

An Overview of Self-Monitoring Research in Assessment and Treatment

William J. Korotitsch and Rosemary O. Nelson-Gray
University of North Carolina at Greensboro

Although widely utilized within clinical assessment, self-monitoring has received little direct research attention since the early 1980s. The assessment and treatment functions of self-monitoring are described and illustrated in this article. As an assessment method, self-monitoring can provide valuable information for diagnosis, target behavior selection, functional assessment, and treatment monitoring. Research findings delineating variables known to influence the accuracy and reactivity associated with self-monitoring are reviewed. Some suggestions are made for additional research that might be conducted regarding self-monitoring accuracy, effects on clients, and utility within treatment.

The term *self-monitoring* is used here to refer to an assessment procedure that involves data collection made by the client primarily within naturalistic settings. Self-monitoring involves two component responses. First, the client must discriminate or notice the occurrence of the target behavior. This may be an action, thought, or feeling. Second, the client must produce some record of the occurrence as well as any additional information (e.g., intensity ratings or antecedent stimuli) that is relevant to the particular goals of assessment. This latter response is referred to here as *self-recording*. These two component responses may occur relatively close together or may be separated temporally in cases where target behaviors are recorded at intervals. They are best regarded as independent responses that may be influenced by different variables (Simkins, 1971). For example, a client may effectively discriminate occurrences of a target behavior, but the client may be reluctant to record these for reasons of social approval or, similarly, the client may record behaviors that have not occurred. Ideally, both responses would be controlled by the occurrence of the target response. Both are necessary for accurate data to be collected.

Self-monitoring has generally served two functions within behavior therapy. The first is an assessment function for which accuracy is an important consideration. In addition, self-monitoring may also be utilized as part of treatment due to its reactive effects on those behaviors being monitored. These reactive effects tend to occur in the therapeutically desired direction (i.e., undesirable behaviors decrease in frequency while desirable behaviors increase in frequency). Both functions are discussed in greater detail in later portions of this article.

History of Self-Monitoring in Clinical Practice and Research

Accurate assessment of responses and their controlling variables is a cornerstone of behavior therapy. Self-monitoring is, of course,

not unique to behavior therapy. Historically, the procedure has been traced to classical psychophysics and introspective methods (Kazdin, 1974b). An historical example of self-monitoring is provided by Thoreson and Mahoney (1974), who describe a self-monitoring procedure used by Benjamin Franklin in his attempts to increase development of 13 virtues. Self-monitoring procedures probably predate these documented examples and are a reasonable component of any self-guided approach to behavior change.

Within clinical assessment, self-monitoring procedures were certainly popularized by behavior therapists and enjoyed widespread use, particularly within behavioral self-control procedures (Bornstein, Hamilton, & Bornstein, 1986). Bornstein, Bridgewater, Hickey, and Sweeney (1980) surveyed assessment methods reported in behavioral journals and concluded that self-monitoring represented 38% of the assessment procedures used. In a more recent survey, 83% of behaviorally oriented practitioners reported using self-monitoring with 44% of their clients (Elliott, Miltenberger, Kaster-Bundgaard, & Lumley, 1996). This continued widespread use of self-monitoring procedures is *prima facie* evidence not only of its utility but also of its flexibility as an assessment method. As behavior therapists have become increasingly concerned with cognitive responses and more comprehensive treatments targeting multiple covarying responses, self-monitoring continues to occupy a prominent role in assessment.

Although most self-monitoring is probably designed for idiographic assessment of clinical cases, there are also nomothetic and standardized approaches. Currently, self-monitoring procedures are described and recommended within most empirically supported treatments. For example, Agras and Apple (1997) describe a daily food record for use in their cognitive-behavioral treatment program for bulimia. Using this form, the client records the type and amount of food consumed, episodes of bingeing and purging, and the use of laxatives or diuretics. For each food intake, the client also records the time, place, and events that influenced eating. An example of a similar form can be found in Wilson and Vitousek (1999, this issue). Similarly, Craske, Barlow, and O'Leary (1992) described the use of a worry record by clients participating in the Mastery of Your Anxiety and Worry program. Clients are instructed to use sharp increases in anxiety as cues to complete this form. Anxiety ratings and symptoms are recorded as

William J. Korotitsch and Rosemary O. Nelson-Gray, Department of Psychology, University of North Carolina at Greensboro.

Correspondence concerning this article should be addressed to William J. Korotitsch, Department of Psychology, Room 296 Eberhart Building, University of North Carolina at Greensboro, Greensboro, North Carolina 27412-5001.

well as information regarding the situation and the content of anxiety-related thoughts.

Historically, research on self-monitoring has focused on both its assessment and therapeutic functions. The main concern regarding the assessment function of self-monitoring has been with variables influencing accuracy, whereas research regarding therapeutic functions has focused on variables contributing to reactivity. Research regarding self-monitoring was widespread during the 1970s. The results of this early research form the basis for much of the current discussion. Even though the use of self-monitoring continues to be widespread within clinical settings, research regarding this assessment method has markedly declined since the early 1980s (Korotitsch, Gaynor, & Nelson-Gray, 1998; Taylor, 1999).

Behavioral assessment methods have been placed on a continuum of directness, with direct observations in naturalistic setting being the most direct, interview and self-report measures being the most indirect, and self-observation falling between these two (Cone, 1978). Unfortunately, these measures can also be placed on a continuum of convenience that appears to be precisely the inverse of directness. The most direct measures are the most difficult to use, particularly in clinical settings where resources are limited. The current trend in assessment appears to be toward the development of more rapid and easily administered measures (i.e., questionnaires and rating scales). This appears in part to reflect the influence of the current environment of managed care. The number of self-report questionnaires and rating scales reported in the literature has increased dramatically since the early 1980s (Froyd, Lambert, & Froyd, 1996). This increase directly parallels the decrease in research applied to self-monitoring and other more direct methods of behavioral assessment (Froyd et al., 1996; Korotitsch et al., 1998). However, there is evidence that clinicians frequently use self-monitoring and value this more direct assessment method (Elliott et al., 1996). The climate of managed care also demands high-quality assessment that can enhance formulation of effective intervention as well as sensitive measures that can detect changes in the behaviors of interest. Behavior therapists, applying principles from basic behavioral research, have been instrumental in shaping the standards for defining efficacious treatments (Nelson-Gray, Gaynor, & Korotitsch, 1997). Behavior therapists, relying on continued research in behavioral assessment methods, including self-monitoring, could similarly be instrumental in defining the standards for measuring treatment effects.

Assessment Functions of Self-Monitoring

As an assessment tool, self-monitoring has several advantages. In many cases, problem behaviors may be under the control of many different variables that may not be obvious to the client or present in the therapy context. In these cases, clinically relevant behaviors may be best assessed within the patient's everyday life. Direct observation may be costly or difficult to implement in these settings. Self-monitoring provides an inexpensive alternative for data collection that requires few clinician resources for implementation. Additionally, many behaviors of clinical interest are covert by nature, such as obsessions or paranoid thoughts, or are private by social convention, such as sexual activity. In these cases, the nature of the target behavior obviates the use of external observers or alternative in vivo assessment methods.

Bornstein et al. (1986) noted several additional advantages of self-monitoring: (a) The procedure implicitly emphasizes the client's potential control over behavior; (b) it provides clients with continuous immediate feedback regarding behavior outside of the therapy session; and (c) it can potentially provide a more complete and more thorough account of behavior than can be obtained through many other methods.

Self-monitoring may be used in all stages of assessment. For example, it may be used to clarify a diagnosis or diagnosis-related treatment selection. Further, it may be used in conducting a functional analysis. Finally, self-monitoring data may be used in evaluating treatment outcome. Each of these assessment uses is elaborated in the following discussion. Detailed examples of the application of self-monitoring in the assessment and treatment of particular behavior disorders can be found in other articles in this Special Section.

