A School Adjustment Program (SAP) which began in 1966 is the focus of a paper which describes the program (now serving emotionally disturbed children from ages 5 to 15 in six classrooms) and compares it to the program used by Hewitt. The behavior modification procedures and individualized programming used in the classes are discussed. Research conducted with a 13-year old girl who showed a decline in reading performance is reported. In order to eliminate her bizarre answers to oral comprehension questions, the teacher and an aide implemented a series of different reaction patterns involving the amount of attention paid to right and wrong answers. Results showed that the longer the amount of attention paid, the more the frequency of the response increased (i.e., lots of attention paid to right answers, short attention to wrong resulted in a low proportion of errors and vice versa). A combination of long attention and praise for correct answers and short attention for incorrect ones produced the lowest percent of error. Caution in dispensing attention is suggested. (RJ)
The School Adjustment Program: Individualized Intervention for Children with Behavior Disorders

Robert P. Hawkins

Kalamazoo Valley Intermediate School District

The School Adjustment Program: Individualized Intervention for Children with Behavior Disorders

Robert P. Hawkins
Kalamazoo Valley Intermediate School District

Public school classrooms for children with behavior disorders ("emotionally disturbed" or severely maladjusted children) are a recent development in this country, but already there are thousands of special education classes of this nature. The general approach used in these programs ranges from the psychodynamic strategies of Bettelheim (1950) and Redl and Wineman (1952) through the neurological strategy of Cruickshank (Cruickshank, Bentzen, Ratzeburg and Tannhauser, 1961), to the behavior modification strategy (Whelan and Haring, 1966; Ross, 1967). In actual practice, most programs are probably eclectic; that is, they do not apply any one of the above strategies consistently. The present paper describes some of the characteristics of a program that consistently applies a behavior modification strategy and presents some research data from that program.

The behavior modification strategy has at least two characteristics that differentiate it from other approaches. First, there is little interest in traditional diagnosis, which emphasizes etiology and the assignment of diagnostic labels. The etiology is generally considered to be of little value because it consists mostly of past events that cannot be changed and is largely a matter of "guesswork" based on numerous unfounded assumptions. Diagnostic labels, such as "brain damaged," "weak ego boundaries," or "adjustment reaction"

1 The program and research presented here were made possible by the support of Mr. Marland E. Bluhm and Mr. Vincent Farrell, successive directors of special education, Kalamazoo Valley Intermediate School District.
are found to be of little value to those employing a behavior modification strategy because they are extremely vague, do not refer to observable behavior, and have little scientific basis (Reger, 1965; Lovitt, 1967; Werry, 1968). Even some diagnostic descriptions that refer to observable behavior such as "hyperactive" or "withdrawn," are of limited value, as will be discussed later, because they are too vague. A second characteristic of the behavior modification strategy is the orientation toward science. The behavior modifier designs and selects techniques based on scientific principles of behavior; or at least any techniques he adopts will have first been evaluated in the light of these principles. In addition, he is very likely to be conducting research to evaluate the techniques he uses and encouraging those he works with to do the same; in fact Risley (1970) has defined behavior modification as the experimental analysis of (hopefully) therapeutic techniques, regardless of the theoretical origin of the techniques.\(^2\)

Hewett (1967, 1968) has been the primary person recognized among special educators as representing the behavior modification approach in public school programs for children with behavior disorders, and his pioneering work has had a widespread effect on such programs. In his book, Hewett (1968) describes his program in considerable detail and provides a variety of research data regarding the effect of certain general aspects of the program. The program, called "the engineered classroom," is designed largely on the basis of Hewett's "developmental sequence of educational goals," a series of general goals (attention, response, order, exploratory, social, mastery, and achievement) that Hewett

\(^2\)This open-minded, empirically-oriented definition of behavior modification should appeal to those who recognize that all wisdom, ingenuity, and perceptiveness regarding human behavior does not reside within current behavior theory or within the behavior modification movement.
Hawkins considers to be descriptive of the development children normally go through in their school experience. The engineered classroom also employs a token economy with edibles and toys as backup reinforcers. Surprisingly, however, Hewett considers the engineered classroom design is "truest to the respondent model and is more concerned with maintenance of classroom success and changing negative attitudes toward learning than rigid adherence to the well-known principle of learning that rewarding maladaptive behavior may cause it to increase in frequency" (Hewett, 1968, p. 267). Hewett does not explain what respondent behaviors are being modified, but it may be fair to assume that he is referring to the fact that by providing reasonable tasks and much positive reinforcement in school, his program may change the stimuli of a school environment from aversive stimuli and elicitors of emotional responses, such as fear or anger, into reinforcing stimuli and elicitors of more positive emotional responses.

Although several authors have described aspects of behavior modification programs for children with behavior disorders (Whelan and Haring, 1966; Quay, Werry, McQueen and Sprague, 1966; O'Leary and Becker, 1967; Martin, Burkholder, Rosenthal, Thorp and Thorne, 1968; Kroth, Whelan, and Stables, 1970), Hewett is the only one who appears to have provided a fairly complete description of such a program. For this reason, Hewett's program will be used as a basis for comparison repeatedly throughout this paper in order to highlight certain characteristics of the behavior modification program to be described. This should not be interpreted as a claim that the program to be described serves children more effectively than does Hewett's engineered classroom; only objective data could provide such an evaluation.
The School Adjustment Program (SAP) was begun in 1966 with two special education classrooms. Currently, the direct service portion of the program consists of six special classrooms, a parent-training component and a prevention component that serves maladjusted children in regular classrooms. The six SAP classrooms are located in six different school buildings in different local school districts in Kalamazoo County.

The children served in the SAP classrooms currently range in age from five to fifteen. Each classroom serves an age range of three or four years and an achievement range up to approximately eight years. Each room has a teacher with a certificate to teach emotionally disturbed children, and an aide who serves as an assistant teacher.

