JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1994, 27, 413–418 NUMBER 2 (SUMMER 1994)

FUNCTIONAL ANALYSIS METHODOLOGY: SOME CLOSING COMMENTS

BRIAN A. IWATA

THE UNIVERSITY OF FLORIDA

Functional analysis methodologies have emerged as powerful tools in research on the assessment and treatment of severe behavior disorders. Although the defining feature of these approaches is the systematic identification of environmental determinants of behavior, research methodology has varied widely with respect to the arrangements used to demonstrate experimental control as well as the types of variables subject to analysis. This commentary summarizes methodological and conceptual issues that have been raised in current research and offers some suggestions for their resolution.

DESCRIPTORS: functional analysis

The research and commentaries contained in this issue of *JABA* highlight the many uses to which functional analysis methodologies have been put and suggest a number of interesting avenues for future research. Having read all of the manuscripts during the editing process, I was left with the fear that not much remained to be said. However, in our own research and during frequent discussions with students and colleagues, a number of methodological and conceptual questions often arise, several of which either were not addressed in the current papers or were exemplified in ways that may not be immediately apparent (i.e., although the question might be inferred from the data, it was not a central aspect of the study). My commentary, therefore, will focus selectively on some personal concerns about the current status of research on functional analysis methodologies.[1]

Increasing Complexity

Our initial attempts to develop an experimental approach to behavioral assessment met with a number of false starts and uninterpretable results. After over a year of pilot work, we finally settled on a few "test" conditions and a suitable control; these were designed to be maximally different so as to produce clear effects. Yet we still found that it often required a number of exposures to a given condition before response differentiation was observed, so much of our subsequent tinkering (see Iwata, Pace, et al., this issue, for details) has been aimed at facilitating discrimination on the part of subjects. By contrast, many of the complex variations evident in some current research are very difficult to understand even after several careful readings, which leads one to wonder how these arrangements ever manage to exert control over a subject's behavior. David Wacker and his colleagues (e.g., Derby et al., 1992) have been quick to point out that brief assessment yields clear results only about half the time, but they have been exceptional in describing the limitations as well as the advantages of their approach. One gets the impression in reading some other studies that most subjects' behavior is clearly differentiated on the first exposure to an assortment of complex assessment conditions.

The increasing complexity found in some studies also might be interpreted as a mandate to test for every possible event that might serve as reinforcement for a given behavior. However, our experience in all but a few cases has been that contingencies presented in a fairly straightforward manner account for much of behavior and that their influence usually is not difficult to detect. Only in the case of initial failure (followed by an attempt to deter-

I am thankful to my students, whose relentless and probing questions often served as the occasion for me to express many of the views presented here, and to Nancy Neef, whose gracious invitation to participate in the organization of this special issue served as an inducement to get my words onto paper.

Reprints may be obtained from Brian Iwata, Psychology Department, The University of Florida, Gainesville, Florida 32611.

[1] In raising some of these concerns, I have purposely refrained from citing the published sources of specific problematic examples, but I am confident that most readers will not have difficulty finding them.

mine if procedures are being implemented correctly) does it seem necessary to add further permutations of a particular assessment condition. In commenting two decades ago on the value of a much different "judgmental aid," statistical analysis, in interpreting data, Michael (1974) noted that, "Determining to what degree the judgmental aid is appropriate to the particular experiment . . . is likely to require reaction to features of the situation that are fully as complex as the features that the aid is supposed to simplify" (p. 648). I hope we can avoid reaching that point with respect to functional analysis methodologies by adopting the general practice of designing our assessments to be no more complicated than is necessary.

The Absence of Contingencies During Assessment

When attempting to determine if a response is maintained by social reinforcement (either positive or negative), it seems reasonable to expose behavior to the contingencies under consideration. In a number of studies, however, differential consequences are not delivered during assessment (baseline). Although in such cases behavior is shown to be under the control of antecedent events, the assumption that certain types of antecedent events are uniformly correlated with particular consequences might not always be correct. For example, inappropriate behavior evoked by instructions (demands) may often reflect the influence of negative reinforcement (escape), but data presented by Rortvedt and Miltenberger (this issue) and in several previous studies suggest that the instructional situation may also serve as an opportunity to receive positive reinforcement (additional attention) contingent on failure to comply with instructions.

