/her functioning. In addition, self-monitoring can be considered a less intrusive form of assessment than the use of external observers

(Morris and Hoschouer, 1980). Whenever children record their own behavior they are aware that assessment is taking place and provide consent to proceed either explicity or implicity. Unfortunately, children are not always accorded this "privilege" when other forms of assessment are used. Another humanistic aspect of self-monitoring is its emphasis on skill development. Clients who learn to self-record their own behavior acquire valuable skills which have broad application. The child who self-monitors arithmetic performance may extend its use to other target behaviors which he/she may find desirable to change.

A second characteristic of self-monitoring is its relative cost-effectiveness. Continued emphasis on the consultation model of service delivery in school psychology has highlighted the need to involve educational personnel in assessment and treatment (Bergan, 1977; Piersel, in press). Consistent with this trend, self-monitoring provides the opportunity for the client to share in this responsibility. This prospect has a number of potential advantages. Clearly, enlisting youth in the assessment process reduces the demands on the psychologist and other educational staff. Greater numbers of students can then be served, thereby maximally utilizing the consultant's knowledge (Bergan, 1977). Piersel (in press) has also noted that consultant attention to minimizing teacher time required for assessment is a key variable which can predict the eventual outcome of consultation. From this perspective, self-monitoring appears considerably more cost-effective than other procedures requiring external assessors.

A final aspect of self-monitoring which should influence its selection for assessment is the quality of data it generates. As previously indicated, self-monitored data have a high potential for being inaccurate and reactive relative to other observation methods (Nelson, 1977b; Nelson et al., 1978). It is strongly recommended, therefore, that the variables shown to enhance the data's validity be included in any self-monitoring procedure. A related issue raised by Haynes and Wilson (1979) is whether even accurate self-monitoring reflects client performance when assessment is not being conducted. For example, a student self-recording on-task behavior during spelling and math may evidence rates near 80%. However, during reading and language arts this rate may be considerably lower (e.g., 30%) and representative of preassessment levels. Discrepancies such as these may be due in large part to the reactivity of the assessment procedure. Consequently, data derived from this method may not be generalizable outside actual assessment occasions. On the positive side, private events such as cognitions, emotions, and certain social relations are perhaps most validly assessed via self-monitoring (Nelson, 1977b). Compared to retrospective methods of self-report, self-monitoring is more direct and less likely to be influenced by bias and limited recall. However, in view of the inherent difficulty in assessing the accuracy of private events, these measures should be supplemented where possible with other sources of information.

Methodological Issues

Many of the issues relating to self-monitoring discussed thus far speak to its advantages as an assessment technique. However, two

primary methodological issues must be considered to put its role in assessment into perspective. The first concerns the potential reactive effects associated with self-monitoring, that is, changes in performance occurring as a function of the assessment procedure alone. A second issue pertains to the accuracy of criterion-related validity of the measurement. In the following sections, reactivity and validity will be examined along with the variables which influence them.

Reactivity of Measures

When external observers are introduced into a natural setting, the stimulus conditions exerting control over behavior undergo significant changes. In many cases, the presence of observers has been demonstrated to alter client behavior independent of formal treatment (Kazdin, 1974; Mercatoris and Craighead, 1974). Yet, other authors have found the reactive effects of observation less problematic (see Nelson et al., 1978). The issue of reactivity also becomes a factor when clients are engaged in self-monitoring. The act of self-observation and self-recording results in similar changes in the stimulus situation and the sequence of events that would otherwise exist (Ciminero et al., 1977). Resulting behavior changes are referred to as the reactivity of self-monitoring.

For the school psychologist employing self-monitoring procedures, the consequences of reactivity are a mixed blessing. On one hand, reactive effects may be therapeutic to the extent that behavior change parallels the goals of treatment. In this regard, self-monitoring may either be used as the sole therapeutic variable or

integrated into a comprehensive treatment package. However, when self-monitoring is used primarily to assess and evaluate the effects of a specific treatment, reactivity becomes problematic (Kazdin, 1974). Simultaneous application of self-monitoring and another intervention leads to confounded data, making it impossible to ascertain the contribution of either variable. For example, Hundert and Batstone (1978) had five boys enrolled in a special education class score their own arithmetic workbook. The number of problems answered correctly was self-recorded on a "report slip" and submitted to the teacher on a daily basis. Concurrently, pupils were awarded points exchangable for classroom privileges based on their self-reported arithmetic performance. Using a multiple baseline across subjects design, the percentage of correct problems was found to increase under the reinforced self-recording condition relative to teacher scored baseline performance. While these authors have evidence for a clinically relevant treatment package, the results are ambiguous in terms of the variables responsible for change.

Experimental Controls for Reactivity

In circumstances where there is concern for identifying the source of behavior change, the reactive effects of self-monitoring will need to be accounted for (Nelson, 1977a). Fortunately, there are a number of design options available for this purpose.

When time-series data are collected, Jeffrey (1974) suggested using an A/B/C/A/B/C within-subject experimental design to isolate the effects of reactivity. In this sequence, A is an independently

measured baseline, B is self-monitoring and C is a combination of self-monitoring plus another intervention. This design strategy is appropriate when A equals B (i.e., no reactivity is apparent); otherwise the design is confounded by sequence effects (Hersen and Barlow, 1976) and contiguous treatment effects (Kazdin, 1980). When reactivity is present, an A/B/A/C/A/C series provides some control for the problems of continuous treatments, although multiple treatment interference is still a problem. Obtaining data across other settings or subjects in a multiple baseline pattern allows the series to be counterbalanced, thus accounting for sequence effects (e.g., subject 1 receives A/B/A/C while subject 2 receives A/C/A/B). Nelson (1977a) noted that in some instances acquiring data for the A phase may prove troublesome. The unavailability of independent assessments may prompt the use of client estimates of preintervention levels of behavior (see Berecz, 1972). Obviously, retrospective baseline estimates are less desirable than direct measures; however, they may provide a rough means of gauging the presence of reactivity.

When using between-subject experimental designs, the use of control groups is a common method of examining the contribution of reactivity. Nelson and McReynolds (1971) and Jeffrey (1974) suggested including a self-monitoring only group in addition to no-treatment and experimental treatment plus self-monitoring groups. As Nelson (1977a) points out, this strategy does not permit evaluation of the independent effects of treatment, only the interaction between treatment and self-monitoring relative to self-monitoring only. One solution to this

problem is the completely crossed factorial design (see Myers, 1979). Four groups of subjects are exposed to separate conditions. In Cell 1, clients receive the package of self-monitoring plus an additional intervention, allowing for examination of their interactive effects. Cells 2 and 3 permit evaluation of the independent contributions of treatment and self-monitoring, respectively. Subjects in the fourth cell serve as no treatment controls against which the experimental conditions are compared. Although subject to the inherent limitations of group designs (see Hersen and Barlow, 1976; and Kratochwill, 1978), this design strategy does allow between condition comparisons in the absence of multiple treatment interference. Furthermore, the design can be strengthened by including a trial factor to facilitate assessment of effects over time.

Evidence of Reactivity

The reactive elements of self-monitoring have been demonstrated across a broad range of subjects, settings and disorders (Haynes and Wilson, 1979). In the child literature, self-monitoring has been used to initiate desirable behavior change (Roberts and Dick, in press; Broden et al., 1971), as well as in the maintenance of effects following intervention (Drabman et al., 1973; Bornstein, et al., 1977). In addition to playing a supplemental role in treatment, self-monitoring has been used as the primary therapeutic variable. It is in this capacity that the effects of reactivity are most clearly evident.

A case study presented by Hallahan et al., (1979) illustrates the therapeutic potential of self-monitoring. A 7-year-old boy with attentional problems was taught to self-monitor his on and off-task behavior using a tape recorder to cue self-recording responses. During handwriting and math periods, the child was provided with a self-recording sheet containing a picture of a boy reading a book and the caption, "Was I paying attention?" Each time a tape recorded tone sounded (on a variable interval schedule), the student was to self-record whether he was on or off-task at the time of the tone. Using a time-series withdrawal design across handwriting and math periods, on-task behavior increased markedly relative to baseline periods in both settings. In subsequent phases, behavior change remained stable as self-monitoring was phased out and substituted with a covert self-praise procedure.

In another case study by Anthony (1978), a 9-year-old boy referred for trichotillomania (excessive hair pulling) was instructed in the use of a wrist counter to self-record the number of times he pulled his hair and the number of times he started to pull his hair but stopped. The child was also taught to graph his self-recorded data. In addition to self-monitoring, the client wore a baseball cap during waking hours to discourage hair pulling responses. Initially, self-monitoring was to serve primarily as an assessment procedure to be followed by a token economy and relaxation training. However, the client failed to return for treatment following instruction in self-monitoring. Eight weeks later, the therapist received a letter of

thanks from the child's mother along with data indicating that hair pulling had ceased. Independent observations made by the mother in addition to measures of hair length were offered in support of her claims. Although lacking an experimental design and control procedures, this case does raise compelling questions about the potential strength of reactive effects.

A series of four single-case studies reported by Piersel and Kratochwill (1979) further exemplify the therapeutic value of reactivity. In the first of these, a 7-year-old girl increased the number of completed worksheets across two academic subjects from a baseline of zero to approximately one per day during self-monitoring. The self-recording strategy involved writing her worksheet scores on a card taped inside her desk. The second referral was a 13-year-old boy who was taught to self-record his disruptive talk in the classroom. Teacher statements to "tone it down" were recorded by the student as an indirect measure of his disruptive verbalizations. Self-recorded data were found to be highly reactive as well as accurate over the course of intervention. Case 3 involved having a 9-year-old girl identified as "hyperactive" self-record the frequency of her verbal interruptions. The level of the target behavior dropped significantly with the onset of self-monitoring. Interestingly, independent teacher ratings were found to be more "reactive" than student self-monitored data. Finally, a 15-year-old hyperactive boy was instructed to monitor the number of language units completed accurately on a sheet of notebook paper. His performance improved dramatically and was maintained by self-monitoring

over a period of 50 consecutive school days. The importance of these cases lies in demonstrating strong and often enduring reactive effects after other staff-planned strategies had failed.

When the therapeutic impact of self-monitoring is compared with conventional treatments, mixed results have occurred. Nelson, Lipinski and Black (1976) contrasted the effectiveness of a token economy with self-monitoring on increasing community living skills in retarded adults. Across multiple treatment phases, self-recording was found to be superior in altering client behavior. Similarly, Sagotsky et al. (1978) found the benefits of self-monitoring to significantly exceed those achieved by goal-setting on children's academic on-task behavior. In contrast to these findings, Mahoney, Moura, and Wade (1973) and Mahoney (1974) reported self-monitoring alone to be less effective for weight reduction than a combination of self-monitoring and self-reward. Moreover, other studies have found minimal or no reactive properties associated with self-monitoring (e.g., Fixen et al., 1972; Miller and Kratochwill, 1979; and Romanczyk, 1974).

These notable inconsistencies across studies have important implications for both clinical practice and research. For the practicing school psychologist, the primary goal of intervention is to effect clinically significant changes in behavior. In most cases, this is often achieved by integrating a number of therapeutic strategies into a comprehensive treatment package. Concern for the contribution of the individual components is typically subordinate to the need to produce meaningful results. For this reason, the reactive potential of

self-monitoring may be best utilized in combination with other treatment approaches, rather than as the solitary therapeutic variable.

While inconsistent reports of reactivity create ambiguity for the practitioner, these findings have led to a wave of research aimed at explaining these differences (Nelson, 1977b). Efforts have concentrated on identifying the variables which enhance reactive effects. Since the majority of work in this area has been conducted with adults, appropriate caution should be exercised in generalizing to children.

Variables Affecting the Reactivity of Self-Monitoring

Most researchers have conceputalized the occurrence, extent and duration of reactive effects from self-monitoring as dependent variables under the control of other factors (Haynes and Wilson, 1979). Nelson (1977a, 1977b) has reviewed this literature and identified nine variables implicated in the reactive process. These nine variables are briefly discussed below.

