JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1968, 1, 91-97 NUMBER 1 (SPRING, 1968)

SOME CURRENT DIMENSIONS OF APPLIED BEHAVIOR ANALYSIS[1]

DONALD M. BAER, MONTROSE M. WOLF, AND TODD R. RISLEY

THE UNIVERSITY OF KANSAS

The analysis of individual behavior is a problem in scientific demonstration, reasonably well understood (Skinner, 1953, Sec. 1), comprehensively described (Sidman, 1960), and quite thoroughly practised (*Journal of the Experimental Analysis of Behavior*, 1957 —). That analysis has been pursued in many settings over many years. Despite variable precision, elegance, and power, it has resulted in general descriptive statements of mechanisms that can produce many of the forms that individual behavior may take.

The statement of these mechanisms establishes the possibility of their application to problem behavior. A society willing to consider a technology of its own behavior apparently is likely to support that application when it deals with socially important behaviors, such as retardation, crime, mental illness, or education. Such applications have appeared in recent years. Their current number and the interest which they create apparently suffice to generate a journal for their display. That display may well lead to the widespread examination of these applications, their refinement, and eventually their replacement by better applications. Better applications, it is hoped, will lead to a better state of society, to whatever extent the behavior of its members can contribute to the goodness of a society. Since the evaluation of what is a "good" society is in itself a behavior of its members, this hope turns on itself in a philosophically interesting manner. However, it is at least a fair presumption that behavioral applications, when effective, can sometimes lead to social approval and adoption.

Behavioral applications are hardly a new phenomenon. Analytic behavioral applications, it seems, are. Analytic behavioral application is the process of applying sometimes tentative principles of behavior to the improvement[2] of specific behaviors, and simultaneously evaluating whether or not any changes noted are indeed attributable to the process of application—and if so, to what parts of that process. In short, analytic behavioral application is a self-examining, self-evaluating, discovery-oriented research procedure for studying behavior. So is all experimental behavioral research (at least, according to the usual strictures of modern graduate training). The differences are matters of emphasis and of selection.

The differences between applied and basic research are not differences between that which "discovers" and that which merely "applies" what is already known. Both endeavors ask what controls the behavior under study. Non-applied research is likely to look at any behavior, and at any variable which may conceivably relate to it. Applied research is constrained to look at variables which can be effective in improving the behavior under study. Thus it is equally a matter of research to discover that the behaviors typical of retardates can be related to oddities of their

[1]Reprints may be obtained from Donald M. Baer, Dept. of Human Development, University of Kansas, Lawrence, Kansas 66044.

[2]If a behavior is socially important, the usual behavior analysis will aim at its improvement. The social value dictating this choice is obvious. However, it can be just as illuminating to demonstrate how a behavior may be worsened, and there will arise occasions when it will be socially important to do so. Disruptive classroom behavior may serve as an example. Certainly it is a frequent plague of the educational system. A demonstration of what teacher procedures produce more of this behavior is not necessarily the reverse of a demonstration of how to promote positive study behaviors. There may be classroom situations in which the teacher cannot readily establish high rates of study, yet still could avoid high rates of disruption, if she knew what in her own procedures leads to this disruption. The demonstration which showed her that would thus have its value.

chromosome structure and to oddities of their reinforcement history. But (currently) the chromosome structure of the retardate does not lend itself to experimental manipulation in the interests of bettering that behavior, whereas his reinforcement input is always open to current re-design.

Similarly, applied research is constrained to examining behaviors which are socially important, rather than convenient for study. It also implies, very frequently, the study of those behaviors in their usual social settings, rather than in a "laboratory" setting. But a laboratory is simply a place so designed that experimental control of relevant variables is as easy as possible. Unfortunately, the usual social setting for important behaviors is rarely such a place. Consequently, the analysis of socially important behaviors becomes experimental only with difficulty. As the terms are used here, a non-experimental analysis is a contradiction in terms. Thus, analytic behavioral applications by definition achieve experimental control of the processes they contain, but since they strive for this control against formidable difficulties, they achieve it less often per study than would a laboratory-based attempt. Consequently, the rate of displaying experimental control required of behavioral applications has become correspondingly less than the standards typical of laboratory research. This is not because the applier is an easy-going, liberal, or generous fellow, but because society rarely will allow its important behaviors, in their correspondingly important settings, to be manipulated repeatedly for the merely logical comfort of a scientifically sceptical audience.