The goal of SAP is to modify each child's behavior sufficiently that he can return to (or, in some cases, attend for the first time) a regular classroom and make an adequate academic and social adjustment there and in the other facets of his life.\(^3\) Thus, the program is a treatment program as well as an educational program in the academic sense. Probably any treatment program (at least any that claims to be innovative) should devote approximately 20% of its resources to research; and, although that particular percentage has not been achieved in SAP, the SAP staff has collected a considerable quantity of data.

Generally, a treatment program should collect at least four kinds of data: (1) data descriptive of the subjects being treated, in this case the maladjusted children being served; (2) data descriptive of the general

\(^3\)This is the goal stated by most public school programs for children with behavior disorders, but few programs even approximate the goal.
independent variable, the overall treatment program; (3) data regarding the
general effectiveness of the program in producing behavioral change; and (4)
data on effects of specific program characteristics or specific therapeutic
techniques. As the data become more adequate (comprehensive, objective, quan-
titative, socially significant, specific, etc.) the information in them becomes
more useful, and the program becomes more accountable and responsive to its own
feedback. Interestingly, only the fourth kind of data are typically obtained
by behavior modifiers, the data on effects of specific program techniques
(e.g., Quay et al., 1966; O'Leary and Becker, 1967; Hall and Broden, 1967;
Kroth et al., 1970). But if the other three kinds of data are not obtained
it will be impossible for the outside observer to determine whether the pro-
gram, as a whole, is worth imitating, and if it is, what its precise character-
istics are. It will also be difficult for the program to continue improving
beyond a particular point, because it will not have adequate feedback.

In SAP all four kinds of data are being collected. Some of these data
will be presented in describing the program and presenting one experimental
analysis conducted within the program.

Aspects of School Adjustment Program Design

Behavioral Assessment

The term "diagnosis" is not popular among behavior modifiers because of
its association with vague, mentalistic descriptions of behavior based on a
medical model, often resulting from psychological tests of highly questionable
validity, and typically having little relevance to the treatment undertaken.
However, some form of assessment obviously must take place before behavioral
objectives can be set and an appropriate intervention implemented.

The assessment used by Hewett (1968) is based on his developmental sequence.
Based on a 29-item rating form, children are diagnosed as having problems at the attention level, the response level, the order level, the exploratory level, the social level, the mastery level or the achievement level. The level at which the child has difficulty is considered to indicate the degree of structure the teacher should provide and what kinds of reinforcers will be effective with the child. For example, a child with difficulty at the response level is considered to profit best from only limited teacher structure and to be rewarded by social attention and tangible rewards, but not by task completion; whereas a child with difficulty at the exploratory level needs an emphasis on teacher structure and is rewarded by sensory stimulation, task completion, social attention and tangibles, but not by social approval.

For behavioral assessment in SAP we utilize no theoretical model, such as Hewett's (1968) developmental sequence, but instead obtain simple descriptions of relatively specific behaviors and make those our behavioral objectives or therapeutic goals for the individual child. The most relevant diagnosis is probably a list of the child's behavioral excesses and deficits that originally caused him to be referred to SAP, so we rely heavily on the referring teacher, principal, social worker and others for the behavioral objectives. At first the referring persons tended to give descriptions that were vague, incomplete, and described the child only in terms of hypothetical constructs (poor "self-image" was a common one). So we devised a "contract" to be filled out and signed by the teacher and principal before the screening meeting at which the child's admission to SAP is considered. This contract begins by saying "When the following goals have been accomplished, this child will be considered ready to return to his regular classroom again." The teacher and principal write in the behavioral objectives they wish to see achieved and sign their names at the
bottom. They are provided guidance in this by only a form letter and a few examples of behavioral objectives from previous cases, for the most part written in terms of behaviors to be strengthened rather than behaviors to be eliminated. Although we often need to obtain clarification or further detail at the screening meeting, many sets of objectives are quite adequate without further modification.4

The following is a verbatim example of the behavioral objectives provided by one regular classroom teacher referring a child to SAP:

1. Classroom Conformity
   a. David should remain in seat during work periods.
   b. Make an effort to follow directions, attempt assignments, and complete assignments regularly.
   c. Refrain from disturbing the class with pencil tapping and mouth noises (blowing on arm or giggling).
   d. Stop interrupting class with irrelevant remarks.

2. Social Attitudes and Behavior
   a. Refrain from distorting the truth.
   b. Control temper (crying, name calling, getting red in face, fighting) when children will not accept his fantastic tales.
   c. Refrain from saying things like, "No one at this school likes me".
   d. Abide by rules and display good-sportsmanship when peers make an effort to include him in games.
   e. Learn to come in from playground when bell rings and be prompt for other scheduled changes. (He likes the added attention of having someone coming to look for him.)
   f. Cease making frequent trips to the office to make unfounded reports about other children's actions.
   g. Do routine things without prompting such as (1) taking off coat and going to his seat upon entering

4Two other procedures that have proven helpful are observation of the child in his regular classroom by the prospective SAP teacher and/or one of the school psychologists working in the program (J. Eric Hayes, Michael P. O'Leary, Thomas Shikoski), and the writing of anecdotes by the referring teacher to illustrate the child's problems. Both are standard assessment procedures in SAP.
the room, (2) taking care of his other personal properties without having to be told.

h. Accept group instruction rather than requiring individual instruction.

This particular list required very little rewording, elaboration, or clarification at the screening-in meeting. One exception is that it was necessary to specify the behaviors being referred to in the last item when the teacher stated the child should "accept group instruction."

This diagnosis does not sound nearly as eloquent or insightful as saying the child has an "unresolved Oedipal conflict," "dyslexia," "cerebral dysfunction," "regressive fantasies," or "inadequate reality testing." However, such diagnostic terms may serve more as means for the clinician to obtain social reinforcement in the form of acceptance and prestige with his colleagues (and with laymen) than as useful goals for treatment, or as cues regarding what kinds of treatment techniques to apply. Such diagnostic terms also have the characteristic of sounding like explanations for the observable behavior; but, as Skinner (1953) has pointed out, they are usually fictional explanations founded on circular reasoning. The behavioral objectives given us by teachers are realistic and relevant, and they help us determine what kinds of techniques to apply.

