DOCUMENT RESUME

ED 240 765 EC 161 928

AUTHOR Evans, Ian M.; Voeltz, Luanna Meyer
TITLE The Selection of Intervention Priorities in Educational Programming of Severely Handicapped Preschool Children with Multiple Behavioral Problems. Final Report.
INSTITUTION Hawaii Univ., Honolulu.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC. Div. of Innovation and Development.
PUB DATE Dec 82
GRANT G007901960
NOTE 190p.; The document was developed through the Departments of Psychology and Special Education.
PUB TYPE Reports - Research/Technical (143) -- Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS *Behavior Modification; *Behavior Problems; Decision Making; *Intervention; *Models; Preschool Education; *Severe Disabilities

ABSTRACT

The report summarizes findings from an investigation of response interrelationships in severely handicapped children with behavior problems in order to provide empirical guidelines for determining intervention priorities. A state-wide sample of severely handicapped, behavior problem children (2-8 years old) was observed longitudinally in their public school program. Real-time observational techniques using microprocessors recorded percent duration of over 95 inappropriate behaviors as well as teacher and environmental variables. Excess behaviors showed small decreases over 3 years, as contrasted with previous epidemiological findings for institutionalized children. For a subsample of the Ss individual interventions were undertaken to determine whether response patterns could be identified. Although reliable clusters proved difficult to identify, both the child studies and investigations of teachers' decision making suggested that programing based on the acquisition of functional skills was the most valuable strategy for reducing excess behaviors. A model was developed to help teachers and other clinicians choose interventions. (Author/CL)

Final Report

Project No. 443CH 90024
Grant No. G00-790-1960

Ian M. Evans and Luanna Meyer Voeltz
Departments of Psychology and Special Education
University of Hawaii
2430 Campus Road
Honolulu, Hawaii 96822

THE SELECTION OF INTERVENTION PRIORITIES IN EDUCATIONAL PROGRAMMING OF SEVERELY HANDICAPPED PRESCHOOL CHILDREN WITH MULTIPLE BEHAVIORAL PROBLEMS

December 1982

U.S. DEPARTMENT OF EDUCATION

Special Education Programs

Division of Innovation and Development
(Field-Initiated Research Program)

Abstract

The project's purpose was to investigate response interrelationships in severely handicapped children exhibiting behavior problems, in order to provide empirical guidelines for determining intervention priorities. A statewide sample of such children, aged from two to eight years, was observed longitudinally in their public school educational programs. Real-time observational techniques employing microprocessors allowed recording of percent duration of over 95 inappropriate behaviors, as well as teacher and environmental variables. Excess behaviors showed small decreases over three years and these data stand in contrast to previous epidemiological findings for institutionalized children. For a sub-sample of the children individual interventions were carried out, to determine whether response patterns could be identified. Methodological issues involved in such analyses are discussed in detail. Although reliable clusters proved difficult to identify, both the child studies and investigations of teachers' decision-making suggested that programming based on the acquisition of functional skills was the most valuable strategy for reducing excess behaviors. A model was developed to assist teachers and other clinicians select interventions,and the educational validity issues in evaluating interventions were discussed.

FINAL REPORT

Project No. 443CH 90024
Grant No. G00-790-1960

The Selection of Intervention Priorities in Educational Programming of Severely Handicapped Preschool Children with Multiple Behavioral Problems

Ian M. Evans and Luanna M. Voeltz
University of Hawaii
Honolulu, Hawaii

December 1982

U.S. Department of Education
Special Education Programs
Division of Innovation and Development

Table of Contents

List of Tables

List of Tables (Contd.)

List of Figures

Chapter 1

Introduction

Approximately two percent of the school-aged population is identified as severely handicapped for purposes of educational programming. These children possess "severe language and/or perceptual cognitive deprivations" and additionally display any or all of the following excess or abnormal behaviors: (1) failure to respond to pronounced social stimuli; (2) self-manipulation; (3) self-stimulation; (4) manifestations of intense and prolonged temper tantrums, and; (5) absence of rudimentary forms of verbal control[1]. Children evidencing such characteristics include the profoundly and severely mentally retarded, those with two or more serious handicapping conditions (e.g., the mentally-retarded blind and the cerebral-palsied deaf) and those seriously emotionally disturbed children diagnosed as autistic and schizophrenic. In some cases, such children may also have an extremely fragile physiological condition--often true for the profoundly retarded but almost never the case for an autistic child.