Thus, the evaluation of a study which purports to be an applied behavior analysis is somewhat different than the evaluation of a similar laboratory analysis. Obviously, the study must be *applied, behavioral,* and *analytic;* in addition, it should be *technological, conceptually systematic,* and *effective,* and it should display some generality. These terms are explored below and compared to the criteria often stated for the evaluation of behavioral research which, though analytic, is not applied.

Applied

The label applied is not determined by the research procedures used but by the interest which society shows in the problems being studied. In behavioral application, the behavior, stimuli, and/or organism under study are chosen because of their importance to man and society, rather than their importance to theory. The non-applied researcher may study eating behavior, for example, because it relates directly to metabolism, and there are hypotheses about the interaction between behavior and metabolism. The non-applied researcher also may study bar-pressing because it is a convenient response for study; easy for the subject, and simple to record and integrate with theoretically significant environmental events. By contrast, the applied researcher is likely to study eating because there are children who eat too little and adults who eat too much, and he will study eating in exactly those individuals rather than in more convenient ones. The applied researcher may also study bar-pressing if it is integrated with socially important stimuli. A program for a teaching machine may use bar-pressing behavior to indicate mastery of an arithmetic skill. It is the arithmetic stimuli which are important. (However, some future applied study could show that bar-pressing is more practical in the process of education than a pencil-writing response.[3])

In applied research, there is typically a close relationship between the behavior and stimuli under study and the subject in whom they are studied. Just as there seem to be few behaviors that are intrinsically the target of application, there are few subjects who automatically confer on their study the status of application. An investigation of visual signal detection in the retardate may have little immediate importance, but a similar study in radar-scope watchers has considerable. A study of language development in the retardate may be aimed directly at an immedi-

[3]Research may use the most convenient behaviors and stimuli available, and yet exemplify an ambition in the researcher eventually to achieve application to socially important settings. For example, a study may seek ways to give a light flash a durable conditioned reinforcing function, because the experimenter wishes to know how to enhance school children's responsiveness to approval. Nevertheless, durable bar-pressing for that light flash is no guarantee that the obvious classroom analogue will produce durable reading behavior for teacher statements of "Good!" Until the analogue has been proven sound, application has not been achieved.

ate social problem, while a similar study in the MIT sophomore may not. Enhancement of the reinforcing value of praise for the retardate alleviates an immediate deficit in his current environment, but enhancement of the reinforcing value of 400 Hz (cps) tone for the same subject probably does not. Thus, a primary question in the evaluation of applied research is: how immediately important is this behavior or these stimuli to this subject?

Behavioral

Behaviorism and pragmatism seem often to go hand in hand. Applied research is eminently pragmatic; it asks how it is possible to get an individual to do something effectively. Thus it usually studies what subjects can be brought to do rather than what they can be brought to say; unless, of course, a verbal response is the behavior of interest. Accordingly a subject's verbal description of his own non-verbal behavior usually would not be accepted as a measure of his actual behavior unless it were independently substantiated. Hence there is little applied value in the demonstration that an impotent man can be made to say that he no longer is impotent. The relevant question is not what he can say, but what he can do. Application has not been achieved until this question has been answered satisfactorily. (This assumes, of course, that the total goal of the applied researcher is not simply to get his patient-subjects to stop complaining to him. Unless society agrees that this researcher should not be bothered, it will be difficult to defend that goal as socially important.)

Since the behavior of an individual is composed of physical events, its scientific study requires their precise measurement. As a result, the problem of reliable quantification arises immediately. The problem is the same for applied research as it is for non-applied research. However, non-applied research typically will choose a response easily quantified in a reliable manner, whereas applied research rarely will have that option. As a result, the applied researcher must try harder, rather than ignore this criterion of all trustworthy research. Current applied research often shows that thoroughly reliable quantification of behavior can be achieved, even in thoroughly difficult settings. However, it also suggests that instrumented recording with its typical reliability will not always be possible. The reliable use of human beings to quantify the behavior of other human beings is an area of psychological technology long since well developed, thoroughly relevant, and very often necessary to applied behavior analysis.