Self-Monitoring Within Diagnostic-Descriptive Assessment

Therapists have increasingly utilized diagnostic results both to document treatment needs and to guide the selection of efficacious treatments that are often organized by diagnostic category (e.g., Chambless et al., 1996). The gold standard of diagnostic evaluation is the semistructured diagnostic interview. However, self-monitoring is frequently incorporated into assessments that are geared at diagnostic or syndromal classification. In addition to providing information regarding the frequency and intensity of symptoms, self-monitored data can be used to document the disruption that symptoms may produce in multiple settings. Self-monitored data may also be used between assessment sessions to gather additional data relevant to diagnostic questions. For example, to establish a diagnosis of generalized anxiety disorder, the clinician must determine that the client's worry is not secondary to any other Axis I diagnosis. This may be difficult to establish in cases where an existing Axis-I disorder disrupts multiple areas of client functioning. For instance, a young adult with an attention deficit hyperactivity disorder diagnosis may report excessive worry related to academic, vocational, and relationship areas. These worries may be difficult to distinguish from distress stemming from the disruption posed by inattentive and hyperactive-impulsive symptoms. Asking the client to self-monitor episodes of worry and to include more detailed information regarding content and situations may help to clarify the diagnostic formulation.

Once a diagnosis has been established, self-monitored data may be helpful in deciding which aspects of treatment may be most beneficial to a particular client. For example, when using cognitive-behavioral therapy for bulimia, the clinician might propose a treatment that emphasizes behavioral interventions if self-monitored data revealed poor eating habits or a treatment that is more cognitive if self-monitoring revealed dysfunctional cognitions.

Target Behavior Selection

Self-monitored data may also be used more specifically to identify particular targets for treatment. Clients may present with vague treatment goals, and potential targets for treatment may initially be difficult to identify. Diary formats are particularly useful during this stage of assessment. These are relatively open-

ended and allow the client to impart large amounts of information regarding their difficulties and the circumstances surrounding problem behaviors that may aid target behavior selection. For example, for a depressed client, self-monitored data might reveal that the person experiences few pleasant events, a large number of unpleasant events, and/or dysfunctional cognitions. Specific treatment goals can be formulated using this information.

Self-Monitoring Within Functional Assessment

In the initial stage of functional assessment, the goals, broadly stated, concern identification of clinically relevant responses and potential controlling variables. Controlling variables may be multiple and difficult to delineate. Therefore, a diary format for self-monitoring would initially be advisable. As assessment proceeds, more precise information regarding a more limited range of responses may be desired, and self-monitoring may increasingly focus on these behaviors.

The first step in identifying a functional relationship is determining the covariance of responses and their potential controlling variables (Haynes, Leisen, & Blaine, 1997; O'Brien, 1995). The most common method of determining covariance is through therapist judgment (Nelson & Hayes, 1986). O'Brien (1995) provided a cautionary discussion and demonstration of inaccuracies in covariance estimation. Touchette, MacDonald, and Langer (1985) described a scatter plot method of recording frequency data that may be helpful in narrowing the range of potential controlling variables. Once these variables are identified for a given client, covariance can be further assessed through self-monitoring. For example, a clinician can evaluate the covariation between a potentially important antecedent and a target response by asking clients to record occurrences of the antecedent both when the target response occurs and when it does not.

Conditional probabilities may also be used to evaluate potential controlling variables. For example, Schlundt, Johnson, and Jarrell (1985) evaluated the covariance of various antecedents with the eating and purging behaviors of 8 bulimic and 23 overweight patients. Participants used self-monitoring booklets organized in a multiple-choice format in which they recorded all food intake. For each instance of food intake, participants recorded time, date, location, social context, activities engaged in before eating, whether or not they overate, and, for bulimic patients, whether or not they purged after food intake. The authors identified a functional relationship when the conditional probability of eating given the presence of an antecedent differed statistically from the unconditional probability of eating. Using this approach, they gathered information regarding bulimic and overeating groups that could be important to treatment planning. For example, in the afternoon and evening hours, bulimic participants were at greater risk for bingeing and purging when at home than when away from home. Alternatively, overweight patients were at greater risk for overeating when eating away from home. Negative mood was positively related to bingeing and purging in bulimic participants, whereas both negative and positive mood was related to overeating in obese participants.

Use of Self-Monitoring in Treatment Monitoring

Once a treatment plan has been formulated and initiated, self-monitoring can be an integral part of the evaluative phase of

therapy. Self-monitoring can provide a relatively continuous measure of client progress as treatment proceeds, thereby providing ongoing feedback regarding effectiveness to both the client and the practitioner. Depending on this feedback, the practitioner can continue or modify interventions without waiting to conduct more elaborate midtreatment and posttreatment evaluations.

In addition to monitoring progress through changes in target responses, practitioners can also use self-monitoring to check compliance with therapeutic interventions and treatment integrity. For example, in dialectical behavior therapy (DBT) for the treatment of borderline personality disorder (Linehan, 1993a, 1993b), clients are asked to complete diary cards during assessment and treatment. Using this format, clients monitor the frequency and intensity of target behaviors such as parasuicidal acts or medication use. Clients also record their use of behavioral coping skills acquired in treatment and the effectiveness of their implementation of these skills in situations where target behaviors occur. Similarly, Beck, Rush, Shaw, and Emery (1979) described a Daily Record of Dysfunctional Thoughts (DRDT) for use in cognitive therapy for depression. Using this form, clients record information regarding situations in which unpleasant emotions occur, the specific emotional response, and automatic thoughts that precede or accompany the emotion. In addition, clients monitor rational responses to these automatic thoughts and the outcome (including ratings of the degree of belief associated with an automatic thought and the intensity of emotion). These examples illustrate the potential benefit of the collection of self-monitored data in gauging the client's application of therapeutic procedures outside of sessions and in identifying any difficulties that may be encountered in their implementation.

Accuracy of Self-Monitored Data

Efforts have been made to establish the validity of self-monitoring, and a number of studies addressing this issue are discussed within the articles in this Special Section. Additionally, treatment outcome studies frequently utilize multiple measures of process and change in addition to self-monitoring, providing evidence for the convergent validity of self-monitoring methods. An important issue with regard to the assessment function of self-monitoring is the accuracy of the data collected, which is one dimension of validity (see Jackson, 1999, [this issue] for a discussion). True accuracy cannot be determined, especially for covert or private behavior, and accuracy is expected to vary with situational factors. Accuracy of self-monitoring has typically been estimated using three types of criteria. The first is comparison to data obtained by independent direct observers. The second is comparison of self-monitored data to mechanical devices. Third is comparison to behavioral byproducts. For instance, self-monitored caloric intake might be evaluated relative to weight changes (Mahoney, Moura, & Wade, 1973).

Variables Affecting Accuracy

As mentioned earlier, much of the research concerning self-monitoring has been aimed at establishing the accuracy of self-recorded data. A number of studies have been concerned with identifying variables that enhance or attenuate accuracy. It has been suggested that self-observers have a unique capacity for

accuracy, because they are in the position of being able to potentially observe all occurrences of their target behaviors (Kazdin, 1974b). In addition, numerous variables have been identified that influence the accuracy of self-monitored data. This information helps to define optimal methods for applying this assessment method. Variables affecting the accuracy of self-monitoring are now addressed.

Awareness of accuracy checks. It has been repeatedly demonstrated that self-monitors are more accurate when they are informed that the accuracy of their recordings will be substantiated (Lipinski, Black, Nelson, & Ciminerio, 1975; Lipinski & Nelson, 1974; Nelson, Lipinski, & Black, 1975). In a typical study, data collected by self-monitors and by external observers are compared, with self-monitors either aware or unaware of this comparison. Lipinski and Nelson (1974) found that agreement between self-monitors and trained external observers was .86 when self-monitors were aware of accuracy checks, but it was only .52 when accuracy checks were made without their knowledge. Similar results for verbal responses were reported by Bailey and Peterson (1977).