Even if antecedent events *are* predictive of a reinforcement contingency, one would expect to find extinction if the contingency is not present. That is, if demands evoke behavior that has a history of being reinforced by escape, the continued presentation of demands without escape during assessment should result in extinction. There have been few opportunities to observe this outcome because many of the studies in which contingencies are absent also contain few data points per condition; even then, however, decreasing trends are sometimes evident in the data.[2]

The Use of Subtle Manipulations

A related concern that becomes more serious when contingencies are not manipulated during assessment is the use of extremely subtle variations in antecedent events. For example, deprivation from attention has been used in some research as an establishing operation to evoke attention-maintained behavior problems. In a few cases, the difference between the deprivation and no-deprivation conditions amounted to no more than 20 to 30 s of attention, yet large and immediate differences in behavior were observed during repeated reversals. By contrast, Hagopian, Fisher, and Legacy (this issue) showed that small (but noticeable) differences in responding were obtained given relatively large differences in the schedule of attention (10 s vs. 5 min). Although the Hagopian et al. data were collected during a comparison of treatment effects, I found them enlightening because I think they provide a good model for producing behavioral differentiation during assessment: Try to make differences among conditions large.

Alternative Assessment Strategies

Advocating simplicity in assessment may appear incongruent with favoring experimental approaches over simpler alternatives. For example, many of us have found that parents, teachers, and other caregivers sometimes can describe the functional characteristics of a client's behavior problem with uncanny accuracy. While screening subjects for a given study or conducting clinical consultation, we also have had the experience of observing an individual for only a few minutes before confidently concluding that the behavior of interest was maintained

[2] I am speaking here of extinction during baseline, which is rarely observed. It does, after all, have a way of ruining the experiment. A more serious and perhaps more common problem arises when extinction does not result in behavioral reduction during baseline due to its brevity, in which case an evaluation of treatment effects implemented following baseline is compromised even further.

by contingent attention. Very clearly then, verbal reports and sources of data other than those obtained through experimental manipulation have some value. What remains to be demonstrated is that nonexperimental methodologies have sufficient reliability and validity to be valuable *most of the time.*

In summarizing evaluative data on questionnaire methods for gathering information about behavioral function, Sturmey (in press) concluded that "Recurrent problems in this area include the lack of replicability of the reliability of these instruments." Other methods are equally lacking in the area of supportive data. For example, in spite of the relative ease with which scatter-plot data (Touchette, MacDonald, & Langer, 1985) can be gathered, I am unaware of any published replications in which the actual data were presented, or attempts made to compare results obtained through such measures with more precise indicators of behavioral function. Finally, results from several studies examining the descriptive analysis approach to behavioral assessment indicate that the procedures are no less time consuming or complicated than experimental approaches, that outcomes are not always consistent with those obtained from experimental analyses, and that correspondence could be a function of designing the descriptive analysis based on prior experimental results (Lerman & Iwata, 1993; Mace & Lalli, 1991; Sasso et al., 1992).

The above findings are not presented as an indictment against nonexperimental approaches to assessment but, rather, as a statement of the extent to which sufficient research has been conducted to establish their utility. It is not clear what types of questions would produce reliable and accurate data in an interview, whether or not scatter-plot results obtained for a significant number of individuals would reveal time-correlated distributions in behavior, and how descriptive data should be collected so as to best reveal naturally occurring contingencies. One approach to resolving these issues might consist of asking lots of questions and taking lots of scatter-plot or other descriptive data for cases in which the results of an experimental analysis are extremely clear (and I wish I had thought of this a long time ago). Given an unambiguous benchmark as a standard for comparison, other sources of data could be recombined and analyzed repeatedly to determine what combinations provide a best fit. For example, most research on descriptive analyses has been based on observational data carefully collected so as to be representative of an individual's daily routine. However, as we have found with experimental analyses, clearer results might be obtained by arranging observations under conditions selected for maximal dissimilarity.