Terms such as "hyperactive," "shy," "hostile," and "irresponsible" are accepted in the list of behavioral objectives, despite their vagueness, so long as there is further specification. Even terms that seem to refer to hypothetical constructs, such as "attitude" or "self-concept," are accepted sometimes, provided there is an accompanying list of the observable behaviors that are the basis for assigning the particular description to the child.

Though the list of specific behaviors is an adequate assessment for treatment in a SAP classroom, when treatment is conducted in the child's regular
classroom, rather than in a SAP classroom, assessment must also include the child's environment to see what factors may be producing the behavior and what resources there may be for changing or overcoming those influences. When the child is treated in a SAP classroom this environmental assessment is only necessary after we have worked with the child for a period and have been unsuccessful in modifying some particular behavior.

While the child is in SAP, his progress must be periodically re-evaluated in order to keep him progressing. His behavioral objectives become the basis for a systematic re-evaluation that occurs every three months. Often the SAP teacher and her consultant have thought of additional behavioral objectives after becoming thoroughly acquainted with the child, and these are added to the progress evaluation form (see Appendix I).

Thus far, I have not mentioned measurement. Many, perhaps most, behavior modifiers would define behavioral assessment as measuring the rate or other relevant dimension of the target behavior (e.g. Lovitt, 1967). This, however, is not the first or most important step in behavioral assessment. Obviously, to measure a behavior one must first have identified that behavior as relevant, and it is this identification aspect that I have been discussing thus far. It should be evident from the list of behavioral objectives above that recording objective data on each target behavior in the regular classroom as a standard part of the assessment process would be quite impractical. It would take a well trained observer several full days to objectively define even a few behaviors, become a familiar part of the regular classroom so that the data were not distorted by his presence, and obtain an adequate baseline.  

5The author must admit that this was not obvious to him before establishing SAP but became obvious rapidly when such measurement was attempted.
kind of expenditure of time and skill would be wasteful. Instead, our strategy has been to take precise measurement when it appears that the measurement itself will be necessary to effect the desired behavior change (or when we are interested in simply investigating the effect of an independent variable). And then the measurement is done only in the SAP classroom, not in the regular classroom before the child is placed in SAP, because most of the target behaviors of children placed in SAP are changed immediately when the child enters a SAP room; and any measurement that had been done in the regular classroom would have been wasted effort.  

Behavior Modification Programming (Treatment) in SAP

This paper will not offer any description of two treatment components of SAP: the informal parent-training groups, and the treatment of certain cases in the regular classroom rather than in the special SAP classroom. It will suffice to say that both are valuable, recent additions to SAP and both utilize the behavior modification approach. The aspect of treatment that will be briefly presented is the program provided within the largest component of SAP, the six special classrooms.

The SAP classrooms differ from one another in many specific respects; in

---

6 This is not to say that such measurement would be of no value. It would be a meaningful part of evaluating the effect of the total program as an intervention. However, it is quite common, in special education classrooms for children with behavior disorders, to observe an improvement in the behavior of most children immediately upon admission, almost regardless of the approach being used in the room. This is referred to as the "honeymoon period" because the improvement is usually temporary in most programs. In SAP it is usually permanent.

7 This number is expected to increase in the 1971-72 school year. The current teachers and their assistants -- whose dedication, enthusiasm and cooperation make a program like SAP possible -- are: Jayne Visser, Jean Musall, Joyce Roberts, Sue Crossman, Doris Mosier, Bob Mosier, Nancy Head, Daryl Coolman, Auleen Lutes, David Brown, Lynne Smith and Ann Velderas.
fact teachers are often encouraged to try some particular strategy that is not in general use in the program. However, there are many commonalities among the SAP classrooms that are worthy of description.

Individual therapeutic programming. First, and perhaps of greatest significance, is the fact that therapeutic programming is individualized. The individualization is evident in several areas. First is the familiar individualization in terms of general difficulty level of academic assignments. The profound importance of this kind of individualization needs no elaboration here; it is familiar to all special educators, psychologists, social workers and others whose particular interest is the individual child. Unfortunately, regular classroom teachers and their supervisors often fail to provide this basic kind of individualization and thus exacerbate or even create some of the severe adjustment problems that lead to a child's placement in a program like SAP. Also unfortunately, many special education programs for maladjusted children "soft pedal" the child's academic instruction on the assumption that he needs relief from the pressure of academic work. Then the child falls farther behind academically and will probably have greater difficulty adjusting to a regular classroom. Our approach is to teach the child to adjust to the demands of academic work; and we cannot do this in a situation which does not make such demands.

The second kind of individualization is the setting of reinforcement criteria within an academic subject to solve specific academic problems. For one child, the teacher may be emphasizing the conditioning of a higher rate of reading, for another she may be emphasizing accuracy, and for yet another she may be emphasizing comprehension.

The third kind of individualization also relates directly to academic performance; it is the individualized programming for Requisite Antecedent
Behaviors, or RABs, as Stanley Sapon at the University of Rochester has called them (Reading Newsreport, 1969). These are the behaviors that Hewett (1968) would include under the headings of "attention level," "response level" and "order level" problems. They are behaviors necessary if the child is to be taught academic skills in a group setting. We make certain adjustments to produce these RABs in children who do not consistently exhibit them, but we do not find it advantageous, as does Hewett (1968), to minimize structure for children with "attention" and "response" problems; nor do we find it necessary (with rare exception) to greatly reduce our expectations of such children, such as by substituting a variation of the "shell game" for academic tasks in order to teach attentive behavior or allowing a child to draw maps all day because that is the only academic activity he will willingly participate in. Generally we find that we can produce consistent attentiveness, responsiveness to tasks, cooperativeness, and obedience to classroom rules by simply providing a combination of consequences for these behaviors, and for their opposites, that provides strong motivation for the appropriate behavior. That is, it appears that nearly every child served by SAP has the appropriate RABs in his repertoire, but for many of the children adequate motivation has not been provided in the regular classroom to produce consistently appropriate performance. The degree of individualization that we do find necessary is primarily in the magnitude and frequency of reinforcement or other conseuation for RABs.