Motivation. The extent to which subjects are motivated to alter the target response appears related to reactivity. Evidence for this comes from a number of studies in which smoking behavior was self-monitored. McFall and Hammen (1971) included only those smokers who were motivated to reduce their habit. All four groups in the study decreased their smoking regardless of the smoking measure they self-recorded. In a direct comparison of motivated versus nonmotivated smokers, Lipinski et al. (1975) found that self-recording decreased

smoking only in subjects volunteering for an experiment to reduce smoking. Reactivity did not occur in subjects volunteering for a general experiment for smokers.

Valence. Several investigators have determined that the valence associated with the self-monitored target behavior can affect the occurrence and direction of reactivity. Specifically, self-monitoring tends to increase the strength of desirable behaviors and decrease the strength of undesirable ones. For example, Nelson et al. (1977) had teachers self-record their positive and negative statements during different phases of a study. Positive verbalizations increased while negative comments decreased relative to their baseline rates. Similar results were obtained by Litrownik and Freitas (1980) in a study with moderately retarded adolescents. Participants were divided into four groups: 1) self-record when they finished a bead-string task within a time limit (positive aspect); 2) self-record when they did not finish the task within the time limit (negative aspect); 3) self-record when they strung the beads (neutral aspect); and 4) string beads without self-recording. The results indicated that youths recording positive or neutral aspects of their behavior out-performed the negative aspect and control groups.

Findings obtained by others examining the effects of self-monitoring negatively valenced behaviors indicate a potential for undesirable side-effects. Self-recording negative behaviors was associated with lower self-evaluations and performance anxiety

(Kirschenbaum and Karoly, 1977), decreased accuracy in performance (Cavior and Marabotto, 1976), and negative affect (Piersel et al., 1977). Similar undesirable concomitants have not been reported with positively valenced target behaviors.

Experimental Instructions. Another variable examined is the effects of experimenter-provided expectations concerning the direction of behavior change to result from self-monitoring. Nelson, Lipinski and Black (1975) informed different groups of college students that the frequency of face touching would either increase, decrease, or remain unchanged as a function of self-monitoring. However, all groups experienced a decrease in the rate of face touching, suggesting that the salience of experimenter instructions may have been overshadowed by personal valences or motivation (Ciminero et al., 1977). Similar outcomes were reported by Hutzell (1976), Orne (1970), and Nelson et al. (1978) suggesting that experimenter induced expectancies have a negligible effect on reactivity. The notable exception is experimenter-assigned valence to target behaviors discussed earlier.

Target Behaviors. The nature of the target behavior used in self-monitoring may also influence the reactivity of the measures. Hayes and Cavior (1977) compared the relative reactivity of self-recording verbal value judgments, verbal nonfluencies and facial touching. Of the three behaviors, only verbal nonfluencies and facial touching were reactive. In a weight control study, Romanczyk (1974) found that greater weight loss occurred when both daily weight and

caloric intake were recorded as compared to monitoring daily weight only. And a recent study by Harmon et al. (1980) reported improved ratings of depression and involvement in pleasant activities when either mood or activity engaged in were self-monitored. However, recording activity resulted in somewhat greater reactivity than when mood was the target response. Unfortunately, few patterns emerge that allow prediction of which target behaviors will be most reactive, although some authors suggest that nonverbal behaviors are more reactive than verbal ones (Peterson, House and Alford, 1975) and monitoring urges is more effective than monitoring the actual behavior (Gottman and McFall, 1972).

Goals, Reinforcement and Feedback. Setting performance goals and the availability of performance feedback and reinforcement tend to enhance reactivity. Kazdin (1974b) found that providing subjects with a performance standard when self-monitoring was more effective than self-monitoring alone in increasing subjects' use of target pronouns. Response frequency was further increased when feedback was available. Reinforcement contingent on self-monitored performance has similarly heightened reactive effects. Changes in self-recorded behavior were reported when college students were given money for decreasing facial touches (Lipinski et al., 1975) and prompt access to breakfast was made contingent on children's tent cleaning (Lyman, Rickard and Edler, 1975). In short, goals, reinforcement and feedback have a predictable effect on behavior irrespective of how it is assessed (i.e., via self-monitoring or external observation).

Timing. Reactivity also seems to be a function of the temporal order of self-recording in relation to the occurrence of the target behavior. In general, self-monitoring a behavior prior to its occurrence promotes greater reactivity than recording it after the fact (Kanfer, 1970a). This relationship was demonstrated by Bellack, Rozensky and Schwartz (1974) when subjects recording their food intake before eating achieved greater weight loss than those monitoring consumption after eating. Similar findings were reported by Gottman and McFall (1972) who found self-monitoring urges to make comments in class more reactive than recording actual comments. The pattern, however, did not hold in a study with young children. Nelson et al. (1977) found that self-monitoring prior to or after verbalizations did not differentially influence the reactivity of appropriate or inappropraite classroom verbalizations.

Concurrent Monitoring of Multiple Behaviors. Two studies have examined the effects of concurrently self-monitoring multiple behaviors. Hayes and Cavior (1977) trained subjects to self-record the frequency of face touching, nonfluencies, and value judgments, all considered negatively valenced behaviors. Groups of subjects monitored either one, two or three of the target behaviors. Tracking of a single behavior produced the greatest reactivity as measured by ratio change scores. Similar results were obtained when this study was replicated with positively valenced behaviors (Hayes and Cavior, 1980). In this instance, self-monitoring produced increases in looking, present tense

verbs, and feelings statements only when one behavior was recorded. Although decreased reactivity with multiple behaviors may enhance the utility of self-monitoring for assessment purposes, these benefits are offset by the finding that accuracy is also diminished.

Schedule of Self-Monitoring. There is some evidence to indicate that the frequency or schedule of self-monitoring may contribute to reactivity. Fredrikson et al. (1975) suggest that the optimal schedule for self-monitoring, when reactivity is desirable, is related to the rate and variability of the behavior. High rate behaviors may require immediate and continuous recording in order to facilitate response discrimination, while low frequency behaviors may still be reactive when monitored on an intermittent basis. This pattern of results was apparent in a study by Mahoney et al. (1973) assessing performance on a standardized general aptitude test. Using a teaching machine, subjects who self-monitored their accuracy on a continuous schedule improved their performance relative to those who monitored accuracy on every third problem. Similarly, Frederiksen et al. (1975) found continuous monitoring of each cigarette smoked to be more effective in reducing smoking behavior than nightly or weekly recording of the total number of cigarettes smoked for that time period.

Nature of the Self-Recording Device. A final variable implicated in the reactive process is the type of self-recording device used. According to some authors, obtrusive recording devices may function as a discriminative stimulus for the self-recording response

which may cue the individual to the environmental contingencies controlling behavior (Rachlin, 1974; Nelson and Hayes, 1981). Suggestive evidence for this comes from a series of case studies by Maletzky (1974) in which a wrist counter was used to self-record maladaptive behaviors. Reactive effects were demonstrated for all subjects during self-monitoring; however, in each case, when wrist counters were removed and self-recording terminated, response increases were noted. This phenomenon was also reported by Broden et al. (1971) when the removal of slips of paper used for self-recording study behavior resulted in a drop in study rates. The obtrusiveness of the recording device was experimentally manipulated by Nelson et al. (1978) with retarded adolescents self-recording appropriate verbalizations. Hand-held counters were found to be more reactive than belt-worn counters, although differences were not statistically significant.

To summarize, equivocal evidence for the reactive effects of self-monitoring has not been treated as whimsical. Systematic evaluation of the self-monitoring process indicates that reactivity may be a function of the specific conditions present during self-assessment. Among the variables affecting reactivity are: 1) client motivation, 2) the valence of the recorded response, 3) experimenter instruction, 4) the nature of the target behavior, 5) goals, reinforcement and feedback, 6) the timing of self-recording, 7) concurrent monitoring of multiple behaviors, 8) the schedule of self-monitoring, and 9) the nature of the self-recording device. Although empirical support for these variables exists, it should be

noted that, with the exception of valence, reinforcement and feedback, evidence is not strong. In most cases, findings should be considered tentative due to a limited number of studies, conflicting results, and inadequate methodology. Moreover, few investigators have examined these variables with children, thus further limiting the application of these findings to school psychology. Needless to say, future research should target this population as well as the psychological and educational problems they encounter. Finally, the search for the determinants of reactivity should be extended to other potential factors. For example: 1) demographic variables including age, sex, and intellectual functioning, 2) the parameters of training in self-monitoring, 3) behaviors versus behavior products (Haynes and Wilson, 1979), 4) the salience of the contingencies linked to performance, 5) interaction of reactive effects with other treatment strategies, 6) setting factors, and 7) cognitive factors.

Validity of Self-Monitoring

As indicated earlier, self-monitoring can serve either a therapeutic or assessment function in the context of behavioral intervention. When used for assessment purposes, the validity of self-monitored data becomes a primary concern. Validity is determined by the degree of association between self-monitored data and concurrent measures of the same target behavior using other methods (Haynes, 1978; Cone, 1981). This type of validity is known as criterion-related validity and refers to the accuracy of measuring a specific behavior compared to some standard or criterion. For example, if interobserver

agreement between a teacher and her aide is .90 on some measure of social withdrawal, data from either source can serve as the criterion measure. Validity is assessed by the degree to which the child's self-monitored data correspond to concurrent measures of the criterion (i.e., interobserver agreement between the child and the teacher or aide).

As Nelson (1977a, 1977b) indicated, the accuracy or criterion related validity of self-monitored data can be evaluated by calculating agreement between: 1) self-monitored data and data obtained concurrently by two or more external observers; 2) self-monitored data and data stimultaneously recorded by mechanical devices; and, 3) self-monitored data and some product-of-behavior measure directly related to the self-monitored behavior. Using these methods, several studies have directly examined the accuracy of self-recorded data. In reviewing this literature, Ciminero et al. (1977), Nelson (1977b), and Haynes and Wilson (1979) found estimates of accuracy to vary widely across studies and target behaviors. In studies with adult subjects, reports ranged from no agreement (.11, Hayes and Cavior, 1977) to excellent accuracy or agreement (.98, Azrin and Powell, 1969). Similar fluctuation is evident in the child literature. At the low end of the continuum were agreement measures of .42 between student and teacher ratings of disruptive behavior (Drabman et al., 1973), and .50 for boys' self-ratings versus teacher ratings of room cleanliness (Fixsen et al., 1974). In contrast, several investigators report good to excellent accuracy (interobsever agreement) when children self-

monitored their behaviors. For example, Bornstein et al. (1977) obtained 92 percent agreement between a child's self-monitored data on out-of-seat behavior and data collected by school personnel. Similar accuracy rates (.96) were reported by Miller and Kratochwill (1979) corresponding to agreement between child records of stomach complaints and those kept independently by the child's mother, teacher, and school nurse. Moreover, agreement between the nurse and teacher was. 98 in the school, providing a better estimate of the "true" accuracy of the data.

Variables Affecting the Validity of Self-Monitored Data

Highly variable estimates of validity evident in the studies above suggest that specific factors operate to affect the accuracy of self-monitored data. Recently, researchers have begun to analyze various conditions of self-monitoring to determine which variables influence accuracy. From a clinical standpoint, this knowledge can be very useful in designing self-monitoring procedures which enhance the data's validity. Accurate self-assessment can obviate the need for costly external observers and/or excessive demands on teacher or parent time. Furthermore, a greater number of children and school personnel can be served when clients participate in assessing their own behavior. Below is a listing of variables which have been empirically demonstrated to affect the accuracy of self-monitored data (McFall, 1977; Nelson, 1977a, 1977b; Haynes and Wilson, 1979). Since much of this literature involves adult subjects, appropriate caution should be exercised in generalizing these findings to children. Clearly needed

are both replication studies with child populations and exploratory research into the variables endogenous to children which influence accurate self-recording.