The accuracy of self-monitored data might be checked against less frequently available observational data. For example, daily self-monitored data from children may be compared to data collected by parental observations 1 day each week (Cohen, Gelfand, Dodd, Jensen, & Turner, 1980).

Topography of the target behavior. Bailey and Peterson (1977) suggested that the accuracy of self-monitored verbal responses may be lower than that of motoric responses. Peterson, House, and Alford (1975) found 64.3% agreement for self-monitored face touching, 0% agreement for the phrase "you know," and 31.4% agreement for the phrase "and all that." Similarly, Hayes and Cavior (1977) found higher agreement for self-recorded face touching than for self-recorded value judgments or speech dysfluencies. However, Willis and Nelson (1982) found no differential agreement as a function of target responses in a study using nonclinical participants.

Training. A number of studies have demonstrated that training improves the accuracy of self-monitored data (e.g., Mahoney, 1977; Nelson, Lipinski, & Boykin, 1978). Mahoney (1977) provided some suggestions for effective training: (a) explicit definitions of target behaviors and examples should be provided; (b) self-monitoring instructions should be made explicit; (c) appropriate use of self-recording instruments should be modeled by the therapist; (d) the definitions of target behaviors and self-monitoring instructions should be repeated by the client; and (e) several occurrences of the target behavior as defined within the assessment should be monitored by the client.

Compliance. The accuracy of any assessment depends in large part on issues of compliance. Compliance with self-monitoring procedures can be enhanced if verbal commitments or contracts are made (Kanfer, Cox, Greiner, & Karoly, 1974; Levy, 1977). If clients are asked to make monetary deposits with refunds contingent on completing self-monitoring, then compliance increases (Ersner-Hersfield, Connors, & Maisto, 1981). Regular data collection through random telephone calls to clients (Christensen, Johnson, Phillips, & Glasgow, 1980) or by asking clients to mail self-monitored data on a daily basis (Harmon, Nelson, & Hayes, 1980) has been shown to increase compliance.

Bornstein, Hamilton, Miller, Quevillon, and Spitzform (1977) also reported increased compliance by (a) using foot-in-the-door approaches (Freedman & Fraser, 1966), where small requests with which the client is likely to comply are made before the requests for data collection; (b) increasing self-esteem by emphasizing the self-monitors' desire to be accurate; (c) inducing guilt by emphasizing the negative effects or costs of inaccurate data; and (d) using written reminders on data sheets emphasizing the need for accuracy and honesty.

Shelton and Levy (1981) suggested common reasons for non-compliance: (a) The client lacks necessary skills and knowledge to complete the task; (b) cognitions may interfere with completing the task; and (c) the client's environment does not support compliance. When confronted with noncompliance, it is important to consider these factors and provide additional training, more detailed rationale, or any appropriate support that may be needed. In some cases, due to stressful environments or other challenges to maintaining self-monitoring, alternative assessment methods might be more appropriate.

Accuracy-contingent reinforcement. Improved accuracy is also observed when reinforcement is provided contingent on the accuracy of self-monitored data (e.g., Lipinski et al., 1975). Additionally, punishment has been used to decrease discrepancies in self-monitored data (Seymour & Stokes, 1976). Some studies have examined the maintenance of self-monitoring accuracy after a reduction in reinforcement. These studies have tended to produce mixed results, with some studies failing to maintain accuracy (e.g., Turkewitz, O'Leary, & Ironsmith, 1975) and with others maintaining it (e.g., Layne, Rickard, Jones, & Lyman, 1976). Training in self-monitoring of cognitions has been attempted (Meyers, Mercatoris, & Artz, 1976) but is complicated by the lack of independently assessed criteria for accuracy.

Nature of the recording device. Despite suggestions that obtrusiveness of recording devices might influence accuracy, Nelson et al. (1978) found no differential accuracy between an obtrusive handheld counter and a less obtrusive belt-worn counter. Additional research is needed that addresses characteristics of recording devices that may influence accuracy.

Concurrent response requirements. The accuracy of self-monitored data decreases when self-monitors are required to engage in other responses concurrently. For example, Epstein, Miller, and Webster (1976) observed increased errors in self-monitoring of respiration when participants engaged in a concurrent operant lever-pressing task. Hayes and Cavior (1977) reported significantly higher accuracy in self-monitoring of one target behavior as opposed to two or three responses. Baird and Nelson-Gray (1999) suggested that improved accuracy might be observed if additional target behaviors were gradually introduced into the self-monitoring procedure.

Valence of the target behavior. Kanfer (1977) suggested that the accuracy of self-monitoring may be lower for negatively valenced behaviors than for positively valenced behaviors. He suggested that negative self-evaluation prompted by attending to undesirable behaviors may precipitate some avoidance of observing and recording these responses. Consistent with this hypothesis, Nelson, Hay, Devany, and Koslow-Green (1980a) found that children more accurately self-monitored appropriate classroom verbalizations as opposed to inappropriate verbalizations. Nelson, Hay, and Carstens (1977) examined the accuracy of teachers'

self-monitoring negative and positive verbalizations in the classroom with similar results. However, Willis and Nelson (1982) found that female undergraduates self-recorded negatively valenced targets more accurately than positively valenced targets.

Alternatives to Self-Monitored Data

Self-monitoring can be more economical and more convenient than direct observation. However, other measures may be more convenient than self-monitoring, such as asking the client to report retrospectively on occurrences of target behaviors within the therapy session or using questionnaire measures. The utility of self-monitoring in assessment hinges in part on the assumption that it is a more accurate and sensitive measurement than these other more convenient alternatives. Evaluations of questionnaire measures have tended to focus on their diagnostic accuracy, and we are aware of no direct comparisons between self-monitoring and questionnaire methods. Research addressing the relative accuracy and sensitivity of self-monitoring and self-report methods is needed (Taylor, 1999).

In evaluating the frequency of a target behavior, it may sometimes be the case that self-monitoring provides little incremental advantage over clients reporting retrospectively the number of times responses of interest occurred. Hasher and Zacks (1979) have hypothesized that information regarding the frequency of events is "automatically encoded." They suggested that, rather than frequency estimation involving effortful recall and counting of instances, information regarding frequency is inevitably registered in memory as a result of attending to events and requires no additional effort for encoding or retrieval. They reviewed a number of studies demonstrating that retrospective reports of the frequency of a range of observable events can be extremely accurate and relatively impervious to variables that influence more effortful recall such as instructions, practice, age, and ability.

Contrary to the automatic-encoding model, some studies have demonstrated that frequency estimation can be sensitive to variables affecting recall. For example, some studies have observed a deleterious effect of participants engaging in secondary tasks during encoding (Fisk & Schneider, 1984; Greene, 1984). Other studies have demonstrated effects of levels of processing manipulations, with semantic encoding conditions yielding better recall as well as more accurate frequency estimates (e.g., Fisk & Schneider, 1984; Jonides & Naveh-Benjamin, 1987). These studies support the suggestion that effortful recall and counting multiple instances may serve as the primary basis for frequency estimates (see Watkins & LeCompte, 1991) as opposed to less effortful retrieval of summative frequency information. The mixed findings suggest that more basic research examining the processes through which frequency estimates are generated is needed. At the moment, however, it cannot be assumed that frequency estimates are invariably subject to distortion from the many factors that have been shown to deleteriously affect effortful recall.