The fourth kind of individualization, and perhaps the kind which most differentiates SAP from the Engineered Classroom (Hewett, 1968) is the individualized programming for non-academic behaviors. It should be evident from the list of behavioral objectives presented earlier that not all of the adjustment problems observed in an individual child are academic behaviors or even RABs.
A variety of social and personal behavioral deviations are seen in many SAP children that appear to severely limit the reinforcers available to them currently and/or in their future. The following behavioral goals, extracted from several different cases, illustrate such problems.

Tina should show emotion, when appropriate. She should laugh, get angry, show disappointment, show excitement, etc.

Michael should not use crude, profane language at inappropriate times and places.

Richard should seek social contact on his own initiative.

Gary should learn how to get more favorable social reactions from peers rather than always being the target of teasing. (Here the actual behavior was not well specified, because it was difficult to specify what skills the child lacked, without extensive observation).

Mike should look at persons to whom he is talking.

Leonard should stop emitting bizarre behaviors such as slapping his face, making odd vocalizations, sniffing people and objects (especially neckties), rolling on the floor at inappropriate times, giggling inappropriately, spinning around, etc.

Ron should learn appropriate grooming habits, especially cutting his extremely long fingernails and washing his hands regularly.

Mike should stop stealing.

It appears that Hewett does not deal systematically with such problems, but rather assumes that the program provided by the Engineered Classroom will weaken undesirable emotional respondents and perhaps through this process eliminate the operants that obstruct the child's adjustment. By contrast, in SAP we do not believe it is generally fruitful to assume that the basic problem is in the child's emotional responses (which is why we prefer not to refer to the children as "emotionally disturbed" but rather as maladjusted or having a behavior disorder) and that these responses must be modified before the child's adjustment can improve. Instead, we generally assume that the child's
maladjustment is in the operant behavior that we can readily observe and specify.

In cases where we do suspect that a child is inordinately fearful, or has some other undesirable, emotional reaction, we believe that the problem can still be solved most readily by directly conditioning the operant behavior upon which such a conclusion is largely based. For example, with a child who appears to be afraid (respondent behavior) to assert himself, we might prompt successively more assertive responses (operant behavior), reinforce them, and then gradually fade the prompts. Typically, the judgement that fear is involved at all is gratuitous; the behavior observed is simply a low frequency of assertive responses. Thus the most parsimonious—and, we feel, the most effective—approach is to simply obtain assertive responses and reinforce them.

Contingent attention and praise. One of the primary generalizations we teach our SAP teachers is that attention is an almost universal reinforcer and that a teacher must be discriminating in the dispensing of that reinforcer. There is considerable applied research literature supporting this generalization (Harris, Wolf and Baer, 1964; Hart, Allen, Buell, Harris, and Wolf, 1964; Harris, Johnston, Kelly, and Wolf, 1964; Allen, Hart, Buell, Harris, and Wolf, 1964; Hall and Broden, 1967; Hall, Lund and Jackson, 1968; Madsen, Becker, Thomas, Koser and Plager, 1968; Schutte and Hopkins, 1970). Therefore SAP teachers generally avoid looking at or talking to a child emitting an undesirable behavior such as inappropriate complaining about physical ailments, whining about the difficulty of assigned work, leaving the seat without permission, daydreaming, or various "clowning" behaviors. Because reprimanding a child verbally involves giving him attention and thus possibly reinforcing a response one wishes instead to weaken, the SAP teachers can very rarely be heard to reprimand a child. If the teacher wishes to rapidly weaken a response, she
is likely to reinforce an incompatible response or apply some non-verbal punisher.

On the other hand, the teachers use a great deal of praise. This year we have had an observer record the frequency of teacher-dispensed verbal approval, so that we could give a quantitative description of this important characteristic of the program. Verbal approval was defined in such a manner as to minimize the degree of subjective judgement required of the observer (see Appendix II). In so defining it we found that we were excluding a large number and variety of social reinforcers given by the teacher, perhaps as many as half of them in some of the classrooms; but it still provided an objective measure of a significant portion of the verbal reinforcers dispensed. The kinds of verbalizations that were excluded were primarily ones that have multiple meanings, such as "okay" and "all right," which are often used as cues to get the classes' attention but are also used extensively by some teachers as a form of verbal approval.

Data were recorded for 47 different sessions ranging in length from 20 minutes to 60 minutes. A total of more than 34 hours of data were collected. Data were recorded on both the teacher and aide (and student teacher; if one were present) together, so that the data represent the frequency with which verbal approval occurred audibly in the room, without regard to which teacher gave it or which child received it. The recording took place during various academic activities within the classroom. The teachers were told only that the observer could not tell them what behavior she was recording because that might cause the teacher to modify that behavior; thus there was an implication that a child's behavior was being recorded. The SAP teachers are accustomed to having visitors in the room, therefore it seems reasonable to assume that the data are representative of the rate of dispensing verbal approval when observers
are not present.

Inter-observer-reliability was checked on five occasions by having a second observer record the same data independent of the first. Independence of the data was assured by keeping the observers far enough apart and oriented in directions sufficiently different that the observers responded to the teachers' behavior and not to each others' behavior. The total number of occurrences of verbal approval recorded by one observer was divided into the total recorded by the other, always dividing the larger into the smaller. When multiplied by 100, this yielded a percentage of agreement. Over the five reliability checks, these scores ranged from 79% to 98%; and the mean was 91%.