A useful tactic in evaluating the behavioral attributes of a study is to ask not merely, was *behavior* changed? but also, *whose* behavior? Ordinarily it would be assumed that it was the subject's behavior which was altered; yet careful reflection may suggest that this was not necessarily the case. If humans are observing and recording the behavior under study, then any change may represent a change only in their observing and recording responses, rather than in the subject's behavior. Explicit measurement of the reliability of human observers thus becomes not merely good technique, but a prime criterion of whether the study was appropriately behavioral. (A study merely of the behavior of observers is behavioral, of course, but probably irrelevant to the researcher's goal.) Alternatively, it may be that only the experimenter's behavior has changed. It may be reported, for example, that a certain patient rarely dressed himself upon awakening, and consequently would be dressed by his attendant. The experimental technique to be applied might consist of some penalty imposed unless the patient were dressed within half an hour after awakening. Recording of an increased probability of self-dressing under these conditions might testify to the effectiveness of the penalty in changing the behavior; however, it might also testify to the fact that the patient would in fact probably dress himself within half an hour of arising, but previously was rarely left that long undressed before being clothed by his efficient attendant. (The attendant now is the penalty-imposing experimenter and therefore always gives the patient his full half-hour, in the interests of precise experimental technique, of course.) This error is an elementary one, perhaps. But it suggests that in general, when an experiment proceeds from its baseline to its first experimental phase, changes in what is measured need not always reflect the behavior of the subject.

Analytic

The analysis of a behavior, as the term is used here, requires a believable demonstra-

tion of the events that can be responsible for the occurrence or non-occurrence of that behavior. An experimenter has achieved an analysis of a behavior when he can exercise control over it. By common laboratory standards, that has meant an ability of the experimenter to turn the behavior on and off, or up and down, at will. Laboratory standards have usually made this control clear by demonstrating it repeatedly, even redundantly, over time. Applied research, as noted before, cannot often approach this arrogantly frequent clarity of being in control of important behaviors. Consequently, application, to be analytic, demonstrates control when it can, and thereby presents its audience with a problem of judgment. The problem, of course, is whether the experimenter has shown enough control, and often enough, for believability. Laboratory demonstrations, either by over-replication or an acceptable probability level derived from statistical tests of grouped data, make this judgment more implicit than explicit. As Sidman points out (1960), there is still a problem of judgment in any event, and it is probably better when explicit.

There are at least two designs commonly used to demonstrate reliable control of an important behavioral change. The first can be referred to as the "reversal" technique. Here a behavior is measured, and the measure is examined over time until its stability is clear. Then, the experimental variable is applied. The behavior continues to be measured, to see if the variable will produce a behavioral change. If it does, the experimental variable is discontinued or altered, to see if the behavioral change just brought about depends on it. If so, the behavioral change should be lost or diminished (thus the term "reversal"). The experimental variable then is applied again, to see if the behavioral change can be recovered. If it can, it is pursued further, since this is applied research and the behavioral change sought is an important one. It may be reversed briefly again, and yet again, if the setting in which the behavior takes place allows further reversals. But that setting may be a school system or a family, and continued reversals may not be allowed. They may appear in themselves to be detrimental to the subject if pursued too often. (Whether they are in fact detrimental is likely to remain an unexamined question so long as the social setting in which the behavior is studied dictates against using them repeatedly. Indeed, it may be that repeated reversals in some applications have a positive effect on the subject, possibly contributing to the discrimination of relevant stimuli involved in the problem.)

In using the reversal technique, the experimenter is attempting to show that an analysis of the behavior is at hand: that whenever he applies a certain variable, the behavior is produced, and whenever he removes this variable, the behavior is lost. Yet applied behavior analysis is exactly the kind of research which can make this technique self-defeating in time. Application typically means producing valuable behavior; valuable behavior usually meets extra-experimental reinforcement in a social setting; thus, valuable behavior, once set up, may no longer be dependent upon the experimental technique which created it. Consequently, the number of reversals possible in applied studies may be limited by the nature of the social setting in which the behavior takes place, in more ways than one.