Training. In general, training in self-monitoring has been found to improve subjects' accuracy during self-monitoring. This is not surprising since training in the correct use of any behavioral assessment procedure should increase the validity of the data (Haynes, 1978). For example, Hamilton and Bornstein (1977) found that subjects who received training in self-monitoring were more accurate in recording target behaviors than those not provided specific training. Similar results were reported by Nelson et al. (1978) and Jones et al. (1977). In both cases, increases in accuracy paralleled exposure to training.

Reinforcement Contingencies. Administration of reinforcers contingent on accurate self-monitoring has also been shown to enhance accurate data collection. With children, Fixen et al. (1972) increased agreement between peer and self-reports of room cleanliness from .76 to .86 when subjects were reinforced for accurate self-ratings. For Layne, et al. (1976) the margin of improvement was even greater, increasing agreement from. 67-.75 to .90-.95 following the introduction of contingent rewards for accurate self-recording of cleanliness. Attempts to thin the schedule of accuracy checks and reinforcement for accuracy have met with mixed results. Layne et al. (1973) and Turkewitz et al. (1975) reported a decline in accuracy measures when

both the assessment and reinforcement of accuracy were gradually decreased.

Nature of the Target Behavior. Some target behaviors lend themselves to easier and more accurate self-recording than do others. Peterson et al. (1975) demonstrated, for example, that accuracies were higher for self-recorded face touches (.64) than for verbal non-fluencies (e.g., "You know", .00; "and all that", .31). Similar findings by Cavior and Marabatto (1976) and Hayes and Cavior (1977) indicated that overt motor behavior may be more accurately self-monitored than verbal behaviors, although considerable variability is reported across verbal response classes depending on the degree to which the verbal behavior is habitual. In addition, the values of the target behavior appear to influence the accuracy of self-monitored data. Nelson, Hay and Koslow-Green (1977) found that children self-recorded appropriate classroom verbalizations more accurately (.81) than their inappropriate verbalizations (.57). Similarly, higher accuracy rates were reported when teachers self-recorded their positive comments than when they monitored their negative comments (Nelson et al., 1977). However, Litrownik and Freitas (1980) did not replicate these findings, possibly because of the presence of external observers during self-monitoring.

Schedule of Self-Monitoring. There is some evidence to suggest that the contiguity between the occurrence of the target behavior and the self-recording response may affect the data's accuracy. Frederiksen et al. (1975) found that subjects were more accurate when

they recorded each cigarette immediately after they smoked it (.93) than waiting until the end of each day (.85) or at the end of each week (.87). This discrepancy would be expected to increase with target behaviors occurring at higher frequencies and/or having no product-of-behavior correlate (i.e., cigarette butts or the number of cigarettes remaining in a pack).

Response Competition. When the number or complexity of responses required of an individual are high, the accuracy of self-monitoring can be expected to decrease. Evidence for this comes from a study by Epstein, Webster and Miller (1975) in which self-recording accuracy diminished when subjects were engaged in a lever pressing operant task while concurrently self-monitoring their respiration rates. These findings suggest that self-monitoring is more accurate when clients are not involved in activities that may interfere with their self-observation and self-recording.

Knowledge of Accuracy Assessment. Several studies have demonstrated that self-monitoring accuracy improves when individuals are informed that their accuracy is being assessed. For example, Lipinski and Nelson (1974) achieved marked increases in accuracy of self-recorded face touches when subjects were told that periodic random accuracy checks would be made. Covert assessments of agreement averaged .52 compared to .86 when subjects were informed that independent measures would also be collected. Similar increases were reported by Nelson et al. (1975) and Lipinski et al. (1975) using face

touches as the target response. Targetting a more clinically relevant behavior, Santogrossi (1974) found the correspondence between pupil's self-recorded reading performance and those obtained concurrently by external observers to increase when either a teacher or peer co-monitored pupil reading performance.

In addition to the variables indicated above, there are a number of other factors that may be functionally related to the accuracy of self-monitoring. McFall (1977) has identified some of these potential factors which await empirical verification. The first of these, client characteristics, is of particular importance to school psychology. Age and type of functional disorder should influence the complexity of the target behavior which can be accurately self-recorded as well as the amount and type of training needed to obtain high subject-rater agreement. Second, the degree of response effort required in self-monitoring will likely affect client performance. As with other observational procedures, self-monitoring accuracy can be expected to decrease as the recording response becomes more arduous and time consuming (Hersen and Bellack, 1981). A third factor, related to sound scientific practice, is the use of systematic recording methods to enhance the accuracy and reliability of the data. In this regard, the use of recording forms, operationalized behaviors, and timing and recording devices may be effective. Fourth, the type of recording device used in self-monitoring may affect accuracy. Along the dimension of obtrusiveness, Nelson et al. (1978) examined the effects of hand-held versus belt-worn counters on the accuracy of self-

monitored data with retarded clients. Although no significant between-group differences were found, accuracy rates were higher for subjects using hand-held counters. A fifth variable not addressed thus far in the research literature is the effect of bias and expectancies on self-monitored data accuracy (Haynes, 1978). For example, a child interested in obtaining positive teacher comments or contingent rewards may overrepresent his or her time on-task because of the contingencies operating on that behavior. Similarly, parents may "exaggerate" the frequency of noncompliant responses on the assumption that a child with more severe problems will have a better chance of being accepted into a treatment program. Thus, the psychologist should always bear in mind the potential for bias in self-monitored data when corroborative data are not available.

Summary and Implications

The role of self-monitoring in the assessment and treatment of childhood disorders is expanding rapidly. In the area of assessment, SM may serve the function of obtaining global information at the preintervention stage, providing data on private and unobservable events, or as the primary dependent measure in the evaluation of an intervention. Data systems parallel those used widely in behavioral assessment and include direct methods such as narrations, frequency counts, duration measures and time sampling, as well as self-ratings, behavioral traces and archival records considered to be indirect measures of child performance. The therapeutic value of SM is associated with its tendency to be reactive under certain conditions.

Factors implicated in the reactive process are the motivation of the client to change, the valence of the self-monitored behavior, the type of experimental instructions provided, the use of goals, reinforcement and feedback, concurrent monitoring of multiple behaviors, the schedule of self-monitoring and the nature of the self-recording device. Yet to emerge in the self-monitoring literature, however, is a data based theory accounting for the reactive effects of self-monitoring. A major purpose of the following investigation is to examine the existing theoretical accounts of the reactivity of SM under laboratory conditions. The results are expected to help clarify conflicts in the literature and guide prospective research efforts in the area.

CHAPTER 3

METHOD

Participants

Undergraduate students enrolled in introductory psychology courses volunteered to participate in the study. No payment or course credit was provided to students in exchange for their involvement in the experiment.

Students ranged in age from 17.3 to 53.6 years-old (M=20.0; SD=7.5). A variance estimate of the dependent measure under preself-monitoring conditions was obtained on 10 graduate students. This statistic was used to estimate the sample size necessary to achieve a power coefficient of .80 for a one-tailed test with alpha .05 and a critical discrepancy (D) of 2.0 (Blommers and Forsyth, 1977). On the basis of this analysis, 60 subjects (13 males and 47 females) participated in the experiment.

The author served as the experimenter during the study. A female with an MA in social science and a male Ph.D. psychologist, uninformed of the hypotheses of the study, acted as data collectors.

Experimental Design

A pretest-posttest control group design (Campbell and Stanley, 1963) was used to evaluate between-group and within-group effects.

Subjects were randomly assigned to one of four experimental conditions or a training only group. The experimental conditions employed were 1) self-monitoring; 2) self-monitoring and goal setting; 3) self-monitoring, goal setting and self-reinforcement; and, 4) goal setting and self-reinforcement. Separate lists were used for females and males in assigning subjects to treatment conditions to ensure that sexes would be equally distributed across groups. Each of the four experimental groups represented conditions consistent with the various theories of reactivity: group (1) Rachlin; group (2) Kanfer; group (3) Nelson and Hayes; and group (4) Kanfer.

Dependent Variable

Criteria for the selection of the dependent variable were that it occur at a moderate to high rate under natural conditions, have a positive or negative valence for most subjects and have been demonstrated to be reactive when self-monitored (e.g., Hayes and Cavior, 1977). Under these guidelines, verbal nonfluencies was chosen as the target measure. A verbal nonfluency was defined as any occurrence of the sounds "um", "ah", "er", "uh", or "you know" (e.g., "Um, he always, you know, does that."). Functional usage of these sounds was not scored (e.g., "Did you know that?").

All experimental sessions were audiotaped using a small portable tape recorder. One data collector listened to tapings on all subjects in a random order. Standard recording forms containing the operational definition of the dependent measure and interval columns were used. Audio tapes were divided into 10-second intervals. The

data collector rated each interval for the occurrence (frequency) or nonoccurrence of the target measure. In order to account for variability across subjects in the amount of time speaking, intervals in which no speaking occurred were excluded from analysis. This procedure permitted the calculation of the rate of verbal nonfluencies and consisted of the ratio of the number of nonfluencies scored to the time each subject spoke. This measure was used in subsequent data analysis.

Independent assessments were conducted by the second data collector on a random sample of 20 subjects. Data from both data collectors on these 20 subjects were used to calculate interobserver agreement. Cohen's (1960) Kappa was used to account for chance agreement between assessors.

In addition to audio-taping each experimental session, the experimenter collected data on the dependent variable for all groups during the preself-monitoring phase. These data were used to inform subjects involved in goal setting of their preself-monitoring frequency of nonfluencies. Data were taken in an unobtrusive manner to minimize the liklihood that subjects would identify the target measure on the basis of the experimenter's data collection behavior. The experimenter recorded each occurrence of the target behavior with a hand-held noiseless counter. The counter was held underneath the table top and out of the subjects' view. Data obtained from the primary data collector on the frequency of verbal nonfluencies for subjects in groups 2, 3 and 4 were used for purposes of establishing interobserver

agreement. A percentage agreement measure was generated using the frequency measures obtained by dividing the smaller of these two values by the larger.

Both assessors were trained in the use of the data collection procedures. Training consisted of 1) verbal instructions describing the method of data collection; 2) the use of standardized recording forms containing operational definitions of the dependent measure; 3) modeling the data collection procedure on a 5-minute sample tape; and 4) independent trials with 5-minute sample tapes. Training was completed when assessors obtained three consecutive agreement measures of .80 or higher.

Procedures

The study was divided into four experimental phases: preself-monitoring, training in self-monitoring, self-monitoring and cognitive assessment. All four phases were conducted in the same experimental room and during a single session for each subject. The room measured 12 by 14 feet and contained a small rectangular table. Two chairs were positioned opposite each other across the table. A portable tape recorder was placed on the experimenter's side of the table.

All subjects were individually exposed to the preself-monitoring phase. A single subject and the experimenter were seated at the table facing each other. The experimenter read a set of introductory instructions (see Appendix A). Immediately after the instructions were read, the experimenter turned the tape recorder on, set the kitchen timer for 3 minutes, and read the first question. Half

of the subjects in each group received question series A (see Appendix A) while the remaining half received question series B (see Appendix B). If, during the 3-minute period alloted the subject to answer, the subject did not speak for approximately 8-10 seconds, the experimenter read one of the following three statements to prompt the subject to resume speaking: 1) "Tell me more about that.", 2) "Tell me why you think that.", 3) "What else can you say about it?". At the end of the 3-minute period, the experimenter reset the timer and read the second question (see Appendices A and B). Experimenter prompts were again provided after 10-second pauses. The above procedure was repeated for a third question which completed the preself-monitoring phase (see Appendices A and B).