The implication of the automatic-encoding model for clinical assessment is that accurate data about target behavior frequency may, under some circumstances, be derived from estimates at intervals rather than from continuously recorded occurrences. Analogue research comparing counting to frequency estimates of word presentations within lists has demonstrated superior accuracy of counting but has also concluded that estimation is not greatly

discrepant from actual frequency (e.g., Nelson-Gray, Herbert, Herbert, Farmer, Badawi, & Lin, 1990). However, more research is needed to extend this line of research to clinical phenomena and populations.

A few studies have compared accuracy of self-monitored data with retrospective estimates and have concluded that retrospective estimates can provide reasonably accurate data. For example, Frederikson, Epstein, and Kosevsky (1975) reported good accuracy for participants who estimated the number of cigarettes they smoked at the end of a day or week compared with participants who continuously recorded the number of cigarettes smoked. However, it is possible that participants used their knowledge of a typical day or week to guide their summative retrospective reports (Farmer & Nelson-Gray, 1990).

Retrospective ratings of mood have also been compared with self-monitoring of mood to determine their accuracy. For example, Parkinson, Briner, Reynolds, and Totterdell (1995) compared global ratings of mood taken at 1-week intervals and at the end of each day with mood ratings taken at 2-hr intervals during waking hours. Participants entered their mood ratings in handheld computers and were not allowed access to these ratings when making global judgments. Accuracy is defined within this study as the correspondence of global summative daily or weekly ratings and the actual momentary ratings made throughout these time periods. Daily ratings showed good correspondence to the average of momentary ratings and did not appear to be biased by any extremes in mood that were recorded over the course of the day. Similarly, weekly ratings showed greater correspondence to average daily ratings than to any extreme daily ratings. Despite some tendency of participants to overestimate positive mood, global ratings provided reasonably accurate information regarding overall mood experienced across intervals. Hedges, Jandorf, and Stone (1985) reported similar results in comparing daily mood ratings with mood ratings recorded at intervals across the day.

In contrast to the aforementioned studies, some researchers have observed significant biases in retrospective estimates as compared to self-monitored data. For example, de Beurs, Lange, and Van Dyck (1992) found a much lower frequency of panic attacks observed in self-monitored data than was reported retrospectively by panic-disordered clients. They cautioned that retrospectively reported overestimation of panic attacks produced overestimation of treatment effects in their study. In this issue, Wilson and Vitousek (1999) discuss studies providing evidence for bias in the retrospective estimates of eating behaviors within clinical populations. Similarly, Barton, Blanchard, and Veazey (1999) discuss evidence for bias in global retrospective estimates of pain.

More research is needed to establish the relative accuracy of frequency estimation versus self-monitoring within clinical assessment situations. Although highly preliminary, this line of research suggests that, under some circumstances, estimation may prove to be a reasonably accurate alternative method for assessing some responses of interest. However, even though some studies have concluded that estimation is "reasonably" accurate compared to continuous recording, there is no conventional standard for the degree of differential accuracy that would warrant the use of self-monitoring over estimation or vice versa. The required accuracy could be ascertained empirically by providing data of varying accuracy to clinicians and evaluating the influence of these data on proposed interventions.

The decision of whether estimates could be used in lieu of self-monitoring would also depend on the goals of the assessment, that is, response occurrence versus functional analysis. The aforementioned findings cannot be generalized to the collection of information about antecedent events and consequences because these may not be as amenable to retrospective report. The observation that individuals have difficulty identifying and reporting controlling variables for their responding has frequently been made (e.g., Kohlenberg & Tsai, 1991; Nisbett & Wilson, 1977). However, rather than looking at the comparative accuracy of self-monitored versus self-estimated controlling variables, it would seem more productive to focus on their comparative utility. This is because, within the context of assessment, we are concerned with a *subset* of functional variables that may be evident, that is, those that may be modified in treatment (Haynes & O'Brien, 1990). The utility of functional variables identified through self-monitoring should be considered within the broader question about the treatment utility of functional assessment. Despite numerous calls for such an evaluation of functional analysis (Hayes & Follette, 1992; Haynes & O'Brien, 1990), little work has been done in this area (although, see Schill, Kratochwill, & Elliott, 1998, for an example of this type of study). Certainly, the relative utility of self-monitored versus self-estimated data could be evaluated within this context (Haynes & O'Brien, 1990).

In summary, on the basis of the existing findings, a prototypical self-monitoring procedure that would maximize accuracy would be one in which (a) target responses are well-defined, overt, and positive in valence; (b) the self-monitor is trained; (c) accuracy checks are possible, and the client is aware that these will be made at least intermittently; (d) the importance of accurate data within the context of treatment is described, emphasized, and illustrated in therapy sessions; (e) reinforcement contingent upon accurate data is provided; (f) concurrent response requirements are minimized; and (g) recordings are made as soon as possible after each occurrence of the target behavior. These conclusions are summarized in Table 1.

Treatment Function of Self-Monitoring

In addition to its assessment function, self-monitoring has been utilized as a form of intervention primarily because of its reactive

effects. As noted earlier, reactivity refers to the change in the frequency of a target behavior that is a result of the self-monitoring procedure. Reactive effects are problematic when self-monitoring is used as an assessment measure, because the measurement method itself alters target frequency. However, these reactive changes occur in the therapeutic direction and can therefore provide some benefits toward behavior change. Reactive effects of self-monitoring have been documented for a wide range of clinically relevant behaviors including hallucinations (Rutner & Bugle, 1969), paranoid ideation (Williams, 1976), ruminative thinking (Frederikson, 1975), insomnia (Jason, 1975), alcohol consumption (Sobell & Sobell, 1973), amphetamine abuse (Hay, Hay, & Angle, 1977), and suicidal ideation (Clum & Curtin, 1993). Reactive effects tend to be fairly small but reliable and do not appear to be susceptible to demand (Nelson, Kapust, Dorsey, & Hayes, 1977) and expectancies (Nelson et al., 1975; Hutzell, 1977). This suggests that reactive effects may make an adjunctive contribution to beneficial treatment effects when used in the context of other interventions. Although unique effects of self-monitoring have been demonstrated in some studies (e.g., Gottman & McFall, 1972; McFall, 1970), the extant findings are mixed (see Kazdin, 1974b, for a review), and additional research is needed to address this question. Although fairly small and transient, reactive effects have the advantage of being fairly immediate. These small but immediate improvements may help to maintain the client's initial investment in therapy.

Other assessment methods, including direct observation, may also produce reactive effects. However, the portability and convenience of self-monitoring make it an appealing method for capitalizing on these reactive effects to the benefit of clients. A number of factors have been found to influence the magnitude of reactive effects associated with self-monitoring. These factors should be considered when self-monitoring procedures are used as a form of intervention.

It is important to note that a number of factors affect both the reactivity and accuracy of self-monitoring. This is true despite the apparent independence of these processes. There is little correspondence between accuracy and reactive effects, and the former is not necessary for the occurrence of the reactive effects of self-monitoring (Brodin, Hall, & Mitts, 1971; Fixen, Phillips, &

Table 1
Variables Affecting the Accuracy of Self-Monitored Data

Variables affecting accuracy	Optimal conditions for accuracy
Awareness of accuracy checks	Accuracy of data is checked randomly, and the client is aware of accuracy checks.
Topography of the target behavior	The target behavior is well-defined and motoric-overt.
Training	The client is trained in the use of self-monitoring, including explicit instructions and modeling of procedures.
Compliance	Steps are taken to maximize and monitor compliance with the self-monitoring procedure.
Reinforcement	Reinforcement is provided contingent on accurate data collection.
Nature of recording device	No effect of obtrusiveness of the recording device on accuracy has been demonstrated.
Concurrent response requirements	A single target behavior is self-monitored.
Valence of the target behavior	The target behavior is positive in valence.
Timing of recordings	Recordings are made as soon as possible after the occurrence of the target behavior.