The lowest rate obtained on any single session was 8 approvals per hour, and the highest was 185. The mean rate of verbal approval for each room was obtained based on as few as 3 and as many as 24 different sessions of observation, and these means ranged from 39.3 approvals per hour to 94.3 approvals per hour. The mean of these means was 60.0 approvals per hour, or exactly one per minute (remarkably). Thus it can be fairly stated that SAP teachers dispense frequent praise or verbal approval. Perhaps it should also be said that the teachers try to develop friendly, accepting relationships with the children, not only because we think such relationships are themselves important for the child's development, but also because the teacher's approval is more likely to serve as a reinforcer and her disapproval as a punisher when such relationships exist.

No data were obtained from one classroom because it was new and the total enrollment was only six, which meant that attendance at any particular time of day was usually below six due to the fact that some children were attending regular classrooms part-time as a part of their phasing out.
Contingent token reinforcers. A token economy is employed in every SAP classroom, with poker chips serving as the tokens. In Hewett's (1968) Engineered Classroom the tokens are checks on a "work record card" and are given on a fixed interval basis at the end of every 15 minute work period. This is also the time when academic work is checked for accuracy. Check marks are given for starting work, continuing to work, and for "being a student," which means a variety of academic behaviors and RABs, depending on the problems of the child. In SAP we give tokens immediately, while the behavior is occurring, as well as at fixed intervals, because we conceive of the tokens primarily as reinforcers for desirable operant behavior rather than as positive events that reduce the child's negative emotional reactions to school and learning (though they may do this as well). In addition, we dispense tokens for a much wider variety of behaviors, depending on our behavioral objectives for the child and for the group. For example, tokens have been used to reward a profoundly withdrawn child for knocking the teacher's books off her desk (an assignment given to teach greater assertiveness), to reward another isolate child for initiating social interaction (Mosier, 1970), to reward a complainer for beginning a task without complaint, to reward a boy who had stolen tokens for going a prescribed period of time (we began with a 30 minute criterion) without stealing, to reward a very shy boy for requesting help rather than crying when he had

---

9The research conducted by Hewett (1968) to evaluate the effectiveness of the Engineered Classroom was essentially an evaluation of the combined effects of employing a token economy and relying consistently on Hewett's Engineered Classroom design (including the developmental sequence, the series of "interventions," the physical arrangement of the room, and other aspects). The results of this research are mixed, but suggest that the children are more attentive while under the token economy. On the other hand, the achievement data show very few effects favoring the token economy and Engineered Classroom conditions.
difficulty on an academic task, and to reward many children for such behaviors as reading expressively, working quietly, reading a line without error, participating in discussion, or ignoring a classmate who is being noisy.

Tokens are used liberally and are often paired with praise, pats on the back, or other social reinforcers. The tokens are dropped in small containers on the children's desks. This produces a sound that makes it unnecessary for the child to even look up momentarily to detect the receipt of a token, and we find it is not disruptive to the child after his first hour or two in the classroom. This year we have measured the rate of token dispensing, just as we measured the rate of verbal approval. The teachers sometimes dispense several tokens at once; but because we did not wish to ask them to tell us how many they were dispensing and thus reveal what we were recording, and because the observer often could not tell how many were dispensed, we actually recorded the frequency of the teacher's token-dispensing responses rather than the precise number of tokens dispensed. Reliability was checked by an independent, second observer on one occasion and was found to be 87%. Over 16 sessions the mean frequency of token dispensing in a classroom ranged from 32.6 tokens per hour to 102.0 per hour, with a mean of 62.1.\(^{10}\)

**Contingent token withdrawal.** On the same sessions in which token dispensing was measured, token withdrawal was also measured. This consequence is used occasionally to weaken an inappropriate response. Reliability of these data was checked on one of the 16 sessions and was found to be 100%. The frequency of token withdrawal in a classroom ranged from zero per hour to 8.8 per hour with a mean of 2.0. This exemplifies the general emphasis on reinforcement rather than punishment in SAP.

\(^{10}\)Again, data were recorded in only five of the six classrooms because of the small enrollment in the sixth.
Contingent time-out. Another consequence used as a punisher is time-out. In SAP this consequence is applied for one of four reasons: the response is potentially dangerous to someone in the group, the response is extremely disruptive, the response is likely to receive or is receiving social reinforcement from peers, or the response has been very resistant to change by reinforcement procedures alone. Each SAP classroom has a 4' X 4' booth in which the teacher can place a child immediately contingent on a response the teacher considers important to eliminate. The child is usually left in the booth approximately five to ten minutes and then allowed out if he is being quiet. If the child has never been in the booth before, or if he has never tried to leave the booth without permission, the door may be left unlocked and ajar. After the time-out is over, the teacher records the incident on a special record sheet and indicates the total time spent in the booth. This year we took certain data from these sheets for the two SAP classrooms where time-out appears to be used most frequently (one teacher has not used it once in four months). The following table presents the data for six months from one classroom and seven months from the other.

Whelan and Haring (1966) were apparently the first to describe the use of a time-out room for emotionally disturbed children in a school setting.
In both classes the most frequent reason for time-out is disruptive behavior (usually mild), and the second most frequent reason is refusal to work (including "daydreaming").