Immediately following the preself-monitoring condition, all subjects received training in self-monitoring. Training consisted of a 5-step procedure developed by Mahoney (1977). First, the target behavior verbal nonfluencies was identified by the experimenter. The behavior was operationally defined to facilitate reliable self-observation and self-recording. In the second step, explicit instructions on how to self-monitor were provided. Third, the experimenter modeled the self-monitoring response using a pencil and sample recording form. The fourth step involved having the subject repeat the target definitions and self-monitoring instructions. The experimenter provided subjects corrective feedback where necessary and asked the subject to repeat the target definition and/or the self-monitoring instruction, where necessary. Fifth, each subject

self-monitored the target behavior over two, 2-minute trials. During each 2-minute trial, subjects were asked to talk about school-related topics of their choice. Mahoney's (1977) 5-step procedure was followed as closely as possible for each subject, although the pace of instruction varied depending on individual questions and learning rates. Since accurate self-monitoring has not been shown to be prerequisite for reactivity to occur (Nelson and Hayes, 1981), no criterion for accuracy was used to complete training.

Upon completion of the self-monitoring training, subjects self-monitored the target behaviors under one of four experimental conditions. All subjects received verbal introductory instructions (see Appendix B). Subjects in the four experimental groups received a written operational definition of the target behavior taped to the table in front of the subject.

Subjects in the self-monitoring group (1) were given a mechanical counter with a visual display. Counters provided subjects with a cumulative total of their monitored responses and served as a form of continuous performance feedback. The experimenter gave the following self-monitoring instructions: "Now let's try something a little different. Place the counter in your hand with the number display facing you. Look at the definition taped on the table in front of you. Each time you say a nonfluency, click the counter and look at your total. Now here is your first question. Remember to speak for the entire 9 minutes." For all experimental groups, subjects were provided three consecutive questions (i.e., the question series, A or

B, they had not yet received) and alloted 3 minutes for each answer (see Appendices A and B). Experimenter prompts were used in a manner identical to the preself-monitoring condition.

Members of the self-monitoring and goal setting group (2) received a mechanical counter with visual display identical to group 1. The self-monitoring instructions provided were identical to group 1 with the following addition: "When you discussed the first three questions you said ___ nonfluencies (the number collected by the experimenter during preself-monitoring). Before we start, set your own goal for reducing the nonfluencies you say during the next 9 minutes. Tell me what your goal is and write it down on the card in front of you. Now here is your first question. Remember to speak for the entire 9 minutes."

Conditions for the self-monitoring, goal setting and self-reinforcement group (3) involved the use of a hand-held counter with a visual display and self-monitoring instructions identical to group 2 with the following addition: "At the end of the 9 minutes, compare your goal to the number of nonfluencies you recorded. If you reached or exceeded your goal, say out loud 'I did a really good job of showing self-control'. Now here is your first question. Remember to speak for the entire 9 minutes." At the end of the 9 minutes, subjects who reached or exceeded their goal were prompted to make the self-reinforcement statement. The experimenter said "You reached your goal. Now say out loud 'I did a really good job of showing self-control'. Now here is your first question. Remember to speak for the entire 9 minutes."

Subjects in the goal setting and self-reinforcement group (4) were not provided mechanical counters and received goal setting instructions identical to those given to group 2 along with the following self-reinforcement instructions: "At the end of the 9 minutes, compare your goal to the number of nonfluencies you think you said. If you reached or exceeded your goal, say out loud 'I did a really good job of showing self-control'. Now here is your first question. Remember to speak for the entire 9 minutes." Experimenter prompts were provided to all subjects in this group to make self-reinforcement statements for goal attainment.

Participants in the training only (5) were provided instructions identical to the preself-monitoring phase and given three different questions (see Appendices A and B). No self-monitoring procedures were employed.

Immediately following the self-monitoring phase, all subjects involved in the experiment underwent the cognitive assessment phase. The purpose of this procedure was to obtain a measure of subjects' self-reported cognitions while they were engaged in self-monitoring. This phase simply consisted of the experimenter saying "Tell me anything that you were thinking at the times you self-monitored (or noticed yourself say) a nonfluency". Subject responses to this instruction were recorded by the experimenter. After all subjects had participated in the experiment, subjects were invited to a debriefing meeting in which the experimenter discussed the purposes of the study, its hypotheses and tenative findings (Kazdin, 1980).

CHAPTER 4

RESULTS

Coefficients of interobserver agreement were calculated between the primary and secondary observers' scores on the dependent variable. Kappa values computed on pretest and posttest scores for each of the 20 subjects randomly selected for interobserver agreement yielded a pretest mean of .6208 (range from -.2832 to .9069), a posttest mean of .7898 (range from .3768 to .9208), and a grand mean of .7324. These values exceed the recommended criterion of .60 and can be considered reliable measures of subjects' verbal nonfluencies (Gelfand and Hartmann, 1976; Hartmann, 1977). Pretest nonfluency data collected by the experimenter and provided to subjects in groups 2, 3, and 4 for purposes of goal setting were compared with the primary observer's scores. Dividing the smaller frequency by the larger frequency for each subject resulted in a mean reliability of .9046 (range from .6667 to 1.000).

A test for the reactive effects of the various self-monitoring conditions within groups employed a repeated measures analysis of variance (ANOVA). The results presented in Table 1 indicate a significant trial effect across the pre and posttest ($F=133.766$, $df=1, 55$, $p<.01$). No significant effects were found for the group factor

Table 1. Repeated Measures ANOVA

Source	SS	df	MS	F
Between subjects	389.955	59		
A (groups)	11.026	4	2.757	.400
Subject w. groups	378.929	55	6.890	
Within subjects	430.707	60		
B (pre/posttest)	290.406	1	290.406	133.766*
AB	20.909	4	5.227	2.408
B x subj w. groups	119.392	55	2.171	
TOTAL	820.663	119		

* - .01 significance level

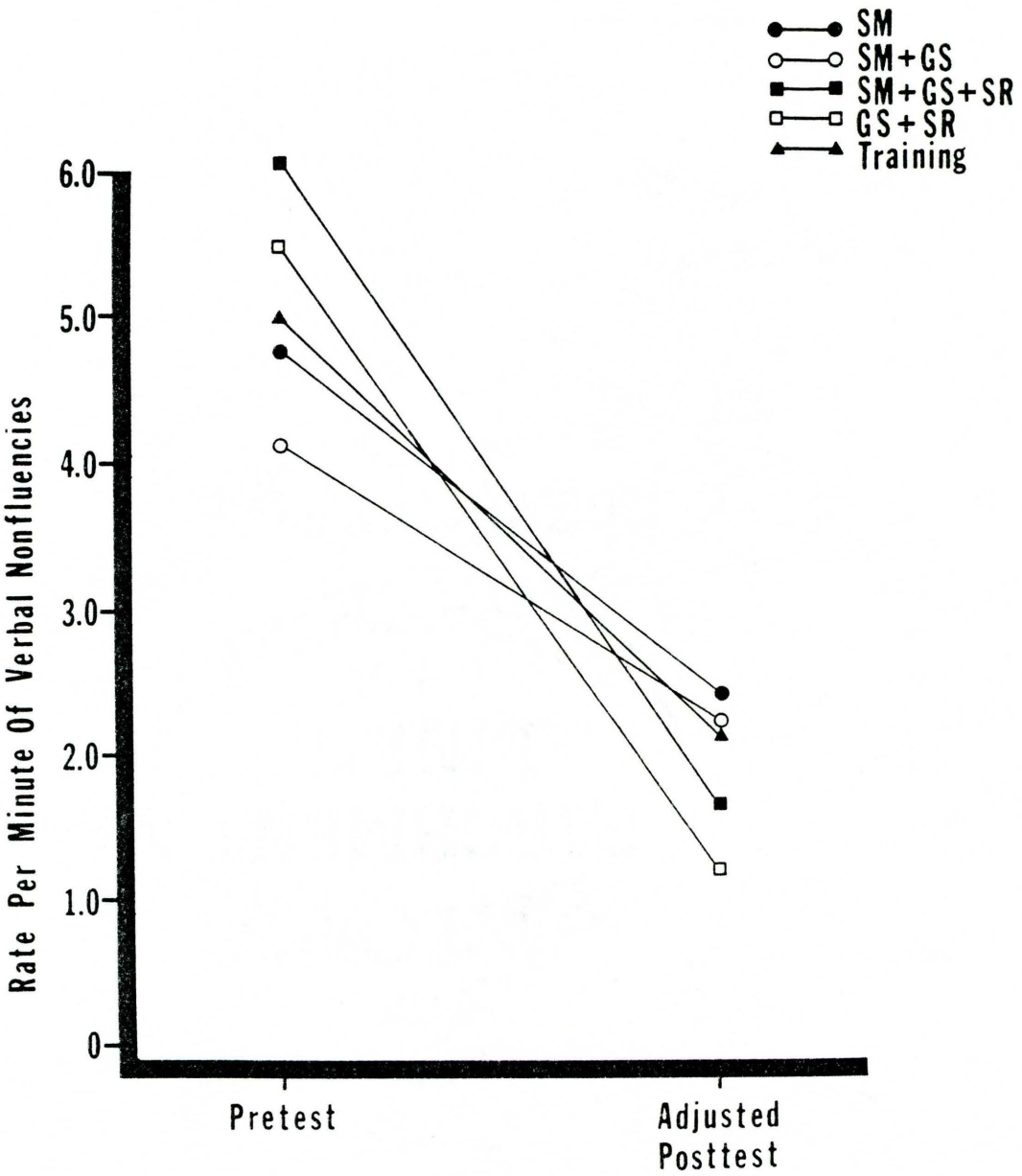

Figure 1. Mean rate per minute of verbal nonfluencies by experimental group on the pretest and the adjusted posttest (pretest as covariate)

($F=.400$, $df=4, 55$, $p > .05$) or the groups by trials interaction (AB) ($F=2.408$, $df=4, 55$, $p > .05$). Post-hoc analyses of the within group factor showed a significant decrease in verbal nonfluencies for all groups. Figure 1 illustrates the reduction in nonfluencies achieved by subjects within the various self-monitoring groups. Each condition exceeded the Tukey critical value ($q=3.85$, $df=2, 55$). Obtained q-values for the pretest-posttest contrasts were as follows: Group 3 (SM + GS + SR, $q=9.706$), group 4 (GS + SR, $q=9.537$), group 5 (TRAINING ONLY, $q=6.744$), group 1, (SM, $q=5.542$), and group 2 (SM + GS, $q=5.046$).

An analysis of covariance (ANCOVA) was conducted on the posttest scores using the pretest scores as the covariate. This analysis was employed due to disparate pretest group means and considerable within group variability (see Table 2) despite group formation via randomization procedures. An analysis of variance (ANOVA) conducted on the pretest scores resulted in no significant between group differences ($F=.8535$, $df=4, 55$). However, this outcome was due largely to a sizeable MS error term (7.5621) overshadowing apparent group differences on the pretest. Indeed, an analysis of the power of this ANOVA indicated a power coefficient of less than .20 for a critical value of .3422 (Kirk, 1968).

Table 3 contains the results of the ANCOVA with the pretest as covariate. Significant between group differences were obtained ($F=2.9086$, $df=4, 54$, $p < .05$). An omega square computed on these data indicated that 19.51% of the total variance is accounted for by the group factor self-monitoring. The power of the ANCOVA F-test was

Table 2. Group Pretest, Posttest and Adjusted Posttest Means and Standard Deviations

Group	(#)	Pretest $\bar{X}$ SD	Posttest $\bar{X}$ SD	Adjusted Posttest $\bar{X}$	Difference Between Pretest $\bar{X}$ and Adjusted Posttest $\bar{X}$
SM	(1)	4.760 2.464	2.403 1.524	2.491	2.269
SM + GS	(2)	4.117 2.475	2.403 1.627	2.237	1.880
SM + GS + SR	(3)	6.057 2.501	1.929 .799	1.655	4.402
GS + SR	(4)	5.473 3.975	1.417 .747	1.206	4.167
SM TRAINING	(5)	4.964 1.885	2.096 1.114	2.127	2.837
Grand Mean		5.074	1.963	1.963	3.111

Table 3. Analysis of Covariance (Pretest Covariate)

Source	SS	df	MS	F	Omega Square
Between groups	10.53844	4	2.635	2.909*	.195
Within groups	48.91383	54	.906		
TOTAL	59.45227	58			

* - .05 significance level

calculated and found to be .67 for a critical value of 1.236. Stated alternatively, there was a 67% chance of detecting actual group differences of 1.236 or greater. Adjusted pretest means were derived from the ANCOVA and are presented in Table 2. This adjustment in the subjects' posttest scores when pretest score variance is accounted for resulted in group posttest-adjusted posttest differences of -.088, -.266, .274, .111 and -.031 for groups 1 through 5, respectively. This indicates that moderate adjustments were made for groups 2, 3 and 4 which correspond to their departure from the grand pretest mean (see Table 2).