Wolf, 1972; Herbert & Baer, 1972; Lipinski & Nelson, 1974; Marshall, Lloyd, & Hallahan, 1993).

Variables Affecting Reactivity

A number of variables have been demonstrated to influence the magnitude and direction of reactive effects. These are described in the following sections.

Target behavior valence. One important variable is the valence of the target behavior, with positively valenced or desirable behaviors increasing in frequency and negatively valenced or undesirable behaviors decreasing in frequency. Kazdin (1974a) demonstrated this effect when differential valence was assigned to the same target behavior. Self-monitoring increased self-reference statements when they were assigned a positive valence and decreased self-reference statements when they were assigned a negative valence. Sieck and McFall (1976) reported similar results when assigning positive and negative valence to eye blinking. Cavior and Marabotto (1976) reported similar directional effects of self-monitoring on behaviors selected by participants that they regarded as positive or negative.

Motivation for change. Motivation of the individual to change the responses being monitored is possibly related to valence effects. A number of studies have demonstrated increased reactivity in persons desiring to change the behavior being monitored. Participants with high motivation have typically been identified as those volunteering for treatment research or requesting help in changing the target behavior. Conversely, low-motivation participants would be those who agree to participate for money or research credit. For example, Lipinski et al. (1975) compared reactivity in groups of self-monitors who volunteered for a project requesting "individuals who want to quit smoking" to those who signed up for a project requesting volunteers "who are cigarette smokers." Self-monitoring reduced smoking only within the motivated group. Similarly, Komaki and Dore-Boyce (1978) found that self-monitoring significantly increased classroom participation only in those students who expressed an interest in changing this behavior.

Topography of the target. Peterson et al. (1975) suggested that reactive effects may be less pronounced when verbal rather than nonverbal responses are self-monitored. They observed greater reactivity for self-monitored frequency of face touching than for the frequency of using two verbal expressions. Hayes and Cavior (1977) reported similar effects of self-monitoring verbal versus motoric responses. Other studies have indicated that reactive effects can depend on the specific target behavior selected for self-monitoring. For example, Romanczyk (1974) found greater weight loss when both weight and caloric intake were self-monitored compared to when weight alone was self-monitored. Gottman and McFall (1972) observed increases in the frequency of actual classroom participation as a result of self-monitoring. Conversely, self-monitoring of unfulfilled urges to participate produced a decrease in the frequency of such urges.

Schedule of recording. Greater reactivity occurs if each occurrence of a target behavior is recorded. For example, Mahoney et al. (1973) reported that continuous self-recording of correct answers resulted in longer study sessions than recording every third correct answer.

Concurrent response requirements. There is evidence that reactivity of self-monitoring decreases when multiple responses are monitored concurrently. Hayes and Cavior (1977) compared reactive effects when participants monitored one, two, or three behaviors. They found significantly greater reactive effects when only one behavior was self-monitored as compared to multiple responses.

Timing of recording. There is also evidence that the timing of recording relative to the target response influences reactivity. Kanfer (1970) suggested that reactivity may be enhanced if self-recording occurred before, rather than after the target response. He reasoned that the act of self-recording would produce a disruption in the chain of behaviors culminating in the target response and would function as an alternative response. Consistent with this hypothesis, when smokers were instructed to record smoking just before its occurrence, reactive effects were increased (Rozensky, 1974). Similarly, Bellack, Rozensky, and Schwartz (1974) found greater weight loss when participants were asked to record food intake before eating, although some contrary findings have also been reported (e.g., Nelson et al., 1977).

Goal-setting feedback and reinforcement. Reactivity can also be enhanced by goal setting, providing feedback to self-monitors, and providing reinforcement contingent on behavior change. Kazdin (1974a) found increased self-reference statements when participants were provided with a goal specifying the number of statements to make during self-monitoring. The amount of feedback given to participants was also manipulated by allowing some participants to view their counter displays. Those participants who were given feedback made more self-reference statements than participants whose counter displays were covered. Richards, McReynolds, Holt, and Sexton (1976) found that students who were relatively unaware of the amount of time they spent studying benefited more from feedback than those who could already report the extent of this time. Kolb, Winter, and Berlew (1968) observed greater behavior change in self-monitors who met weekly to discuss their individual projects. Nelson, Lipinski, and Black (1976) found that when developmentally disabled adults believed edible reinforcers were contingent on changes in response frequency, their frequency of talking and touching external objects increased and face touching decreased. Lipinski et al. (1975) observed similar effects of monetary reinforcement on the frequency of face touching in college students.

Nature of the self-recording device. There is some evidence that the obtrusiveness of the self-recording device can influence reactivity. Nelson et al. (1978) found greater reactivity when participants used more obtrusive hand-held counters as opposed to a less-obtrusive belt-worn counter. Kirby, Fowler, and Baer (1991) also found evidence of greater reactivity when children used an obtrusive self-recording device to record math performance. However, two additional studies failed to observe differential reactivity as a result of the obtrusiveness of recording devices (Nelson, Hay, Devany, & Koslow-Green, 1980b; Schloss, Thompson, Gajar, & Schloss, 1985).

Theoretical Explanations of Reactivity of Self-Monitoring

Three primary theoretical explanations have been offered to account for the reactive effects of self-monitoring. Kanfer (1977) proposed a three-stage mediational model of reactive effects. In the first stage, the self-monitor observes and records the occurrence of

the target behavior. The second stage involves an internal process of self-evaluation. The self-monitor compares information regarding the occurrence of the target behavior to an internalized self-standard or criterion for performance. The third stage involves self-reinforcement in the form of self-praise if the performance criterion has been met or self-punishment if the occurrence of the target behavior fails to meet self-standards. Kanfer suggests that it is these self-administered consequences that increase or decrease the future probability of the target response, thereby producing behavior change in the therapeutically desired direction.

Rachlin (1974) offered an alternative account that has been referred to as the operant or nonmediational model. Rachlin avoids positing a mediational role of covert self-praise or punishment in producing reactive effects. According to his model, both the act of self-monitoring and the administration of self-consequences function as discriminative stimuli that serve as cues for the ultimate environmental consequences for engaging in the target response. For example, self-recording of cigarette smoking functions as a cue for ultimate environmental consequences such as social disapproval or illness. It is these environmental consequences that influence the future probability of the target response.

Nelson and Hayes (1981) offered an elaboration of Rachlin's model. They agreed that the act of self-monitoring functions as a cue for environmental consequences. However, they proposed that, in addition to self-recording and self-consequating, the entire self-monitoring procedure—including instructions, training, and the presence of the self-recording device—serves to cue environmental consequences. They argued that this elaboration of the operant model could better account for the empirical findings concerning reactive effects in self-monitoring. For example, by emphasizing multiple cues, including the presence of the self-monitoring device, the Nelson and Hayes model can provide an account for observations that reactive effects of self-monitoring are maintained even when the frequency of the target behavior is zero (e.g., Maletzky, 1974). Their model also suggests that accurate self-monitoring should not be necessary for reactive effects to occur. Their model accounts for observations of reactive effects even when the accuracy of self-monitoring is poor (Lipinski & Nelson, 1974; Hayes & Cavior, 1977) and for the negligible correlations observed between reactivity and accuracy (Hayes & Cavior, 1977).

In summary, based on the existing findings, a prototypical self-monitoring procedure that would maximize reactivity would

be one in which (a) target responses are well-defined and motoric, (b) the client is motivated to change, (c) explicit goals for change are formulated, (d) reinforcement is provided contingent on reactive effects, (e) recordings are made just before each occurrence of the target behavior, and (f) concurrent response requirements are minimized. Possibly, more obtrusive recording devices would be used; however, more work is needed to establish the effects of obtrusive devices on reactivity. These conclusions are summarized in Table 2.