Analysis of Antecedent Influences on Behavior

As noted in the introduction to this special issue, each of the treatment studies included one or more manipulations involving antecedent stimuli. This focus reflects a growing realization that prior events affecting behavior can be incorporated into the treatment process. To date, however, we have seen few attempts to provide a systematic account, let alone an experimental analysis, of the role of antecedent events. (Michael's work, 1982, 1993, has been an exception in providing very clear direction to our own research.) Instead, we have been comfortable in referring to such processes as "setting events" and "contextual variables." In the absence of ways to (a) describe the effects of setting in terms of behavioral mechanism and (b) establish functional relations between setting events and behavior while ruling out other potential sources of influence, we may find ourselves attributing observed correlations between antecedent events and behavior to the operation of setting influences, with little attempt at more careful analysis.[3]

An example of this problem can be found in a study recently published in *JABA*. The situation was one in which almost perfect correlations were observed between a student's late awakening (getting up in the morning more than 5 min late) and

[3] Or, if pressed for a more specific account, we might simply offer stimulus control as an explanation because it is one of the few and perhaps the most widely known mechanism of antecedent control over operant behavior. But see the examples provided by Michael (1993) showing how this interpretation may be incorrect.

her subsequent engagement in problem behavior throughout the school day. The treatment involved a contingency to promote on-time awakening, and the successful effects of intervention were attributed to this setting-event manipulation. But it is unclear how behavior was changed because, by merely altering temporally distant routines, we do not know what originally occasioned and reinforced behavior or how these local influences were modified. For example, late awakening may have initiated an endless chain of interactions involving deprivation from reinforcement (e.g., missed breakfast, loss of free time, less social interaction) or aversive stimulation (e.g., prompts to "hurry up," more work as a result of arriving late at school) that eventually occasioned behavior and resulted in some type of reinforcement. These unspecified and therefore unknown contingencies thus remained functional, and other "setting events" having the same effect as late awakening would be expected to result in recurrence of the behavior. In fact, the student's behavior problems *were* initially related to events immediately contiguous with behavior, and it would have been interesting to see if the setting event was correlated not only with the occurrence of problem behavior but also with these more immediate (and, most likely, more functional) phenomena.

Few would suggest that setting and context are irrelevant features of the environment, but many are awaiting an elucidation of these terms that is directly tied to known principles of learning (e.g., see recent discussions by Marr, 1993; Shull & Lawrence, 1993; and Staddon, 1993). The extension of functional analysis methodologies to include a thorough study of the effects of complex and temporally distant variables might thus have the dual outcome of contributing to both theory and practice.

Biological Variables

A number of biological theories for the development and maintenance of problematic behavior, particularly self-injurious behavior (SIB), have been proposed over the years and are summarized nicely in Ted Carr's commentary (this issue). Therefore, my consideration is limited to one of these theories, but I believe it applies to others as well. The endogenous opioid hypothesis, in which SIB is viewed as producing a release of endorphins that either (a) serves as reinforcement or (b) eliminates normal response suppression due to pain by elevating the pain threshold, is one of the most interesting biological accounts of SIB because it is entirely consistent with an operant analysis (Cataldo & Harris, 1982). Clinical research related to this hypothesis has focused on the opiate antagonists, naloxone and naltrexone. However, findings with both drugs have been extremely mixed (see Singh, Singh, & Ellis, 1992, for a recent review).

Although a number of explanations might account for these inconclusive results, one possibility is that the suppressive effects of opiate antagonists may be selective for a given function of SIB. If, in fact, endorphin release reinforces SIB in some individuals, the contingency is a nonsocial one (i.e., a very specific example of automatic reinforcement). Under such conditions, several outcomes might be expected, although some qualification is required. First, results of a functional analysis should often reveal lack of sensitivity to social reinforcement such as attention and escape, except in the event that the behavior also contacted these contingencies and is now maintained by them. Second, to the extent that opiate antagonists block the uptake of endorphins, behavior maintained by these substances should be suppressed due to extinction. Third, opiate antagonists should not suppress SIB maintained by attention or escape when they are evaluated on an intact baseline of reinforcement, unless the automatic punishing effects from SIB (due to an increase in the pain threshold) override the effects of reinforcement.

To my knowledge, no studies reported to date on the use of naloxone or naltrexone have incorporated data from functional analysis baselines, although it is clear that such data could greatly increase our understanding of endorphin influences on SIB. By conducting assessment prior to treatment, it might be possible to identify potential "responders" as well as "nonresponders" to drug

intervention. By continuing assessment during treatment, drug effects could be evaluated against a variety of baseline conditions.