Comparisons among group means were conducted via one-tailed planned orthogonal contrasts and two-tailed planned nonorthogonal contrasts. One-tailed planned orthogonal contrasts were applied to the following contrasts: group 2 v. group 3 and group 1 v. group 5. These comparisons were of primary interest because they would expose the relative contributions of self-reinforcement and self-monitoring. Thus, the a priori prediction was that nonfluencies for group 3 would be lower than group 2 and group 1 would be lower than group 5. The remaining eight contrasts of interest were performed using a two-tailed planned nonorthogonal contrasts analysis. To maintain the alpha rate constant for the family of comparisons, Duncan's Multiple Range Test was employed for these nonorthogonal contrasts (Kirk, 1968).

Table 4 presents the differences among adjusted posttest means, the significance test used and the level of statistical significance obtained. Group 4 (GS + SR) significantly reduced their nonfluencies

when compared to group 1 (SM, D=1.845, $p < .01$), group 2 (SM + GS, D=.9315, $p < .05$), and group 5 (TRAINING ONLY, D=.8207, p=.05). Marginally significant contrasts also favored group 3 (SM + GS + SR) over group 1 (SM, D=.8351, $p < .06$), and at higher significance levels for group 2 (SM + GS, D=.5821, p=.08). The remaining contrasts including group 1 (SM) with group 5 (TRAINING ONLY), group 1 (SM) with group 2 (SM + GS), group 2 (SM + GS) with group 5 (TRAINING ONLY), group 3 (SM + GS + SR) with group 4 (GS + SR) and group 3 (SM + GS + SR) with group 5 (TRAINING ONLY) were nonsignificant to the .25 level. A summary of the contrasts performed, the test statistic obtained and the level of statistical significance achieved appears in Table 5.

Data pertaining to goal setting were examined for groups 2, 3 and 4. A percentage was computed for each subject consisting of their individually set goal for reducing nonfluencies below their pretest score divided by their pretest score and multiplied by 100. This measure represented the percent reduction of nonfluencies reflected in the individual's goal. A one-way ANOVA found no significant differences among groups with respect to their goal setting (F=1.651, df=2, 33, $p > .05$). The means for groups 2, 3 and 4 were 44.96%, 55.1% and 46.41%, respectively. A second analysis was conducted examining the extent to which individually set goals were achieved on the posttest. Each subject's goal was divided by their posttest frequency and multiplied by 100. A one-way ANOVA performed on these data again revealed no between group differences (F=1.495, df=2, 33). Group means were 125.45% (group 2), 205.66% (group 3), and 198.85% (group 4).

Table 4. Differences Among Adjusted Posttest Means

		$\bar{X}_4$ (GS + SR)	$\bar{X}_3$ (SM + GS + SR)	$\bar{X}_5$ (SM TRAINING)	$\bar{X}_2$ (SM + GS)	$\bar{X}_1$ (SM)
(GS + SR)	$\bar{X}_4$ 1.306	____	.349	.821+	.932+	1.185++
(SM + GS + SR)	$\bar{X}_3$ 1.655		____	.471	.582**	.835*
(SM TRAINING)	$\bar{X}_5$ 2.127			____	.111	.364
(SM + GS)	$\bar{X}_2$ 2.238				____	.253
(SM)	$\bar{X}_1$ 2.491					

* - two-tailed planned orthogonal contrast, .06 significance level

** - one-tailed planned orthogonal contrast, .08 significance level

\+ - two-tailed planned nonorthogonal contrast, .05 significance level

++ - two-tailed planned nonorthogonal contrast, .01 significance level

Table 5. Statistical Tests and Significance Levels for Between Group Comparisons

Comparison	Statistical Test	Significance Level
SM TRAINING vs. SM	t = .936	N.S.
SM TRAINING vs. SM + GS	W = .777	N.S.
SM TRAINING vs. SM + GS + SR	W = .777	N.S.
SM TRAINING vs. GS + SR	W = .819	P < .05
SM vs. SM + GS	W = .777	N.S.
SM vs. SM + GS + SR	W = .830	P < .06
SM vs. GS + SR	W = 1.132	P < .01
SM + GS vs. SM + GS + SR	t = 1.459	P = .08
GS + SR vs. SM + GS + SR	W = .819	N.S.

t - Student's t-Distribution obtained value

W - The critical difference required for significance with the Duncan Multiple Range Test

However, only 40% of the subjects in group 2 (SM + GS) reached their goals, while 75% of the subjects in group 3 (SM + GS + SR) and 91% of the subjects in group 4 (GS + SR) achieved their goals for reducing the target measure. A chi square analysis of these percentages indicated that these groups differed significantly ($X^2=7.9134$, df=2, $p<.05$).

Subject responses to the instruction "Tell me anything you were thinking at the times you self-monitored (or noticed yourself saying) nonfluencies" were rated by independent judges as "self-punishing" or "neutral" for all 60 subjects. None of the subjects' self-reported cognitions were judged to be "self-reinforcing". The extent of interjudge agreement was found to be .90 by dividing the number of agreements by the number of agreements plus disagreements on a random sample of 20 subjects. A Chi Square analysis consisting of a group by statement classification contingency table yielded nonsignificant results ($X^2=4.902$, df=4, $p>.05$). Self-punishing statements were reported by 0%, 16.7%, 33.3%, 25.0%, and 16.7% for groups 1 through 5, respectively.

CHAPTER 5

DISCUSSION

The results of the present study indicated that all of the experimental conditions examined including self-monitoring only, self-monitoring plus goal setting, self-monitoring plus goal setting and self-reinforcement, goal setting plus self-reinforcement, and training in self-monitoring produced significant reductions in verbal nonfluencies. Reactive effects were largest under the two conditions (SM + GS + SR and GS + SR) that employed self-reinforcement procedures. In the present investigation, however, the contribution of self-reinforcement to the reactivity of self-monitoring appeared to be the result of external contingencies rather than as a function of covert processes. These findings contribute to an expanding body of research aimed at delineating a theoretical model to account for the reactive effects of self-monitoring (Nelson and Hayes, 1981; Rachlin, 1975; Spates and Kanfer, 1977).

The finding that all of the experimental conditions yielded reactive effects is of interest. Performance in the group receiving only training in self-monitoring (group 5) was comparable to groups 1 and 2 who, in addition to training, self-monitored (group 1) and set personal goals while self-monitoring (group 2) during the posttest.

This suggests that the training process consisting of learning an operational definition of the target response, learning to self-record the target behaviors and practice in self-monitoring contribute significantly to the reactive process. A major function of the training experience may be to teach the individual to discriminate the occurrence of a given response (cf. Nelson and Hayes, 1981). Increasing awareness of each emission of a behavior may serve, as Rachlin (1975) and Nelson and Hayes (1981) have suggested, to enhance the relationship between the behavior and its controlling contingencies. In this case, the negatively valenced behavior verbal nonfluencies is likely under the control of social variables in the individuals' reinforcement or punishment history. Since, under laboratory conditions, the inclusion of self-monitoring or self-monitoring plus goal setting resulted in no further reduction in the dependent variable, these two factors did not appear to contribute to the salience between the behavior and the event which controls it.

Another important outcome of this study was that group 3 (self-monitoring, goal setting and self-reinforcement) outperformed groups 1 (self-monitoring) and 2 (self-monitoring and goal setting). While the evidence for group 4 differing between groups 1, 2 and 5 is strong (i.e., significance at or beyond .05), marginal significance was obtained for the group 3-group 1 and group 3-group 2 contrasts (i.e., .06 and .08, respectively). However, when groups 3 and 4 were compared, no significant relationship was found. Thus, the common denominator that differentiates groups 3 and 4 from groups 1, 2 and 5 is the inclusion of self-reinforcement.

At first glance, these findings would appear to support Bandura's (1976) position that self-reinforcement is a key variable in the reactive process. However, the results of the cognitive assessment phase and the analysis of goal data tend to undermine the validity of this tenet. For example, no significant between group differences were found in subjects' reports of self-punishing statements (none of the self-reported statements were judged to be self-reinforcing). Although self-punitive statements were more prevalent in groups 3 and 4 (33.3% and 25.0%, respectively), they occurred at low rates in all groups (group 1, 0%, group 2, 16.7% and group 5, 16.7%). According to the cognitive mediational view, the self-control or reactive process is largely automatic (Spates and Kanfer, 1977). That is, the act of self-recording initiates a covert process beginning with the evaluation of one's performance against some standard or goal. Favorable performance is presumed to result in self-reinforcement while falling short of the goal yields self-punishment. If this chain of events were automatic as Kanfer (1975) and Bandura (1976, 1977) contend, we would have expected to see a large percentage of subjects reporting self-punishment across all groups, a pattern not born out in the data. Examination of the goal data for groups 2, 3 and 4 also provides evidence contrary to the cognitive-mediational model. Only 40% of the subjects in group 2 (SM + GS) reached their individually set goals, while 75% of the subjects in group 3 (SM + GS + SR) and 91% of the subjects in group 4 (GS + SR) achieved their goals for reducing the target measure. An obvious explanation for subjects in groups 3 and 4

achieving a greater percentage of their goals is the presence of the self-reinforcement component. Conversely, in the absence of reinforcement the percentage of goal attainment in group 2 was relatively low. This finding again contradicts the view that covert forms of reinforcement automatically follow self-evaluation of goal attainment to control the level of responding, although conclusions drawn on the basis of self-report data should be appropriately qualified.

As noted above, the groups demonstrating the greatest reactivity received an experimentally manipulated form of "self-reinforcement." There is some question, however, to what extent this experimentally controlled form of reinforcement corresponds to the covert form of reinforcement referred to in the cognitive-mediational model. In this study, subjects exposed to the self-reinforcement variable were instructed prior to their posttest to say a positive statement regarding their performance immediately following the posttest contingent upon reaching their goal. Although this tactic is commonly used by cognitive-behavioral researchers to verify that "self-reinforcement" occurred (see Bandura, 1976, 1977), it supplants an external (and observable) event for a covert (and unobservable) event. The external event in this case is the experimenter's stated contingency and the act of commending oneself in front of the experimenter. Because superior performance was evident only in those groups receiving the external form of reinforcement, there appears to be little need to turn to covert variables to explain the observed differences in reactivity.

The results of this study have a number of implications regarding a theory of reactivity in self-monitoring and the use of this procedure in applied psychology. A major finding of this study is the importance of training in self-monitoring to the reactive process. As mentioned earlier, the primary function of training may be to teach response discrimination. Operant psychologists have referred to the phenomenon of response discrimination as self-awareness (Catania, 1975; Rachlin, 1978). That is, observations or discriminations of one's own behavior that correlate closely with the behavior itself accentuate the relationship between the response and its reinforcer (or punisher). The analogy Rachlin provides is that response discrimination "may be like saying to oneself, 'Yes, I did just do that', thereby performing the same function as the feedback click that tells the pigeon it has just pecked a key" (1976, p. 253). When this reliably occurs, the individual would seem to be in a better position to gauge the level of his/her performance with that needed to obtain reinforcement. Interestingly, in this study subjects provided with a mechanical counter for this purpose did no better than those without the recording device. While this finding may contradict those studies that suggest an obtrusive recording device enhances reactivity (e.g., Broden et al., 1971; Maletzky, 1974; Nelson et al., 1978), this effect is apparent only under laboratory conditions free of many of the distracting stimuli found in the natural environment. In applied settings, the use of an obtrusive recording device may well contribute to reactivity by providing a stimulus to both prompt discrimination and record its occurrence.