Suggestions for Additional Research

As mentioned earlier, there has been a dramatic decline in research devoted to self-monitoring after the early 1980s. This decline is well documented (Korotitsch et al., 1998; Taylor, 1999) and appears to mirror a general decline in behavioral assessment research. Some suggestions for potential research in self-monitoring are offered here. In many cases, such research can be incorporated into other projects already utilizing self-monitoring methods, particularly those evaluating the efficacy of therapeutic interventions. The majority of research described earlier has been conducted with analogue populations and simple target behaviors. One general avenue of research could involve replications of this work with clinical populations and more clinically relevant clinical target responses to establish their generalizability to treatment settings and clinically relevant behaviors. Of course, there are additional considerations and precautions needed in clinical research over and above those in analogue research (Kazdin, 1998).

More specific research questions might be focused on accuracy and sensitivity, the effects of self-monitoring on clients, and the utility of self-monitoring in treatment.

Accuracy and Sensitivity

One important question for research concerns the relative accuracy and sensitivity of self-monitoring compared to other assessment methods. Although self-monitoring has frequently been compared to direct observation, future research may compare the accuracy and sensitivity of self-monitoring to other frequently utilized measures such as client self-ratings or global estimates, self-report questionnaires, clinician ratings, and semistructured interview results. The accuracy and sensitivity of measures is of

Table 2
Variables Affecting the Reactivity of Self-Monitoring

Variables affecting reactivity	Optimal conditions for enhanced reactivity
Target behavior valence	Positively valenced behaviors increase in frequency while negatively valenced behaviors decrease in frequency.
Motivation	The client is highly motivated to change the target behavior.
Topography of the target behavior	The target behavior is overt-motoric.
Schedule of recording	Each occurrence of the target behavior is self-monitored.
Concurrent response requirements	A single target behavior is self-monitored.
Timing of recording	Recordings are made just before the occurrence of the target response.
Goal setting, feedback, and reinforcement	Goals for changes in the target behavior are clearly specified, and reinforcement contingent on behavior change is provided.
Nature of the self-recording device	Some studies have indicated that an obtrusive recording device enhances reactivity.

great practical importance to outcome research. Additionally, with increasing demands on clinicians for documentation of beneficial treatment effects, this information can aid practitioners in selecting the most appropriate and efficient measures for this purpose.

A related question concerns the relative accuracy of self-monitoring across diagnostic groups. For example, one might hypothesize that depressed clients might be generally less accurate in self-monitoring than nondepressed or clinical controls. Alternatively, it may be suggested that depressed persons would evidence greater accuracy in self-monitoring negatively valenced target behaviors. If differences in accuracy among diagnosed groups are observed, it may be important to establish whether the relative accuracy of self-monitoring changes over the course of treatment.

Additional research might also more generally evaluate potential bias of self-monitored data as a function of variables such as social desirability or malingering response styles. In all cases, researchers should be sensitive to the distinction between the component responses of detecting and recording within self-monitoring. It is likely that any of the aforementioned variables might differentially influence these responses.

Effects on Clients

Research might also be conducted concerning the effects of self-monitoring procedures on clients. One line of research might examine whether self-monitoring might have negative or detrimental effects for some clients. For example, self-monitoring might exacerbate symptoms such as obsessional thoughts by focusing attention on these responses. Another line of research might examine client satisfaction with self-monitoring compared with other assessment methods.

Utility in Treatment

Finally, the impact of self-monitoring on the quality of treatment may be evaluated. As suggested earlier, this question might be incorporated into a broader examination of the utility of functional assessment in formulating interventions. More generally, this type of research might also address the question of whether self-monitored data that is revealed to clinicians enhances the effectiveness of treatment. This question could be addressed using a group design where self-monitored data are made available to therapists for some clients and not others. It might also be addressed using a multiple baseline design across behaviors where self-monitored data for different behaviors are revealed to therapists in a sequential fashion. The utility and quality of data collected with different self-monitoring formats (e.g., continuous recording vs. momentary time sampling) or with different recording methods (e.g., diaries vs. handheld computers) could also be evaluated in this context. The most important question for any assessment procedure, to quote Meehl (1959) is "in what way and to what extent does this . . . information help us in treating the patient?" (p. 117).

Summary

Despite a decline in research directed at self-monitoring methods, self-monitoring continues to be widely used both within clinical settings and within outcome research. Self-monitoring can

provide detailed information that is useful in diagnostic evaluation, target behavior selection, and functional assessment. Self-monitoring also provides continuous feedback regarding treatment effects thereby enabling the clinician to make necessary adjustments to maximize benefits to clients. Due to its reactive effects, self-monitoring has also been used as an intervention and has been demonstrated to produce beneficial effects for a wide variety of clinically relevant behaviors. Early research has identified a number of variables that influence the accuracy and reactivity of self-monitoring. Information regarding these effects can be used to maximize the effectiveness of self-monitoring, both as an assessment and an intervention strategy. Much of the research regarding self-monitoring has been conducted with nonclinical populations who self-monitored readily observable but nonclinical behaviors. Additional research is needed to generalize these results to clinical settings and populations. Research might also be conducted to determine the benefits or detriments of self-monitoring procedures for categories of clients and the utility of self-monitored data in treatment.

References

- Agras, W. S., & Apple, R. F. (1997). *Overcoming eating disorders: Therapist guide*. New York: Graywind Publications.
- Bailey, M. I., & Peterson, G. L. (1977). *Reactivity and accuracy of self-monitored verbal responses*. Unpublished manuscript.
- Baird, S., & Nelson-Gray, R. O. (1999). Direct observation and self-monitoring. In S. C. Hayes, D. H. Barlow, & R. O. Nelson-Gray (Eds.), *The scientist practitioner* (2nd ed., pp. 353-386). New York: Allyn & Bacon.
- Barton, K. A., Blanchard, E. B., & Veazey, C. (1999). Self-monitoring as an assessment strategy in behavioral medicine. *Psychological Assessment, 11*, 490-497.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Bellack, A. S., Rozensky, R., & Schwartz, J. (1974). A comparison of two forms of self-monitoring in a behavioral weight reduction program. *Behavior Therapy, 5*, 523-530.
- Bornstein, P. H., Bridgwater, C. A., Hickey, J. C., & Sweeney, T. M. (1980). Characteristics and trends in behavioral assessment: An archival analysis. *Behavioral Assessment, 2*, 125-133.
- Bornstein, P. H., Hamilton, S. B., & Bornstein, M. T. (1986). Self-monitoring procedures. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (2nd ed.). New York: Wiley.
- Bornstein, P. H., Hamilton, S. B., Miller, R. K., Quevillon, R. P., & Spitzform, M. (1977). Reliability and validity enhancement: A treatment package for increasing the fidelity of self-report. *Journal of Clinical Psychology, 33*, 861-866.
- Brodin, M., Hall, R. V., & Mitts, B. (1971). The effect of self-recording on the classroom behavior of two eighth-grade students. *Journal of Applied Behavior Analysis, 4*, 191-199.
- Cavior, N., & Marabotto, C. M. (1976). Monitoring verbal behaviors in a dyadic interaction: Valence of target behavior, type, timing, and reactivity to monitoring. *Journal of Consulting and Clinical Psychology, 44*, 68-76.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Johnson, S. B., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on empirically validated therapies. *The Clinical Psychologist, 49*, 5-18.
- Christensen, A., Johnson, S. M., Phillips, S., & Glasgow, R. E. (1980). Cost effectiveness in behavioral family therapy. *Behavior Therapy, 11*, 208-226.