At times when the child's behavior is such that a SAP teacher would probably apply the one consequence of time-out in an empty booth, Hewett (1968) utilizes a series of nine consequences that he calls "interventions," which the teacher can either use in sequence or select from, until the child's behavior improves. The first is to move the child, in a positive manner, to a study booth that consists of three walls partially enclosing a table or desk. The child continues his assigned work there. For many children, this probably functions as a punisher and thus effectively weakens whatever behavior preceded it. The second intervention is to give the child an easier, or at least different, assignment. When viewed in terms of operant principles, this technique appears
to run the risk of getting the child and teacher into a vicious circle of reinforcing each other's inappropriate behavior. It may be similar to the teacher-pupil interaction pattern in which a teacher persists in telling a disruptive child to sit down because the child does obey her for the moment, thus reinforcing the teacher's saying "Sit down," but at the same time the teacher's attention to the child while out of his seat reinforces his out of seat behavior (see Madsen, et al., 1968). In Hewett's intervention the disruptive or other problem behavior of the child seems likely to be reinforced by removal of a difficult assignment (and by the attention of the teacher); but when given an easy assignment he is likely to behave in a manner less aversive to the teacher, so the teacher's behavior of applying this intervention is also reinforced. Thus the child reinforces teacher behavior that is, in the long run, detrimental to him, and the teacher reinforces child behavior that is, in the long run, detrimental to her. Of the other seven interventions—verbal restructuring, sending the child to the exploratory center, sending the child to the order center, taking the child outside to perform a physical activity task, providing individual tutoring and increasing the number of tokens the child may earn, isolating the child in the principal's office and telling him that he may not earn tokens for a specified few minutes, and excluding the child from school—the first five appear to run the same kind of risk as did intervention number two. In SAP we avoid interventions that appear to run such a risk unless there is good experimental evidence that they are effective.

Home consequation of school behavior. The last general SAP technique I will describe is based on research we conducted with seven underachievers in regular classrooms (Hawkins and Sluyter, 1970). We found that parents could readily modify significant school behaviors by providing consequences at home
based on daily feedback from school. McKenzie, Clark, Wolf, Kothera and Benson (1968); Cantrell, Cantrell, Huddleston and Woolridge (1969); Kroth et al., (1970); and Bailey, Wolf, and Phillips (1970) have obtained similar results. Therefore, all SAP teachers send home daily reports on certain classroom performance and encourage the parents to provide appropriate reinforcers. In some cases very special report sheets are printed in order to solve some difficult problem of a particular child.

Before presenting one of the experiments we have conducted within a SAP classroom, I should point out that this paper will not describe the process of gradually phasing a child out of the special classroom, placement into a selected regular classroom, or subsequent follow-up. Nor will the initial pre-test, post-test, follow-up-test data be presented here. For the present, perhaps it will suffice to say that approximately 35 children have been placed in regular classrooms from SAP during the five years of its existence, and only two of these children have been referred back to the program. In both cases the screening out procedures had not been followed properly.

Experimental Research in the School Adjustment Program

Several studies have been conducted as a part of the School Adjustment Program (Hawkins, 1967; Schwarz and Hawkins, 1970; Hawkins and Sluyter, 1970; Shikoski, 1971; Tough, Hawkins, McArthur and Van Ravenswaay, in press) and a few have been conducted within SAP classrooms (Hawkins, McArthur, Rinaldi, Gray and Schaftenaar, 1967; Mosier, 1970). The study to be presented here was also conducted within a SAP classroom. Its primary purpose was not therapeutic, though therapeutic change was an intended outcome. The primary purpose was investigation of a variable that could have considerable significance for the teaching of children with various learning problems.
The experiment was prompted by a decline in the reading performance of one SAP child. The teacher noticed that the child had begun making numerous errors in answering the oral comprehension questions that were asked daily after she had completed reading an assigned story. The teacher and the aide, a handsome young man who was new in the classroom, alternated at working with the child, and both had observed that she not only gave many incorrect answers but often gave answers that seemed bizarre because of their unrelatedness to the story or the question.

In observing the aide teaching the subject it was noted that he was employing a pattern of teaching behavior that is very common for a teacher trying to give extra help and consideration to an individual child. When the child gave a correct answer he praised her genuinely and then asked the next question; but when she gave an incorrect answer he patiently explained to her certain relevant aspects of the story and indicated why her answer was therefore incorrect. His response to an error was of much greater duration than his response to a correct answer. Because there is good evidence that attention itself can serve as a reinforcer, (and because we knew that for this child attention was a particularly powerful reinforcer) we suspected that the aide was reinforcing wrong answers by his long attention more than he was reinforcing right answers by his praise and brief attention. If this were true it would have important implications for the technology of teaching, especially of teaching exceptional children. We decided, therefore, to conduct an experiment to investigate this hypothesis. 15

15 This research was conducted by J. Eric Hayes and the present author, with the able assistance of Doris B. Mosier and Joseph J. Vaai, Jr. The cooperation and enthusiastic support of Larry Lindeman, Principal, was also essential.
Method

Subject

The subject was a 13 year old girl in one of the SAP classrooms. She had been referred to the program when in third grade because of a history of bizarre behavior, hyperactivity, and manipulativeness since entering school, and unmanageability at both school and home. Much of her bizarre behavior related to physical affection; she would try to touch, stroke and kiss boys and even strange men. She also made grotesque faces, laughed at inappropriate times and generally did and said a variety of strange things that resulted in active rejection by her peers. She had a history of seizures, some of which were suspected of being "faked." In approximately 11 months in SAP, the subject had learned to be calm and orderly, work diligently, complete assignments, display no inappropriate degree of affection, and have virtually no seizures. However, she still sought adult approval excessively, was sometimes manipulative, and had not made any friends (though she was no longer disliked).

Situation

Daily experimental sessions were conducted during the subject's reading period, and started when she indicated that she had read the assigned story and was ready to answer comprehension questions. The sessions were approximately five to ten minutes in length. The teacher and aide alternated as instructors and asked the subject a series of approximately 15 prepared, objective questions about the story. The instructors were asked to keep the difficulty of the questions comparable throughout the study and appeared to do so.

Data

The instructor also served as data recorder. For each question he recorded whether the answer was right or wrong and the duration of his own verbal
response to the answer. The latter was measured by starting a stopwatch, mounted on a clipboard, when he started talking about the answer and stopping it when he finished. After the time was recorded, he reset the watch and asked the next question. From these records was calculated the percentage of errors made by the subject, the mean duration of teacher response to correct answers and the mean duration of response to incorrect answers. Thus both the dependent variable and the independent variable were measured in each session. One of the experimenters was frequently present to assure that prescribed experimental and data recording procedures were followed.