Another aspect of this study having theoretical implications is the effect of goal setting on the respondent's verbal nonfluencies. Of the three groups (2, 3 and 4) who were instructed to set personal performance goals, only those receiving contingent reinforcement (groups 3 and 4) showed significant improvement beyond the other experimental conditions. This suggests that goal setting contributes little, if any, to the reactive process in the absence of reinforcement. Moreover, as discussed above, it appears that external reinforcement is required to produce the effect since group 2 (self-monitoring and goal setting) performed no better than groups 1 (self-monitoring) and 5 (TRAINING ONLY). This concluson diverges sharply with the assertion by Spates and Kanfer that "criterion setting (goal setting) represents a critical stage in the model and...training which did not include this component was ineffective" (1977, p. 10, 13). Rather, the present study's finding that goal setting is most effective under conditions of reinforcement corrobates reports by Humphey et al. (1978), Turkewitz et al. (1975) and Santogrossi et al. (1973).

A third theoretical implication evident in this work is the importance of reinforcement to reactive self-monitoring. Among the theoretical views investigated in this study only Rachlin (1974) ascribed a paramount role to external reinforcement. In the absence of external consequences, self-monitoring and goal setting serve no function according to Rachlin (1978). Interestingly, Bandura (1976) concurs with the view that reinforcement, albeit self-administered, is

the critical component in the reactive chain. The major point of contention, however, is whether the source of reinforcement lies within or outside the skin. Evidence for a covert process that supports behavior was not found in this study. Only when reinforcement was available externally did reactivity increase significantly. In addition, subject reports of self-punishing thoughts did not differ significantly across groups and occurred at low rates. Together these findings suggest that self-administered covert consequences may not be assumed to occur automatically in response to self-monitoring and goal setting.

Fourth, support for some of the tenets of the multiple cueing model of Nelson and Hayes (1981) is also difficult to glean from the present study. The results obtained here indicate that training in self-monitoring and reinforcement are key factors in the reactivity of self-monitoring. If the principal function of training is to facilitate response discrimination, the use of an obtrusive recording device (i.e., a hand-held mechanical counter) apparently did not provide additional assistance to subjects in recognizing when they said nonfluencies. It is possible, however, that in the absence of training an obtrusive recording device would enhance reactivity. Further research to assess the separate and combined effects of training and obtrusive recording devices is needed to determine whether their effects are additive or overlapping. The contribution of goal setting as an additive component in reactivity is likewise unclear. The current data indicated that goal setting did little to facilitate

reactivity outside the context of external reinforcement. Under these circumstances, goal setting apparently does not function as a cue (cf. Nelson and Hayes, 1981). Since the present study did not examine reinforcement apart from goal setting, it was not determined whether setting goals interacted favorably with reinforcement. While additional research is needed to clarify this relationship, it is reasonable to assume that factors that increase the salience between behavior and its contingent consequence will enhance reactivity.

Finally, evidence for the importance of response discrimination and reinforcement has implications for Nelson and Hayes' theory of reactivity and the direction future research should take. Nelson and Hayes refer to each of the components in the reactivity equation as cues (e.g., therapist instructions, training, the self-recording device, feedback, self-evaluation, self-administered consequences and the self-recording response). That is, these stimuli are believed to contribute to reactivity by accentuating the relationship between behavior and its environmental consequences. The implication of the present study, however, is that two basic processes are operating, namely response discrimination and environmental consequences (according to reinforcement or punishment paradigms). The value of each of the factors listed above may therefore be in its effect on either of these two processes. For example, training in self-monitoring, obtrusive recording devices, the self-recording response and performance feedback can be viewed as factors which enhance response discrimination. Stated simply, these factors instruct the

individual on what and how to self-record, remind the individual to self-record and provide information on when and how much behavior has occurred. A compelling argument has been made that self-reinforcement (i.e., the delivery of rewards to oneself) serves a similar function (Catania, 1975). By contrast, experimenter instructions (particularly stated contingencies) and self-evaluation (a response to goal setting) may influence the effects of reinforcement or punishment. Stated contingencies generate what Skinner (1957) refers to as rule governed behavior and serve to improve the efficiency of the behavior changing paradigm and avoid adventitious reinforcement. Where goal setting is effective (e.g., Kazdin, 1974), this may also be the result of specifying the amount of behavior required to obtain reinforcement or avoid aversive consequences.

Given the preceding rationale, future research efforts should perhaps focus on variables known or suspected to influence these two basic processes. Regarding response discrimination, any stimuli or combination of stimuli that increase an individual's awareness that behavior has occurred may contribute to the reactivity of self-monitoring. Many of these factors have already received some attention in the empirical literature (Nelson, 1977a, 1977b). Others, however, await empirical scrutiny including 1) the factors that affect the accuracy of self-monitored data while controlling for the effects of reinforcement for the target behaviors, 2) the latency between the behavior and the self-recording response, 3) the effect of competing responses that interfere with response discrimination, and 4) factors

that affect the discriminative properties of performance feedback. The role of environmental consequences remains relatively unexplored. Factors to consider in this area include: 1) the relative impact of punishing versus reinforcing contingencies, 2) schedules of reinforcement, 3) the magnitude of reinforcement, 4) competing contingencies, and 5) the role of behavioral chaining in the reactivity of self-monitoring.

In addition to the theoretical issues cited above, the results of this study offer implications for practitioners using self-monitoring. When self-monitoring is used primarily as a therapeutic variable, steps may be taken to enhance its reactivity. First, clients should receive thorough training in self-monitoring as outlined by Mahoney (1977). Periodic retraining may also be advisable should therapeutic effects decline. Second, reinforcement contingencies should be in effect for the self-monitored behavior. Since the results of this study suggest that covert consequences are not automatic, care should be taken to plan external contingencies. Finally, these suggestions should be used in tandem with recommendations provided by Nelson (1977a) for maximizing the reactivity of self-monitoring. Briefly, these include: 1) providing individuals with specific performance goals, feedback and reinforcement for reaching their goals, 2) using a relatively obtrusive recording device, 3) selecting a single behavior for self-monitoring during a given time interval, 4) recording each instance of the target behavior, and 5) self-monitoring undesirable behaviors prior to rather than after their occurrence.

Generalizations made from the present study should take the following factors into account. First, this study was conducted in a laboratory setting. Extraneous sources of variability were minimized to highlight the effect of the independent variable. It remains to be seen whether the findings reported here will generalize to the natural environment. Second, the subject population used limits the generality of the results. For example, no conclusions may be made regarding school-age, clinical or disadvantaged populations. Third, the target behavior used in this study has generally little clinical relevance. Further research is needed to determine whether comparable effects are found with motor and other verbal behaviors both with and without clinical significance. Fourth, some of the conclusions drawn from this experiment were based on marginally significant results. Additional research is required to assess the validity of these findings. Finally, investigation in this area should be extended to actual clinical cases in various settings. Such efforts are needed to bridge the gap between basic and applied research.

APPENDIX A

Introductory Instructions for the Preself-Monitoring Phase and Question Series A

Instructions

During the next 9 minutes, I will ask you a series of three questions. After each question you will have 3 minutes to discuss your answer. Please speak for the entire 3 minutes. If, during a 3-minute period, you cannot think of anything to say, stop for a few seconds to think and then resume your discussion. Here's your first question.

Question 1

President Reagan's economic program (called Reaganomics) includes tax cuts, budget cuts across most programs, and increases in defense spending. Take the next 3 minutes to tell me your opinions regarding "Reaganomics".

Question 2

There has been support in Europe and more recently in this country for complete nuclear disarmament for both the United States and the Soviet Union. Take the next 3 minutes to tell me your opinions on nuclear disarmament.

Question 3

Due to time limitations and a lack of support from the Reagan administration, the Equal Rights Amendment may not be passed. Take the next 3 minutes to tell me why you are for or against the passage of the ERA.

APPENDIX B

Introductory Instructions for the Self-Monitoring Phase and Question Series B

Instructions

During the next 9 minutes, I will ask you another series of three questions. After each question you will have 3 minutes to discuss your answer.

Question 1

The president and some members of congress support legislation that would eliminate federal funding for abortion for low-income women. Take the next 3 mintues to tell me why you are for or against abortion and whether it should be federally funded.

Question 2

With the recent military build-up, some members of congress have made proposals to reinstate the draft. Take the next 3 minutes to tell me why you are for or against the draft.

Question 3

Every year thousands of laboratory animals are killed or maimed during scientific experiments. Recently, efforts have been made to

stop or closely supervise the use of animals in laboratory experiments. But, researchers claim such action would severely impede scientific research. Take the next 3 minutes to tell me your opinions on using animals for scientific experiments in which they are killed or maimed.

REFERENCES

Anthony, W. Z. Brief intervention in a case of childhood trichotillomania by self-monitoring. Journal of Behavior Therapy and Experimental Psychiatry, 1978, 9, 173-175.

Bandura, A. Principles of behavior modification. New York: Holt, Rinehart & Winston, 1969.

Bandura, A. Self-reinforcement: Theoretical and methodological considerations. Behaviorism, 1976, 4, 135-155.

Bandura, A. Social learning theory. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Bandura, A., & Perloff, B. Relative efficacy of self-monitored and externally imposed reinforcement systems. Journal of Personality and Social Psychology, 1967, 7, 111-116.

Barlow, D. H. On the relation of clinical research to clinical practice: Current issues and new directions. Journal of Consulting and Clinical Psychology, 1981, 49, 147-155.

Bayer, C. A. Self-monitoring and mild aversion treatment of trichotillomania. Journal of Behavior Therapy and Experimental Psychiatry, 1972, 3, 139-141.

Bell, D. R., & Low, R. M. Observing and recording children's behavior. Richland, Wash.: Performance Associates, 1977.

Bellack, A. S., & Hersen, M. Self-report inventories in behavioral assessment. In J. D. Cone and R. P. Hawkins (Eds.), Behavioral assessment: New directions in clinical psychology. New York: Brunner/Mazel, 1977.

Bellack, A. S., Rozensky, R., & Schwartz, J. A comparison of two forms of self-monitoring in a behavioral weight reduction program. Behavior Therapy, 1974, 5, 523-530.

Berecz, J. Modification of smoking behavior through self-administered punishment of imagined behavior: A new approach to aversive therapy. Journal of Consulting and Clinical Psychology, 1972, 38, 244-250.

Bergan, J. R. Behavioral consultation. Columbus, Ohio: Charles E. Merrill, 1977.

Bernard, M.E., Kratochwill, T. R., & Keefauver, L. W. The effects of rational-emotive therapy on chronic hair pulling. Cognitive Therapy and Research, in press.

Bolstad, O. D., & Johnson, S. M. Self-regulation in the modification of disruptive behavior. Journal of Applied Behavior Analysis, 1972, 5, 443-454.

Bornstein, P. H., Hamilton, S. B., & Quevillon, R. P. Behavior modification by long distance: Demonstration of functional control over disruptive behavior in a rural classroom setting. Behavior Modification, 1977, 1, 369-380.

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

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Catania, A. C. The myth of self-reinforcement. Behaviorism, 1975, 3, 192-199.

Cavior, N., & Marabotto, C. M. Monitoring verbal behaviors in a dyadic interaction. Journal of Consulting and Clinical Psychology, 1976, 44, 68-76.

Ciminero, A. R., Nelson, R. O., & Lipinski, D. P. Self-monitoring procedures. In A. R. Ciminero, K. S. Calhoun & H. E. Adams (Eds.), Handbook of behavioral assessment. New York: Wiley & Sons, 1977.

Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.

Cone, J. D. Psychometric considerations. In M. Hersen & A. S. Bellack (Eds.), Behavioral assessment: A practical handbook. New York: Pergamon Press, 1981.

Dapcich-Miura, E., & Hovell, M. F. Contingency management of adherence to a complex medical regimen in an elderly heart patient. Behavior Therapy, 1979, 10, 193-201.