- Clum, G. A., & Curtin, L. (1993). Validity and reactivity of a system of self-monitoring suicide ideation. *Journal of Psychology and Behavioral Assessment, 15*, 375-385.
- Cohen, E. A., Gelfand, D. M., Dodd, D. K., Jensen, J., & Turner, C. (1980). Self-control practices associated with weight loss maintenance in children and adolescents. *Behavior Therapy, 11*, 26-37.
- Cone, J. D. (1978). The Behavioral Assessment Grid (BAG): A conceptual framework and a taxonomy. *Behavior Therapy, 9*, 882-888.
- Craske, M. G., Barlow, D. H., & O'Leary, T. A. (1992). *Mastery of your anxiety and worry: Client workbook*. New York: Graywind Publications.
- de Beurs, E., Lange, A., & Van Dyck, R. (1992). Self-monitoring of panic attacks and retrospective estimates of panic: Discordant findings. *Behaviour Research and Therapy, 30*, 411-413.
- Elliot, A. J., Miltenberger, R. G., Kaster-Bundgaard, J., & Lumley, V. (1996). A national survey of assessment and therapy used by behavior therapists. *Cognitive and Behavioral Practice, 3*, 107-125.
- Epstein, L. H., Miller, P. M., & Webster, J. S. (1976). The effects of reinforcing concurrent behavior on self-monitoring. *Behavior Therapy, 7*, 89-95.
- Ersner-Hersfield, S. M., Connors, G. L., & Maisto, S. A. (1981). Clinical and experimental utility of refundable deposits. *Behaviour Research and Therapy, 19*, 455-457.
- Farmer, R., & Nelson-Gray, R. O. (1990). The accuracy of counting versus estimating event frequency in behavioral assessment: The effects of behavior frequency, number of behaviors monitored, and time delay. *Behavioral Assessment, 12*, 425-442.
- Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 181-197.
- Fixen, D. L., Phillips, E. L., & Wolf, M. M. (1972). Achievement place: The reliability of self-reporting and peer-reporting and their effects on behavior. *Journal of Applied Behavior Analysis, 5*, 19-30.
- Frederiksen, L. W. (1975). Treatment of ruminative thinking by self-monitoring. *Behavior Therapy and Experimental Psychiatry, 6*, 258-259.
- Frederiksen, L. W., Epstein, L. H., & Kosevsky, B. P. (1975). Reliability and controlling effects of three procedures for self-monitoring and smoking. *Psychological Record, 25*, 255-264.
- Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology, 4*, 195-202.
- Froyd, J. E., Lambert, M. J., & Froyd, J. D. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health, 5*, 11-15.
- Gottman, J. M., & McFall, R. M. (1972). Self-monitoring effects in a program for potential high school dropouts: A time-series analysis. *Journal of Consulting and Clinical Psychology, 39*, 273-281.
- Greene, R. L. (1984). Incidental learning of event frequency. *Memory and Cognition, 12*, 90-95.
- Harmon, T. M., Nelson, R. O., & Hayes, S. C. (1980). Self-monitoring mood versus activity in depressed clients. *Journal of Consulting and Clinical Psychology, 48*, 30-38.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108*, 356-388.
- Hay, L. R., Hay, W. M., & Angle, H. V. (1977). The reactivity of self-recording: A case report of a drug abuser. *Behavior Therapy, 8*, 1004-1007.
- Hayes, S. C., & Cavior, N. (1977). Multiple tracking and the reactivity of self-monitoring: Negative behaviors. *Behavior Therapy, 8*, 819-831.
- Hayes, S. C., & Follette, W. C. (1992). Can functional analysis provide a substitute for syndromal classification? *Behavioral Assessment, 14*, 345-365.
- Haynes, S. N., Leisen, M. B., & Blaine, D. D. (1997). Design of individualized behavioral treatment programs using functional analytic clinical case models. *Psychological Assessment, 9*, 334-348.
- Haynes, S. N., & O'Brien, W. H. (1990). Functional analysis in behavior therapy. *Clinical Psychology Review, 10*, 649-668.
- Hedges, S. M., Jandorf, L., & Stone, A. A. (1985). Meaning of daily mood assessments. *Journal of Personality and Social Psychology, 48*, 428-434.
- Herbert, E. W., & Baer, D. M. (1972). Training parents as behavior modifiers: Self-recording of contingent attention. *Journal of Applied Behavior Analysis, 5*, 139-149.
- Hutzell, R. R. (1977). Expectancy effects upon self-recorded eye-blinking behavior. *Behavior Therapy, 8*, 832-839.
- Jackson, J. L. (1999). Psychometric considerations in self-monitoring assessment. *Psychological Assessment, 11*, 439-447.
- Jason, L. (1975). Rapid improvement in insomnia following self-monitoring. *Journal of Behavior Therapy and Experimental Psychiatry, 6*, 349-350.
- Jonides, J., & Naveh-Benjamin, M. (1987). Estimating frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 230-240.
- Kanfer, F. H. (1970). Self-monitoring: Methodological limitations and clinical applications. *Journal of Consulting and Clinical Psychology, 35*, 148-152.
- Kanfer, F. H. (1977). The many faces of self-control, or behavior modification changes its focus. In R. B. Stuart (Ed.), *Behavioral self-management: Strategies, techniques, and outcomes* (pp. 1-48). New York: Brunner/Mazel.
- Kanfer, F. H., Cox, L. E., Greiner, J. M., & Karoly, P. (1974). Contracts, demand characteristics and self-control. *Journal of Personality and Social Psychology, 30*, 605-619.
- Kazdin, A. E. (1974a). Reactive self-monitoring: The effects of response desirability, goal setting, and feedback. *Journal of Consulting and Clinical Psychology, 42*, 704-716.
- Kazdin, A. E. (1974b). Self-monitoring and behavior change. In M. J. Mahoney & C. E. Thorsen (Eds.), *Self-control: Power to the person* (pp. 218-246). Monterey, CA: Brooks-Cole.
- Kazdin, A. E. (1998). *Research design in clinical psychology*. Boston: Allyn & Bacon.
- Kirby, K. C., Fowler, S. A., & Baer, D. M. (1991). Reactivity in self-recording: Otrusiveness of recording procedure and peer comments. *Journal of Applied Behavior Analysis, 24*, 487-498.
- Kohlenberg, R. J., & Tsai, M. (1991). *Functional analytic psychotherapy: Creating intense and curative therapeutic relationships*. New York: Plenum.
- Kolb, D. A., Winter, S. K., & Berlew, D. E. (1968). Self-directed behavior change: Two studies. *Journal of Applied Behavior Science, 4*, 453-471.
- Komaki, J., & Dore-Boyce, K. (1978). Self-recording: Its effects on individuals high and low in motivation. *Behavior Therapy, 9*, 65-72.
- Korotitsch, W., Gaynor, S., & Nelson-Gray, R. O. (1998, November). *Trends in behavior therapy research: A poster presentation analysis*. Poster session presented at the 32nd annual convention of the Association for the Advancement of Behavior Therapy, Washington, DC.
- Layne, C. C., Rickard, H. C., Jones, M. T., & Lyman, R. D. (1976). Accuracy of self-monitoring on a variable ratio schedule of observer verification. *Behavior Therapy, 7*, 481-488.
- Levy, R. L. (1977). Relationship of an overt commitment to task compliance in behavior therapy. *Journal of Behavior Therapy and Experimental Psychiatry, 8*, 25-29.
- Linehan, M. M. (1993a). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press.
- Linehan, M. M. (1993b). *Skills training manual for treating borderline personality disorder*. New York: Guilford Press.
- Lipinski, D. P., Black, J. L., Nelson, R. O., & Ciminero, A. R. (1975). The influence of motivational variables on the reactivity and reliability of