An alternative to the reversal technique may be called the "multiple baseline" technique. This alternative may be of particular value when a behavior appears to be irreversible or when reversing the behavior is undesirable. In the multiple-baseline technique, a number of responses are identified and measured over time to provide baselines against which changes can be evaluated. With these baselines established, the experimenter then applies an experimental variable to one of the behaviors, produces a change in it, and perhaps notes little or no change in the other baselines. If so, rather than reversing the just-produced change, he instead applies the experimental variable to one of the other, as yet unchanged, responses. If it changes at that point, evidence is accruing that the experimental variable is indeed effective, and that the prior change was not simply a matter of coincidence. The variable then may be applied to still another response, and so on. The experimenter is attempting to show that he has a reliable experimental variable, in that each behavior changes maximally only when the experimental variable is applied to it.

How many reversals, or how many baselines, make for believability is a problem for

the audience. If statistical analysis is applied, the audience must then judge the suitability of the inferential statistic chosen and the propriety of these data for that test. Alternatively, the audience may inspect the data directly and relate them to past experience with similar data and similar procedures. In either case, the judgments required are highly qualitative, and rules cannot always be stated profitably. However, either of the foregoing designs gathers data in ways that exemplify the concept of replication, and replication is the essence of believability. At the least, it would seem that an approach to replication is better than no approach at all. This should be especially true for so embryonic a field as behavioral application, the very possibility of which is still occasionally denied.

The preceding discussion has been aimed at the problem of *reliability:* whether or not a certain procedure was responsible for a corresponding behavioral change. The two general procedures described hardly exhaust the possibilities. Each of them has many variations now seen in practice; and current experience suggests that many more variations are badly needed, if the technology of important behavioral change is to be consistently believable. Given some approach to reliability, there are further analyses of obvious value which can be built upon that base. For example, there is analysis in the sense of simplification and separation of component processes. Often enough, current behavioral procedures are complex, even "shotgun" in their application. When they succeed, they clearly need to be analyzed into their effective components. Thus, a teacher giving M & M's to a child may succeed in changing his behavior as planned. However, she has almost certainly confounded her attention and/or approval with each M & M. Further analysis may be approached by her use of attention alone, the effects of which can be compared to the effects of attention coupled with candies. Whether she will discontinue the M & M's, as in the reversal technique, or apply attention with M & M's to certain behaviors and attention alone to certain others, as in the multiple baseline method, is again the problem in basic reliability discussed above. Another form of analysis is parametric: a demonstration of the effectiveness of different values of some variable in changing behavior. The problem again will be to make such an analysis reliable, and, as before, that might be approached by the repeated alternate use of different values on the same behavior (reversal), or by the application of different values to different groups of responses (multiple baseline). At this stage in the development of applied behavior analysis, primary concern is usually with reliability, rather than with parametric analysis or component analysis.

Technological

"Technological" here means simply that the techniques making up a particular behavioral application are completely identified and described. In this sense, "play therapy" is not a technological description, nor is "social reinforcement". For purposes of application, all the salient ingredients of play therapy must be described as a set of contingencies between child response, therapist response, and play materials, before a statement of technique has been approached. Similarly, all the ingredients of social reinforcement must be specified (stimuli, contingency, and schedule) to qualify as a technological procedure.

The best rule of thumb for evaluating a procedure description as technological is probably to ask whether a typically trained reader could replicate that procedure well enough to produce the same results, given only a reading of the description. This is very much the same criterion applied to procedure descriptions in non-applied research, of course. It needs emphasis, apparently, in that there occasionally exists a less-than-precise stereotype of applied research. Where application is novel, and derived from principles produced through non-applied research, as in current applied behavior analysis, the reverse holds with great urgency.

Especially where the problem is application, procedural descriptions require considerable detail about all possible contingencies of procedure. It is not enough to say what is to be done when the subject makes response R_1; it is essential also whenever possible to say what is to be done if the subject makes the alternative responses, R_2, R_3, *etc.* For example, one may read that temper tantrums in children are often extinguished by closing the child in his room for the duration of the tantrums plus ten minutes. Unless that pro-

cedure description also states what should be done if the child tries to leave the room early, or kicks out the window, or smears feces on the walls, or begins to make strangling sounds, *etc.*, it is not precise technological description.