Drabman, R. S., Spitalnick, R., & O'Leary, K. D. Teaching self-control to disruptive children. Journal of Abnormal Psychology, 1973, 82, 10-16.

Emmelkamp, P. M. Self-observation versus flooding in the treatment of agoraphobia. Behavior Research and Therapy, 1974, 12, 229-237.

Emmelkamp, P. M. & Kraanen, J. Therapist-controlled exposure in vivo versus self-controlled exposure in vivo: A comparison with obsessive compulsive patients. Behavior Research and Therapy, 1977, 15, 491-495.

Emmelkamp, P. M., & Kwee, K. G. Obsessional ruminations: A comparison between thought stopping and prolonged exposure in imagination. Behavior Research and Therapy, 1977, 15, 441-444.

Emmelkamp, P. M., & Ultee, K. A. A comparison of "successive approximation" and "self-observation" in the treatment of agoraphobia. Behavior Therapy, 1974, 5, 606-613.

Epstein, L. H., Webster, J. S., & Miller, P. M. Accuracy and controlling effects of self-monitoring as a function of concurrent responding and reinforcement. Behavior Therapy, 1975, 6, 654-666.

Ernst, F. Self-recording and counterconditioning of a self-mutilative compulsion. Behavior Therapy, 1973, 4, 144-146.

Feuerstein, M., & Adams, H. E. Cephalic vasomotor feedback in the modification of migraine headaches. Biofeedback and Self-Regulation, 1977, 3, 241-254.

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

Flowers, J. V. Behavior modification of cheating in an elementary school student: A brief note. Behavior Therapy, 1972, 3, 311-312.

Fredericksen, L. W., Epstein, L. H., & Kosevsky, B. P. Reliability and controlling effects of three procedures for self-monitoring smoking. Psychological Record, 1975, 25, 255-264.

Friedman, S. Self-control in the treatment of Gilles de la Tourette's Syndrome: Case study with 18-month follow-up. Journal of Consulting and Clinical Psychology, 1980, 48, 400-402.

Gelfand, D. M., & Hartmann, D. P. Child behavior analysis and therapy. New York: Pergamon Press, 1976.

Glynn, E. L., & Thomas, J. D. Effect of cueing on self-control of classroom behavior. Journal of Applied Behavior Analysis, 1974, 7, 299-306.

Glynn, E. L., Thomas, J. D., & Shee, S. M. Behavioral self-control of classroom behavior in an elementary classroom. Journal of Applied Behavior Analysis, 1973, 6, 105-113.

Gottman, J. M., & McFall, R. M. Self-monitoring effects in a program for potential high school dropouts: A time-series analysis. Journal of Consulting and Clinical Psychology, 1972, 39, 273-281.

Green, L. Temporal and stimulus factors in self-monitoring by obese persons. Behavior Therapy, 1978, 9, 328-341.

Hallahan, D. P., Lloyd, J., Kosiewicz, M. M., Kauffman, J. T., & Graves, A. W. Self-monitoring of attention as a treatment for a learning disabled boy's off-task behavior. Learning Disability Quarterly, 1979, 2, 24-32.

Hamilton, S. B., & Bornstein, P. H. Increasing the accuracy of self-recording in speech anxious undergraduates through the use of self-monitoring training and reliability enhancement procedures. Journal of Consulting and Clinical Psychology, 1975, 43, 577-583.

Hannum, J. W., Thoresen, C. E., & Hubbard, D. R. A behavioral study of self-esteem with elementary teachers. In M. J. Mahoney and C. E. Thoresen (Eds.), Self-control: Power to the person. Monterey, Calif.: Brooks-Cole, 1974.

Harmon, T. M., Nelson, R. O., & Hayes, S. C. The differential effects of self-monitoring mood versus activity in depressed patients. Journal of Consulting and Clinical Psychology, 1980, 48, 30-38.

Hartmann, D. P. Considerations in the choice of interobserver reliability estimates. Journal of Applied Behavior Analysis, 1977, 10, 103-116.

Hay, L. R., Hay, W. M., & Angle, H. V. The reactivity of self-recording: A case report of a drug abuser. Behavior Therapy, 1977, 8, 1004-1007.

Hayes, S. C., & Cavior, N. Multiple tracking and the reactivity of self-monitoring: I. Negative behaviors. Behavior Therapy, 1977, 8, 819-831.

Hayes, S. C., & Cavior, N. Multiple tracking and the reactivity of self-monitoring: II. Positive behaviors. Behavior Therapy, 1980, 2, 283-296.

Haynes, S. N. Principles of behavioral assessment. New York: Halstead Press, 1978.

Haynes, S. N., & Wilson, C. C. Behavioral assessment: Recent advances in methods, concepts and applications. San Francisco: Jossey-Bass, 1979.

Herbert, E. W., & Baer, D. M. Training parents as behavior modifiers: Self-recording of contingent attention. Journal of Applied Behavior Analysis, 1972, 5, 139-149.

Hersen, M., & Barlow, D. H. Single case experimental designs: Strategies for studying behavior change. New York: Pergamon Press, 1976.

Hersen, M., & Bellack, A. S. Behavioral assessment: A practical handbook. New York: Pergamon Press, 1981.

Horne, D. J. Behavior therapy for trichotillomania. Behavior Research and Therapy, 1977, 15, 192-196.

Hundert, J., & Batstone, D. A practical procedure to maintain pupils' accurate self-rating in a classroom token program. Behavior Modification, 1978, 2, 93-111.

Hundert, J., & Bucher, B. Pupils' self-scored arithmetic performance: A practical procedure for maintaining accuracy. Journal of Applied Behavior Analysis, 1978, 11, 304.

Hutzel, R. R. Analogue self-recording therapy: Expectancy effects. Dissertation Abstracts International, 1976, 36, 4161.

Hutzel, R. R., Platzek, D., & Logne, P. E. Control of symptoms of Giles de la Tourette's Syndrome by self-monitoring. Journal of Behavior Therapy and Experimental Psychiatry, 1974, 5, 71-76.

Jackson, B. Treatment of depression by self-reinforcement. Behavior Therapy, 1972, 3, 298-307.

Jeffrey, D. B. Self-control: Methodological issues and research trends. In M. J. Mahoney and C. E. Thoresen (Eds.), Self-control: Power to the person. Monterey, Calif.: Brooks-Cole, 1974.

Johnson, S. M., & White, G. Self-observation as an agent of behavior change. Behavior Therapy, 1971, 2, 488-497.

Jones, J. C., Trap, J., & Cooper, J. O. Technical report: Students' self-recording of manuscript letter strokes. Journal of Applied Behavior Analysis, 1977, 10, 509-514.

Kanfer, F.H. Self-regulations: Research, issues, and speculations. In. C. Neuringer and J. L. Michael (Eds.), Behavior modification in clinical psychology. New York: Appleton-Century-Crofts, 1970. (a)

Kanfer, F. H. Self-monitoring: Methodological limitations and clinical applications. Journal of Consulting and Clinical Psychology, 1970, 35, 148-152. (b)

Kanfer, F. H. Maintenance of behavior by self-generated stimuli and reinforcement. In A. Jacobs and L. B. Sachs (Eds.), The psychology of private events. New York: Academic Press, 1971.

Kanfer, F. H. Self-management methods. In F. H. Kanfer and A. P. Goldstein (Eds.), Helping people change. New York: Pergamon, 1975.

Kanfer, F. H. The many faces of self-control, or behavior modification changes its focus. In R. B. Stuart (Ed.), Behavior self-management: Strategies, techniques, and outcomes. New York: Brunner/Mazel, 1977.

Kanfer, F. H., & Karoly, P. Self-control: A behavioristic excursion into the lion's den. Behavior Therapy, 1972, 3, 398-416.

Karoly, P. Behavioral self-management in children. Concepts, methods, issues and directions. In M. Hersen, R. Eisler, & P. Miller (Eds.), Progress in behavior modification, vol. 5. New York: Academic Press, 1977.

Katz, R. C. A procedure for concurrently measuring elapsed time and response frequency. Journal of Applied Behavior Analysis, 1973, 6, 719-720.

Kaufman, K. F., & O'Leary, K. D. Reward, cost, and self-evaluation procedures for disruptive adolescents in a psychiatric hospital school. Journal of Applied Behavior Analysis, 1972, 5, 293-309.

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

Kazdin, A. E. Self-monitoring and behavior change. In M. J. Mahoney & C. E. Thoresen (Eds.), Self-control: Power to the person. Monterey, Calif.: Brooks-Cole, 1974. (b)

Kazdin, A. E. Assessing the clinical or applied importance of behavior change through social validation. Behavior Modification, 1977, 1, 427-452.

Kazdin, A. E. Unobtrusive measures in behavioral assessment. Journal of Applied Behavior Analysis, 1979, 12, 712-724.

Kazdin, A. E. Research design in clinical psychology. New York: Harper & Row, 1980.

Kilmann, P. R., Wagner, M. K., & Sotile, W. M. The differential impact of self-monitoring on smoking behavior. An exploratory study. Journal of Clinical Psychology, 1977, 33, 912-914.

Kirschenbaum, D. S. & Karoly, P. When self-regulation fails: Tests of some preliminary hypotheses. Journal of Consulting and Clinical Psychology, 1977, 45, 1116-1125.

Knapczyk, D. R. & Livingston, G. Self-recording and student teacher supervision: Variables within a token economy. Journal of Applied Behavior Analysis, 1973, 6, 481-486.

Kolb, D. A., Winter, S. K. & Berlew, D. E. Self-directed behavior change: Two studies. Journal of Applied Behavior Science, 1968, 4, 453-471.

Knox, D. Marriage happiness. Champaign, Ill.: Research Press, 1972.

Kratochwill, T. R. Single subject research: Strategies for evaluating change. New York: Academic Press, 1978.

Kunzelmann, H. D. (Ed.), Precision teaching. Seattle: Special Child Publications, 1970.

Layne, C. C., Rickard, H. C., Jones, M. T., & Lyman, R. D. Accuracy of self-monitoring on a variable ratio schedule of observer verification. Behavior Therapy, 1976, 7, 481-488.

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

Lindsley, O. R. A reliable wrist counter for recording behavior rates. Journal of Applied Behavior Analysis, 1968, 1, 77-78.

Lipinski, D. P., Black, J. L., Nelson, R. O., & Ciminero, A. R. Influence of motivational variables on the reactivity and reliability of self-recording. Journal of Consulting and Clinical Psychology, 1974, 42, 118-123.

Litrownik, A. J., & Freitas, J. L. Self-monitoring in moderately retarded adolescents: Reactivity and accuracy as a function of valence. Behavior Therapy, 1980, 11, 245-255.

Lubor, J. F., & Shouse, M. N. Use of biofeedback in the treatment of seizure disorders and hyperactivity. In B. Lahey and A. E. Kazdin (Eds.), Advances in clinical child psychology, vol. 1. New York: Plenum Press, 1977.

Lyman, R. D., Rickard, H. C., & Elder, I. R. Contingency management of self-report and cleaning behavior. Journal of Abnormal Child Psychology, 1975, 3, 155-162.

Mahoney, M. J. The self-management of covert behavior: A case study. Behavior Therapy, 1971, 2, 575-578.

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

Mahoney, M. J. Some applied issues in self-monitoring. In J. D. Cone & R. P. Hawkins (Eds.), Behavioral assessment: New directions in clinical psychology. New York: Brunner/Mazel, 1977.

Mahoney, M. J., & Arnkoff, D. Cognitive and self-control therapies. In S. L. Garfield and A. E. Bergin (Eds.), Handbook of psychotherapy and behavior change. New York: Wiley & Sons, 1978.

Mahoney, M. J., Moore, B. S., Wade, T. C., & Moura, N. The effects of continuous and intermittent self-monitoring on academic behavior. Journal of Consulting and Clinical Psychology, 1973, 41, 65-69.

Mahoney, M. J., Moura, N. G., & Wade, T. C. The relative efficacy of self-reward, self-punishment and self-monitoring techniques for weight loss. Journal of Consulting and Clinical Psychology, 1973, 40, 404-407.