- self-recording. *Journal of Consulting and Clinical Psychology*, 43, 637-646.
- Lipinski, D. P., & Nelson, R. O. (1974). Problems in the use of naturalistic observation as a means of behavioral assessment. *Behavior Therapy*, 42, 118-123.
- Mahoney, M. J. (1977). Some applied issues in self-monitoring. In J. D. Cone & R. P. Hawkins (Eds.), *Behavioral assessment: New directions in clinical psychology* (pp. 241-254). New York: Brunner/Mazel.
- Mahoney, M. J., Moura, N. G. M., & Wade, T. C. (1973). The relative efficacy of self-reward, self-punishment, and self-monitoring techniques for weight loss. *Journal of Consulting and Clinical Psychology*, 40, 404-407.
- Malctzky, B. M. (1974). Behavior recording as treatment: A brief note. *Behavior Therapy*, 5, 107-111.
- Marshall, K. J., Lloyd, J. W., Hallahan, D. P. (1993). Effects of training to increase self-monitoring accuracy. *Journal of Behavioral Education*, 3, 445-459.
- McFall, R. M. (1970). The effects of self-monitoring on normal smoking behavior. *Journal of Consulting and Clinical Psychology*, 35, 135-142.
- Meehl, P. E. (1959). Some ruminations on the validation of clinical procedures. *Canadian Journal of Psychology*, 13, 102-128.
- Meyers, A., Mercatoris, M., & Artz, L. (1976). On the development of a cognitive self-monitoring skill. *Behavior Therapy*, 7, 128-129.
- Nelson, R. O., Hay, L. R., Devany, J., & Koslow-Green, L. (1980a). The reactivity and accuracy of children's self-monitoring: Three experiments. *Child Behavior Therapy*, 2, 1-24.
- Nelson, R. O., Hay, L. R., Devany, J., & Koslow-Green, L. (1980b). The reactivity and accuracy of teachers' self-monitoring of positive and negative classroom verbalizations. *Behavior Therapy*, 8, 972-985.
- Nelson, R. O., Hay, L. R., Hay, W. M., & Carstens, C. B. (1977). The reactivity and accuracy of teachers' self-monitoring of positive and negative classroom verbalizations. *Behavior and Therapy*, 8, 972-985.
- Nelson, R. O., & Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification*, 5, 3-14.
- Nelson, R. O., & Hayes, S. C. (1986). The nature of behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 3-41). New York: Guilford Press.
- Nelson, R. O., Kapust, J. A., Dorsey, B. L., & Hayes, S. C. (1977). *Cautions in the use of self-monitoring with young children*. Unpublished manuscript.
- Nelson, R. O., Lipinski, D. P., & Black, J. L. (1975). The effects of expectancy on the reactivity of self-recording. *Behavior Therapy*, 6, 337-349.
- Nelson, R. O., Lipinski, D. P., & Black, J. L. (1976). The relative reactivity of external observations and self-monitoring. *Behavior Therapy*, 7, 314-321.
- Nelson, R. O., Lipinski, D. P., & Boykin, R. A. (1978). The effects of self-recorders' training and the obtrusiveness of the self-recording device on accuracy and the reactivity of self-monitoring. *Behavior Therapy*, 9, 200-208.
- Nelson-Gray, R. O., Gaynor, S. T., & Korotitsch, W. J. (1997). Behavior therapy: Distinct but acculturated. *Behavior Therapy*, 28, 563-572.
- Nelson-Gray, R. O., Herbert, D. L., Herbert, J. D., Farmer, R., Badawi, I., & Lin, K. (1990). The accuracy of frequency estimation as compared with actual counting in behavioral assessment. *Behavioral Assessment*, 12, 157-178.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- O'Brien, W. H. (1995). Inaccuracies in the estimation of functional relationships using self-monitoring data. *Journal of Behavior Therapy and Experimental Psychiatry*, 26, 351-357.
- Parkinson, B., Briner, R. B., Reynolds, S., & Totterdell, P. (1995). Time frames for mood: Relations between momentary and generalized ratings of affect. *Personality and Social Psychology Bulletin*, 21, 331-339.
- Peterson, G. L., House, A. E., & Alford, H. F. (1975, March). *Self-monitoring: Accuracy and reactivity in a patients' recording of three clinically targeted behaviors*. Paper presented at the meeting of the Southeastern Psychological Association, Atlanta, GA.
- Rachlin, H. (1974). Self-control. *Behaviorism*, 2, 94-107.
- Richards, C. S., McReynolds, W. T., Holt, S., & Sexton, T. (1976). Effects of information feedback and self-administered consequences on self-monitoring study behavior. *Journal of Counseling Psychology*, 23, 316-321.
- Romanczyk, R. G. (1974). Self-monitoring in the treatment of obesity: Parameters of reactivity. *Behavior Therapy*, 5, 531-540.
- Rozensky, R. H. (1974). The effect of timing of self-monitoring behavior on reducing cigarette consumption. *Journal of Behavior Therapy and Experimental Psychiatry*, 5, 301-303.
- Rutner, I. T., & Bugle, C. (1969). An experimental procedure for the modification of psychotic behavior. *Journal of Consulting and Clinical Psychology*, 33, 651-653.
- Schill, M. T., Kratochwill, T. R., & Elliott, S. N. (1998). Functional assessment in behavioral consultation: A treatment utility study. *School Psychology Quarterly*, 13, 116-140.
- Schloss, P. J., Thompson, C. K., Gajar, A. H., & Schloss, C. N. (1985). Influence of self-monitoring on heterosexual conversation behaviors of head trauma youth. *Applied Research in Mental Retardation*, 6, 269-282.
- Schlundt, D. G., Johnson, W. G., & Jarrell, M. P. (1985). A naturalistic functional analysis of eating behavior in bulimia and obesity. *Advances in Behaviour Research and Therapy*, 7, 149-162.
- Seymour, F. W., & Stokes, T. F. (1976). Self-recording in training girls to increase work and evoke staff praise in an institution for offenders. *Journal of Applied Behavior Analysis*, 9, 41-54.
- Shelton, J. L., & Levy, R. L. (1981). *Behavioral assignments and treatment compliance*. Champaign, IL: Thomas.
- Sieck, W. A., & McFall, R. M. (1976). Some determinants of self-monitoring effects. *Journal of Consulting and Clinical Psychology*, 44, 958-965.
- Simkins, L. (1971). The reliability of self-recorded behaviors. *Behavior Therapy*, 2, 83-87.
- Sobell, L. C., & Sobell, M. B. (1973). A self-feedback technique to monitor drinking behaviors in alcoholics. *Behaviour Research and Therapy*, 11, 237-238.
- Taylor, S. (1999). Behavioral assessment: Review and prospect. *Behaviour Research and Therapy*, 37, 475-482.
- Thoresen, C. E., & Mahoney, M. J. (1974). *Behavioral self-control*. New York: Holt, Rinehart, & Winston.
- Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying stimulus control of problem behavior. *Journal of Applied Behavioral Analysis*, 18, 343-351.
- Turkewitz, H., O'Leary, K. D., & Ironsmith, M. (1975). Generalization and maintenance of appropriate behavior through self-control. *Journal of Consulting and Clinical Psychology*, 43, 577-583.
- Watkins, M. J., & LeCompte, D. C. (1991). Inadequacy of recall as a basis for frequency knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 1161-1176.
- Williams, J. E. (1976). Self-monitoring of paranoid behavior. *Behavior Therapy*, 7, 562.
- Willis, S. E., & Nelson, R. O. (1982). The effects of valence and nature of target behavior on the accuracy and reactivity of self-monitoring. *Behavioral Assessment*, 4, 401-412.
- Wilson, G. T., & Vitousek, K. M. (1999). Self-monitoring in the assessment of eating disorders. *Psychological Assessment*, 11, 480-489.

Received November 22, 1998

Revision received August 19, 1999

Accepted August 19, 1999 ■