In order to determine objectivity of data recording, inter-observer reliability was measured. On ten occasions, at least once in every phase of the experiment, a second observer recorded the same data as did the instructor and in the same manner. Inter-observer agreement on the correctness data was then determined by dividing the percentage of errors obtained by one observer by the percentage of errors obtained by the other observer, always dividing the smaller by the larger. This was then multiplied by 100, which yielded a reliability index called "percent agreement." The same method was used to calculate a reliability index for the mean duration of teacher response to errors and the mean duration of response to correct answers. Reliability ranged from 84% agreement to 100% with a mean of 94%.

Procedure

The experiment was conducted in six phases, the first four constituting an ABAB experimental design. In all, four different patterns of teacher reaction were implemented as independent variables.

Phase I. The first phase consisted of implementing consistently the teaching pattern of the aide that was suspected of having produced the high percentage of errors from the child. This strategy appeared to offer more possibility
Hawkins

of yielding understandable results than the usual procedure of measuring performance under relatively unspecified, unquantified and variable baseline conditions. The instructors consistently gave longer responses to the subject's wrong answers to comprehension questions than to her right answers. This teaching pattern was called Teaching Condition 1.

Phase II. After 41 sessions of Teaching Condition 1, Teaching Condition 2 was implemented. This condition was the reverse of Condition 1; the instructors consistently responded longer to correct answers than to errors. This was found to be a somewhat difficult and unnatural behavior pattern to emit, initially; it is easier to point out why an answer is wrong than why it is right.

Phase III. On session 72 Teaching Condition 1 was again implemented. The instructors responded longer to errors than to correct answers.

Phase IV. On session 90 Teaching Condition 2 was implemented again. The instructors responded longer to correct answers than to errors.

Phase V. Up to this point in the study the teaching pattern always involved differential response to right and wrong answers. On session 101 the instructors began responding approximately the same length of time to right answers as to wrong ones. They gave a long response to either kind of answer. This was called Teaching Condition 3.

Phase VI. On session 115 Teaching Condition 4 was implemented. Like Condition 3, this consisted of equal duration of teacher response to correct and incorrect answers; but in Condition 4 both durations were short. This condition was in effect for 10 sessions.

Results and Discussion

Data on both the behavior of the instructors and the behavior of the child are presented in Figure 1. During Phase I the instructors' mean duration of
Figure Caption

Fig. 1 Percent answers incorrect (heavy line) as a function of differential length of teacher response to correct (dashed line) and incorrect answers (solid, thin line). Note that ordinate represents both percent and time (in seconds).
response to errors ranged from 8 seconds to 34 seconds, with an overall mean (of these daily means) of 16 seconds, while the mean response duration to correct answers ranged from 3 seconds to 13 seconds, with an overall mean of five seconds.\textsuperscript{16} The child's answers ranged from 30\% incorrect to 82\% incorrect, with a mean of 55\%.

In Phase II the instructors' responses to correct answers averaged nearly the same as had their responses to incorrect answers in the previous condition (mean, 14 sec.), but showed less variability between daily means. Likewise, their mean response to incorrect answers in Phase II nearly matched their response to correct answers in Phase I (mean 4 sec.). Thus, in terms of duration, the instructors' responses in Teaching Condition 2 are nearly a perfect reversal of their responses in Condition 1. The child's percentage of errors gradually declined over Phase II, reaching a mean of only 14\% on the last three sessions.

In Phase III the instructors' overall mean duration of response to errors was 16, exactly as in Phase I. Their responses to correct answers averaged three seconds during Phase II, somewhat less that in Phase I. The child's errors immediately increased under these conditions, finally reaching a level comparable to her performance in Phase I. She averaged 51\% errors over Phase III.

The instructors' responses in Phase IV replicated those of Phase II very well. The overall mean response to correct answers was 15 seconds, and the overall mean response to errors was four seconds. The child's errors declined immediately and averaged 24\% over Phase IV. This, then, replicated the results obtained in Phases I and II. The differential duration of teacher response to

\textsuperscript{16}On session 12 the instructor, through a misunderstanding, accidentally reversed conditions, which may account for the change in the subject's performance on that day.
errors and correct answers controlled the correctness of the child's performance within a rather wide range. When long responses followed errors, and short responses followed correct answers, the proportion of errors was high. When long responses followed correct answers, and short responses followed errors, the proportion of errors was low. This confirmed our original hypothesis regarding the cause of the subject's poor daily performance in reading and thus guided us in designing a solution to her reading problem.

However, we were then curious whether the child's improvement under Teaching Condition 2 was a result of the decreased duration of attention following errors or the increased duration of attention following correct responses. We implemented Teaching Condition 3 to determine whether long responses to both kinds of answers would also produce a low proportion of errors. In Phase V the instructors' responses to errors and correct responses averaged 16 and 15 seconds respectively. The child's answers averaged 41% during this condition, higher than during Teaching Condition 2, but lower than during Condition 1. This suggests that the high percentage of errors under Condition 1 could be attributed only partially to the long duration of attention following errors.

In Phase VI the instructors' responses were all of short duration, averaging only three seconds for both right and wrong answers. Under this condition errors reached nearly as high proportions as during Teaching Condition 1, 48%. These results combined with those of Phase V suggest that both a long response to errors and a short response to correct answers contributed to the high percentage of errors in Teaching Condition 1, and the combination of the two tends to produce a consistently high percentage of errors (Phases I and III).