Conceptual Systems

The field of applied behavior analysis will probably advance best if the published descriptions of its procedures are not only precisely technological, but also strive for relevance to principle. To describe exactly how a preschool teacher will attend to jungle-gym climbing in a child frightened of heights is good technological description; but further to call it a social reinforcement procedure relates it to basic concepts of behavioral development. Similarly, to describe the exact sequence of color changes whereby a child is moved from a color discrimination to a form discrimination is good; to refer also to "fading" and "errorless discrimination" is better. In both cases, the total description is adequate for successful replication by the reader; and it also shows the reader how similar procedures may be derived from basic principles. This can have the effect of making a body of technology into a discipline rather than a collection of tricks. Collections of tricks historically have been difficult to expand systematically, and when they were extensive, difficult to learn and teach.

Effective

If the application of behavioral techniques does not produce large enough effects for practical value, then application has failed. Non-applied research often may be extremely valuable when it produces small but reliable effects, in that these effects testify to the operation of some variable which in itself has great theoretical importance. In application, the theoretical importance of a variable is usually not at issue. Its practical importance, specifically its power in altering behavior enough to be socially important, is the essential criterion. Thus, a study which shows that a new classroom technique can raise the grade level achievements of culturally deprived children from D– to D is not an obvious example of applied behavior analysis. That same study might conceivably revolutionize educational theory, but it clearly has not yet revolutionized education. This is of course a matter of degree: an increase in those children from D– to C might well be judged an important success by an audience which thinks that C work is a great deal different than D work, especially if C students are much less likely to become drop-outs than D students.

In evaluating whether a given application has produced enough of a behavioral change to deserve the label, a pertinent question can be, how much did that behavior need to be changed? Obviously, that is not a scientific question, but a practical one. Its answer is likely to be supplied by people who must deal with the behavior. For example, ward personnel may be able to say that a hospitalized mute schizophrenic trained to use 10 verbal labels is not much better off in self-help skills than before, but that one with 50 such labels is a great deal more effective. In this case, the opinions of ward aides may be more relevant than the opinions of psycholinguists.

Generality

A behavioral change may be said to have generality if it proves durable over time, if it appears in a wide variety of possible environments, or if it spreads to a wide variety of related behaviors. Thus, the improvement of articulation in a clinic setting will prove to have generality if it endures into the future after the clinic visits stop; if the improved articulation is heard at home, at school, and on dates; or if the articulation of all words, not just the ones treated, improves. Application means practical improvement in important behaviors; thus, the more general that application, the better, in many cases. Therapists dealing with the development of heterosexual behavior may well point out there are socially appropriate limits to its generality, once developed; such limitations to generality are usually obvious. That generality is a valuable characteristic of applied behavior analysis which should be examined explicitly apparently is not quite that obvious, and is stated here for emphasis.

That generality is not automatically accomplished whenever behavior is changed also needs occasional emphasis, especially in the evaluation of applied behavior analysis. It is sometimes assumed that application has failed when generalization does not take place in any widespread form. Such a conclusion has no generality itself. A procedure which is ef-

fective in changing behavior in one setting may perhaps be easily repeated in other settings, and thus accomplish the generalization sought. Furthermore, it may well prove the case that a given behavior change need be programmed in only a certain number of settings, one after another, perhaps, to accomplish eventually widespread generalization. A child may have 15 techniques for disrupting his parents, for example. The elimination of the most prevalent of these may still leave the remaining 14 intact and in force. The technique may still prove both valuable and fundamental, if when applied to the next four successfully, it also results in the "generalized" loss of the remaining 10. In general, generalization should be programmed, rather than expected or lamented.

Thus, in summary, an *applied* behavior analysis will make obvious the importance of the behavior changed, its quantitative characteristics, the experimental manipulations which analyze with clarity what was responsible for the change, the technologically exact description of all procedures contributing to that change, the effectiveness of those procedures in making sufficient change for value, and the generality of that change.

REFERENCES

Journal of the Experimental Analysis of Behavior. Bloomington: Society for the Experimental Analysis of Behavior, 1957-.

Sidman, Murray. *Tactics of scientific research.* New York: Basic Books, 1960.

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

Received 24 December 1967.