Maletzky, B. M. Behavior recording as treatment: A brief note. Behavior Therapy, 1974, 5, 107-111.

Mathews, A., & Shaw, P. Cognitions related to anxiety: A pilot study of treatment. Behavior Research and Therapy, 1977, 15, 503-505.

McFall, R. M. Effects of self-monitoring on normal smoking behavior. Journal of Consulting and Clinical Psychology, 1970, 35, 135-142.

McFall, R. M. Parameters of self-monitoring. In J. D. Cone and R. P. Hawkins (Eds.), Behavior assessment: New directions in clinical psychology. New York: Brunner/Mazel, 1977.

McFall, R. M., & Hammen, C. L. Motivation, structure, and self-monitoring: Role of nonspecific factors in smoking reduction. Journal of Consulting and Clinical Psychology, 1971, 37, 80-86.

McKenzie, T. L., & Rushall, B. S. Effects of self-recording on attendance and performance in a competitive swimming training environment. Journal of Applied Behavior Analysis, 1974, 7, 199-206.

McLaughlin, T. F. Self-control in the classroom. Review of Educational Research, 1976, 46, 631-663.

McNamara, J. R. The use of self-monitoring techniques to treat nail biting. Behavior Research and Therapy, 1972, 10, 193-194.

Meichenbaum, D. Teaching children self-control. In B. Lahey and A. E. Kazdin (Eds.), Advances in child clinical psychology, vol. 1. New York: Plenum Press, 1977.

Mercatoris, M., & Craighead, W. E. Effects of nonparticipant observation on teacher and pupil classroom behavior. Journal of Educational Psychology, 1974, 66, 512-519.

Miller, A. J., & Kratochwill, T. R. Reduction of frequent stomachache complaints by timeout. Behavior Therapy, 1979, 10, 211-218.

Mitchel, K. R., & White, R. G. Behavioral self-management: An application to the problem of migraine headaches. Behavior Therapy, 1977, 8, 213-221.

Monti, P. M., McCrady, B. S., & Barlow, D. H. Effect of positive reinforcement, informational feedback, and contingency contracting on a bulimic anorexic female. Behavior Therapy, 1977, 8, 258-263.

Morris, R. J., & Hoschouer, R. L. Current issues in applied research with mentally retarded persons. Applied Research in Mental Retardation, 1980, 1, 85-94.

Myers, J. L. Fundamentals of experimental design. Boston: Allyn and Bacon, Inc., 1979.

Nelson, C., & McReynolds, W. Self-recording and control of behavior: Reply to Simkins. Behavior Therapy, 1971, 2, 594-597.

Nelson, R. O. Assessment and therapeutic functions of self-monitoring. In M. Hersen, R. Eisler, and P. Miller (Eds.), Progess in behavior modification, vol. 5. New York: Academic Press, 1977. (a)

Nelson, R. O. Issues in assessment via self-monitoring. In J. D. Cone & R. P. Hawkins (Eds.), Behavioral assessment: New directions in clinical psychology. New York: Brunner/Mazel, 1977. (b)

Nelson, R. O. Realistic dependent measures for clinical use. Journal of Consulting and Clinical Psychology, 1981, 49, 168-182.

Nelson, R. O., Hay, L. R., Hay, W. M., & Carstens, C. B. The reactivity and accuracy of teachers' self-monitoring of positive and negative classroom verbalizations. Behavior Therapy, 1977, 8, 972-985.

Nelson, R. O., & Hayes, S. Theoretical explanations for reactivity in self-monitoring. Behavior Modification, 1981, 5, 3-14.

Nelson, R. O., Kapust, J. A., & Dorsey, B. L. Minimal reactivity of overt classroom observations on student and teacher behaviors. Behavior Therapy, 1978, 9, 695-702.

Nelson, R. O., Lipinski, D. P., & Black, J. L. The effects of expectancy on the reactivity of self-recording. Behavior Therapy, 1975, 6, 337-349.

Nelson, R. O., Lipinski, D. P., & Black, J. L. The reactivity of adult retardates' self-monitoring: A comparison among behaviors of different valences and a comparison with token reinforcement. Psychological Record, 1976, 26, 189-201.

Nelson, R. O., Lipinski, D. P., & Boykin, R. A. The effects of self-recorders' training and the obtrusiveness of the self-recording device on the accuracy and reactivity of self-monitoring. Behavior Therapy, 1978, 9, 200-208.

Novaco, R. W. Stress inoculation: A cognitive therapy for anger and its application to a case of depression. Journal of Consulting and Clinical Psychology, 1977, 45, 600-608.

Ollendick, T. H. Self-monitoring and self-administered overcorrection: The modification of nervous tics in children. Behavior Modification, 1981, 5, 75-84.

Orne, M. T. From the subject's point of view, when is behavior private and when is it public: Problems of inference. Journal of Consulting and Clinical Psychology, 1970, 35, 143-147.

Parsonson, B. S., & Baer, D. M. The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), Single subject research: Strategies for evaluating change. New York: Academic Press, 1978.

Peacock, R., Lyman, R. D., & Richard, H. C. Correspondence between self-report and observer report as a function of task difficulty. Behavior Therapy, 1978, 9, 578-583.

Peterson, G. L., House, A. E., & Alford, H. F. Self-monitoring: Accuracy and reactivity in a patient's recording of three clinical targeted behaviors. Presented at the meeting of the Southeastern Psychological Association, Atlanta, March 1975.

Piersel, W. C. Behavioral consultation: An approach to problem solving. In J. R. Bergan (Ed.), School psychology in contemporary society. Columbus, Ohio: Merrill, in press.

Piersel, W. C., Brody, G. H. & Kratochwill, T. R. A further examination of motivational influences on disadvantaged minority group children's intelligence test performance. Child Development, 1977, 48, 1142-1145.

Piersel, W. C., & Kratochwill, T. R. Self-observation and behavior change: Applications to academic and adjustment problems through behavioral consultation. Journal of School Psychology, 1979, 17, 151-161.

Plachetta, K. E. Encopresis: A case study utilizing contracting, scheduling and self-charting. Journal of Behavior Therapy and Experimental Psychiatry, 1976, 7, 195-196.

Plummer, S., Baer, D. M., & LeBlanc, J. M. Functional considerations in the use of procedural timeout and effective alternative. Journal of Applied Behavior Analysis, 1977, 10, 689-705.

Rachlin, H. Self-control. Behaviorism, 1974, 2, 94-107.

Rainwater, N., & Ayllon, T. Increasing academic performance by using a timer as antecedent stimulus: A study of four cases. Behavior Therapy, 1976, 7, 672-677.

Rekers, G. A., Amaro-Plotlin, H. D., & Low, B. P. Sex-typed mannerisms in normal boys and girls as a function of sex and age. Child Development, 1977, 48, 275-278.

Rekers, G. A. & Varni, J. W. Self-monitoring and self-reinforcement processes in a pretranssexual boy. Behavior Research and Therapy, 1977, 15, 177-180.

Ribordy, S. C., & Denny, D. R. The behavioral treatment of insomnia: An alternative to drug therapy. Behavior Research and Therapy, 1977, 15, 177-180.

Roberts, R., & Dick, L. Self-control strategies with children. In T. R. Kratochwill (Ed.), Advances in school psychology, Vol. II. Hillsdale, N.J.: Erlbaum, in press.

Romanczyk, R. G. Self-monitoring in the treatment of obesity: Parameters of reactivity. Behavior Therapy, 1974, 5, 531-540.

Rosensky, R. H. The effect of timing self-monitoring behavior on reducing cigarette consumption. Journal of Behavior Therapy and Experimental Psychiatry, 1974, 5, 301-303.

Ross, A. O. Of rigor and relevance. Professional Psychology, 1981, 12, 318-327.

Rutner, I. T., & Bugle, C. An experimental procedure for the modification of psychotic behavior. Journal of Consulting and Clinical Psychology, 1969, 33, 651-653.

Sagotsky, G., Patterson, C. J., & Lepper, M. R. Training children's self-control: A field experiment in self-monitoring and goal setting in the classroom. Journal of Experimental Child Psychology, 1978, 25, 242-253.

Santogrossi, D. A. Self-reinforcement and external monitoring of performance on an academic task. Paper read at the fifth annual Converence on Applied Behavior Analysis in Education, Kansas City, October, 1974.

Santogrossi, D. A., O'Leary, K. D., Romanczyk, R. G., & Kaufman, K. F. Self-evaluation by adolescents in a psychiatric hospital school token program. Journal of Applied Behavior Analysis, 1973, 6, 277-287.

Schwartz, G. J. College students as contingency managers for adolescents in a program to develop reading skills. Journal of Applied Behavior Analysis, 1977, 10, 645-655.

Seymour, F. W., & Stokes, T. F. Self-recording in training girls to increase work and evoke staff praise in an institution for offenders. Journal of Applied Behavior Analysis, 1976, 9, 41-54.

Sheehan, D. J., & Casey, B. Communication. Journal of Applied Behavior Analysis, 1974, 7, 446.

Sieck, W. A., & McFall, R. M. Some determinants of self-monitoring effects. Journal of Consulting and Clinical Psychology, 1976, 44, 958-965.

Simkins, L. The reliability of self-recorded behaviors. Behavior Therapy, 1971, 2, 83-87.

Skinner, B. F. Science and human behavior. New York: Macmillian, 1953.

Skinner, B. F. Verbal behavior. New York: Appleton-Century-Crofts, 1957.

Sobell, L. C., & Sobell, M. B. Validity of self-reports in three populations of alchoholics. Journal of Consulting and Clinical Psychology, 1978, 46, 901-907.

Spates, C. R. & Kanfer, F. H. Self-monitoring, self-evaluation, and self-reinforcement in children's learning: A test of multistage self-regulation model. Behavior Therapy, 1977, 8, 9-16.

Stollack, G. E. Weight loss obtained under different experimental procedures. Psychotherapy: Theory, research, and practice, 1967, 4, 61-64.

Stuart, R. B., & Davis, B. Slim chance in a fat world. Champaign, Ill.: Research Press, 1972.

Stumphauzer, J. S. Daily behavior card. Venice, Calif.: Behaviormetrics, 1974.

Sulzer-Azaroff, B., & Mayer, G. R. Applying behavior analysis procedures with children and youth. New York: Holt, Rinehart & Winston, 1977.

Switzky, H. N., & Haywood, H. C. Motivational orientation and the relative efficacy of self-monitored and externally imposed reinforcement systems in children. Journal of Personality and Social Psychology, 1974, 30, 360-366.

Tasto, D. L. Self-report schedules and inventories. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), Handbook of behavioral assessment. New York: Wiley & Sons, 1977.

Thomas, D. R. Preliminary findings on self-monitoring for modifying teacher behaviors. In E. A. Ramp and B. L. Hopkins (Eds.), A new direction for education: Behavior analysis, vol. 1. Lawrence: University of Kansas, 1971.

Thomas, E. J., Abrams, K. S., & Johnson, J. B. Self-monitoring and reciprocal inhibition in the modification of multiple tics of Gilles de la Tourette's syndrome. Journal of Behavior Therapy and Experimental Psychiatry, 1971, 2, 159-171.

Thoresen, C. E., & Mahoney, M. J. Behavioral self-control. New York: Holt, 1974.

Turkewitz, H., O'Leary, K. D., & Ironsmith, M. Generalization and maintenance of appropriate behavior through self-control. Journal of Consulting and Clinical Psychology, 1975, 43, 577-583.

Turner, S. M., Hersen, M., & Bellack, A. S. Effects of social disruption, stimulus interference, and aversive conditioning on auditory hallucinations. Behavior Modification, 1977, 1, 249-258.

Vargas, J. M., & Adesso, V. J. A comparison of aversion therapies for nailbiting behavior. Behavior Therapy, 1976, 7, 322-329.

Watson, D. L., & Tharp, R. G. Self-directed behavior change: Self-modification for personal adjustment. Monterey, Calif.: Brooks-Cole, 1972.

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