It appears that all of the results can be accounted for by the hypothesis that for this particular child adult attention is a powerful reinforcer, but
mild praise for correct responses (or simply feedback that an answer was correct) has only a slight reinforcing value (or at least a slightly more reinforcing value than being told her answer was wrong). Under such a hypothesis, Teaching Condition 2 should be most conducive to a low percentage of errors, for correct answers would be receiving both kinds of reinforcers, both the praise and the greatest amount of attention. Such was the case. Teaching Condition 4 should produce the next highest percentage of errors because the attention reinforcement for right and wrong answers was equal and should therefore tend to yield 50% wrong answers, but the praise should draw this percentage down somewhat. This was the case; errors were slightly under 50%. Teaching Condition 3 should produce a slightly lower percentage of errors, because in responding to correct answers for a long duration, the instructors often repeated their praise more than once. For example, one long duration response to a correct answer, taken from a tape recording of a session, was "Very good Jane. It's good that you knew Hubert didn't go to school and that you knew why he didn't go to school. I'm glad that you are able to remember these details." The repetition of praise should reinforce correct answers more than one simple occurrence of praise and thus produce a lower percentage of errors under Condition 3 than under Conditions 1 or 4.

One additional experimental result should be reported. Comparison was made between the subject's performance for the teacher, a woman, and the aide, a young man. This was of interest because before admission to SAP, and for a period after her enrollment, the subject had appeared to be particularly eager to obtain male attention and affection.

Comparisons were made between the child's performance for the aide and her performance for the teacher in Phases I and II. It was anticipated that
the aide would produce the highest percentage of errors in Phase I and the lowest in Phase II. Neither prediction was correct. In Phase I the subject made a mean of 55% errors when the aide acted as instructor and 53% errors when the teacher acted as instructor. In Phase II the subject made a mean of 31% errors for the aide and 32% for the teacher. These differences are too small to be of particular interest.

Following the end of the experiment the aide was quite convinced that it would be wise to praise the child heartily and comment at length about correct answers but to give only brief feedback for errors. The child's reading comprehension, as judged by her answers to comprehension questions, improved greatly. In addition we learned that some very natural methods for trying to help a child can be detrimental to the child instead. Allen and Harris (1966) had found the same phenomenon in the case of a child who scratched herself excessively. Harris, Johnston, Kelly and Wolf (1966) had found it in the case of regressed crawling in a nursery school child. Allen et al., (1964) had found it in the case of an isolate child. Hart et al., (1964) had found it in the case of a child who cried excessively. And Hawkins, Peterson, Shweid and Bijou (1966) found it in the case of a tantruming child (as had Williams, 1965).

The next question, then, has to do with the generality of the phenomenon. One brief replication of the first two conditions of the experiment was conducted with a boy in the same SAP classroom, and the results were similar. We are now in Phase III with a third child in that classroom and have obtained no effect resembling that reported here. Only continued research will answer the question of generality, but the primary fact remains that any teacher -- particularly a teacher of handicapped children (or, probably, a counselor) -- must be cautious in the dispensing of her attention; regardless of her good
Hawkins

intentions she can actually be hurting the child she intends to help. A thorough awareness of behavioral principles can help to prevent such unfortunate effects.
REFERENCES


Mosier, Doris B. Operant conditioning techniques applied to low rate self-initiated conversation and eye-contact. School Applications of Learning Theory, 1970, 2, 49-54. This informal journal for teachers is available from Kalamazoo Valley Intermediate School District, Box 2025, Kalamazoo, Michigan 49003.

Hawkins


Reading Newsreport article, World without words, 1969, 3, 26-31.


APPENDIX I

Goal Attainment Evaluation System (GAES)

The form shown here was developed in the School Adjustment Program of Kalamazoo Valley Intermediate School District to periodically evaluate a child's adjustment progress in terms of the individualized behavioral objectives set for him. The objectives, listed in the extreme left column of the form, are written before and during the screening-in process, largely by the persons referring the child to SAP. Additions are occasionally made by the SAP teacher as she comes to know the child better.

Evaluation every three months is in terms of the teacher's subjective estimate of the current frequency of the behavior listed as an objective. The purpose of the evaluation is to accelerate the treatment process by causing the teacher to systematically review the child's progress in terms of relatively specific, behavioral objectives and to revise her techniques when objectives are not being achieved. In addition, this evaluation process often prompts the SAP teacher to begin the phasing out of a child whose goals have been reached.
GOAL ATTAINMENT EVALUATION SYSTEM (GAES) FOR EVALUATING NON-ACADEMIC PROGRESS FOR Child

Write in your name and the date, then fill in one column using one of the following letters to indicate the current status of the behavior: (C) consistently true, (U) usually true, (S) seldom true, (N) never true, or (Q) questionable. Then add any notes to explain your evaluation (esp. Q) if needed. Then add any new programming decisions to the "Behavioral Prescriptions" sheet, date it, and attach to this sheet.

<table>
<thead>
<tr>
<th>Behavioral Goals set at screening-in and supplemented by yourself.</th>
<th>Evaluation by:</th>
<th>Date:</th>
<th>Evaluation by:</th>
<th>Date:</th>
<th>Evaluation by:</th>
<th>Date:</th>
<th>Evaluation by:</th>
<th>Date:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX II

Definition of approval:

Verbal approval consists of verbal (vocal) comments indicating approval, commendation of correctness, such as "Good," "Fine job," "You're studying well," "I like to see you...," "Right," or "Thank you." The words themselves (not accompanying gestures, expressions or emphasis) must convey approval clearly enough that if seen in print, the reader would judge them to be a statement of verbal approval.

Comments which, if seen in print, would not clearly indicate verbal approval should not be counted. Thus "I see you remembered to put your name on your paper" is not verbal approval, because the statement does not indicate whether the child's having his name on the paper is desirable or not. If "That's great" followed it, one occurrence of verbal approval would be scored.

Comments which require a knowledge of other things the teacher or child have said in order to be identifiable as verbal approval should not be counted. Thus a child may ask "Is this answer right?" and the teacher may reply "Yes," but no verbal approval can be scored because the word "Yes," seen in print, does not by itself indicate verbal approval. By the same token "Okay" and "Allright" are not verbal approval.

Episodes of giving approval may contain multiple instances of verbal approval. Each instance should be scored. Thus if a teacher said "Good job, Sue. I like the way your paper looks," two instances of verbal approval are scored, even though the approval may be referring to the same performance of the child. The words "good" and "I like" are the cues here, for they indicate approval.