Assessment of Low-Frequency or Cyclical Responding

Assessment and treatment become exceedingly complex undertakings when the behavior of interest occurs infrequently or according to unpredictable cycles. In such situations, it may be all but impossible to evoke behavior on many occasions, so when behavior does occur, suppression may require the use of punishment because there is no way to interfere with the unidentified reinforcement contingency, assuming there is one. This very problem was recently described by Osborne et al. (1993).

Functional analysis methodologies might be helpful in identifying the determinants of low-frequency disorders when combined with data from other sources. For example, it is possible that recurrent ailments of one sort or another (e.g., headaches) serve as establishing operations for either attention-seeking or escape behavior. When experiencing a headache, one might find social interaction in the form of comfort more reinforcing than under ordinary circumstances or escape from work-related situations a particularly desirable event.[4] Assessment during both "low" and "high" cycles of responding, when considered along with data on other aspects of the client's behavior (e.g., loss of appetite, elevated temperature, reduced engagement in typically reinforcing activities, etc.), might reveal the interaction of these influences, which not singly but in combination account for sporadic occurrences of a problem behavior. Although this situation is purely hypothetical, it is at least plausible, and the extent to which it actually occurs can be determined through the application of functional analysis methodology.

[4] Few of us persist in work while doubled over in pain, although I am struck by Nate Azrin's story about how he showed up at his lab with a mouth full of cotton and worked all day immediately after having several teeth extracted (maybe for him it was the comfort).

The issues highlighted above represent but a few of the many questions that constitute "unfinished business" for those of us involved in functional analysis approaches to assessment and treatment. Others can be found within each of the preceding studies and commentaries, and I hope these will serve as effective occasions (prompts? setting events? SDs? EOs?) for future research activity.

REFERENCES

Cataldo, M. F., & Harris, J. (1982). The biological basis for self-injury in the mentally retarded. *Analysis and Intervention in Developmental Disabilities, 2,* 21–39.

Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis, 25,* 713–721.

Lerman, D. C., & Iwata, B. A. (1993). Descriptive and experimental analyses of variables maintaining self-injurious behavior. *Journal of Applied Behavior Analysis, 26,* 293–319.

Mace, F. C., & Lalli, J. S. (1991). Linking descriptive and experimental analyses in the treatment of bizarre speech. *Journal of Applied Behavior Analysis, 24,* 553–562.

Marr, M. J. (1993). Contextualistic mechanism or mechanistic contextualism?: The straw machine as tar baby. *The Behavior Analyst, 16,* 59–65.

Michael, J. (1974). Statistical inference for individual organism research: Mixed blessing or curse? *Journal of Applied Behavior Analysis, 7,* 647–653.

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37,* 149–155.

Michael, J. (1993). Establishing operations. *The Behavior Analyst, 16,* 191–206.

Osborne, J. G., Baggs, A. W., Darvish, R., Blakelock, H., Peine, H., & Jenson, W. R. (1993). Cyclical self-injurious behavior, contingent water mist treatment, and the possibility of rapid-cycling bipolar disorder. *Journal of Behavior Therapy & Experimental Psychiatry, 23,* 325–334.

Sasso, G. M., Reimers, T. M., Cooper, L. J., Wacker, D., Berg, W., Steege, M., Kelly, L., & Allaire, A. (1992). Use of descriptive and experimental analyses to identify the functional properties of aberrant behavior in school settings. *Journal of Applied Behavior Analysis, 25,* 809–822.

Shull, R. L., & Lawrence, P. S. (1993). Is contextualism productive? *The Behavior Analyst, 16,* 241–243.

Singh, N. N., Singh, Y. N., & Ellis, C. R. (1992). Psychopharmacology of self-injury. In J. K. Luiselli, J. L.

Matson, & N. N. Singh (Eds.), *Self-injurious behavior: Analysis, assessment, and treatment* (pp. 307–351). New York: Springer-Verlag.

Staddon, J. E. R. (1993). Pepper with a pinch of psalt. *The Behavior Analyst, 16,* 245–250.

Sturmey, P. (in press). Assessing the functions of aberrant behaviors: A review of psychometric instruments. *Journal of Autism and Developmental Disorders.*

Touchette, P. E., MacDonald, R. F., & Langer, S. N. (1985). A scatter plot for identifying stimulus control of problem behavior. *Journal of Applied Behavior Analysis, 18,* 343–351.