When educational intervention is introduced for severely handicapped children it is apparent that those children who exhibit multiple behavior problems present the greatest difficulty to educators. This is especially noticeable now that special educational programs designed to enhance adaptive skills are clearly effective for most developmentally disabled children. Conventional or special educational teaching strategies are minimally effective with children who are socially unresponsive, who do not attend to environmental stimuli, who show no sustained involvement with external objects, who respond to attempts at social intrusion with crying, tantrums, aggressive or self-injurious behavior, or who engage in persistent, repetitive activities that have no significance for learning or development. The presence of such behavior cuts across all the traditional diagnostic categories of severe childhood handicaps. It is also likely that children exhibiting these types of behavior are referred to time consuming and costly consultations by neurologists, psychiatrists and psychologists more frequently than children exhibiting similar levels of adaptive behavior and are also harder to place in long-term educational settings. Paradoxically, therefore, those children in greatest need of training in adaptive skills may receive the least intensive treatment services because of the disruptive effect of these excess behaviors on the educational process.

Numerous guidelines are available to assist teachers and parents in the selection of priority instructional objectives in the various skills domains relevant to the ultimate functioning of severely handicapped persons. Traditional skill sequences are available in the cognitive, communication, motor and sensorimotor, socio-emotional, and self-care or adaptive behavior areas. These sequences may be either developmental or behavioral-task analytic in

[1] Federal Register, 1975, 40:35, 7412.

nature (or they may be a combination of behavioral-developmental), but they are explicitly organized in a test-train format such that priorities emerge directly from the assessment process. The alternative ecological assessment and functional programming model articulated by Brown and his colleagues also provide (if diametrically opposed) criteria for selecting skill acquistion targets on behalf of severely handicapped learners. The Individualized Education Program format required by Public Law 94-142 seems well-designed for documentation of these educational decisions on a year-by-year basis.

Unfortunately, a similar level of sophistication does not exist in the form of empirically-based priorities for programming for the decrease of excess behaviors. With the exception of isolated attempts to assess socio-emotional adaptation level (for example, see Schopler and Reichler's PEP, 1976), no sequenced and prioritized instructional recommendations exist for programming either for the development of appropriate social interaction and play behaviors or the systematic decrease of those behaviors judged to be negative ones in the child's repertoire. Furthermore, those efforts which do exist are primarily intended as gross measures to assess clinical improvement rather than providing a sequence of intervention priorities. Generally, behavior modification with excess (problem) behaviors has proceeded on an ad hoc basis wherein single target behaviors are selected for intervention according to arbitrarily criteria and with little regard for the effects of each such intrusion on the child's total repertoire and longitudinal development.

Finally, there exist no guidelines for teachers and other clinicians to use in determining the appropriate combination of programming for both skill development and the decrease in excess behavior: concomitant needs presented by severely handicapped children who additionally display multiple behavior problems. Such children present many more potential instruction and behavioral objectives than could be systematically programmed at any given point in time. Which needs should be addressed first? In fact, there exists no concensus regarding even the general issue of whether or not excess behavior should be modified before skill instruction or in conjurction with efforts to teach new skills (Koegel, Egel, & Dunlap, 1980; Lathey, 1978; Lovaas, 1981; Gaylord-Ross, 1980; Schroeder, Mulick & Schroeder, 1978) After 15 years of demonstrated (if temporary) efficacy of behavior modification procedures in the decrease of negative behaviors considered one at a time, it would seem appropriate to attempt to establish socially and empirically valid recommendations for selecting particular priority target behaviors and intervention procedures over others,under which conditions, in conjunction with each child's overall educational needs. Given the multiple educational needs of severely handicapped children with severe behavior problems, the selection of particular intervention targets invariably means that others will not be modified. Furthermore, the selection or certain intervention targets may result in intended and unintended positive and negative behavior changes. Obviously, teachers, clinicians and parents should attempt to identify both effective and efficient behavior changes, i.e., those which are associated with maximum benefit to the child. Particularly given the limited educational time available to handicapped children, it seems crucial that those choices assoicated with the most beneficial outcomes must be made.

Careful consideration of the multiple and longitudinal effects of each intervention decision is essential if the precious instructional time of severely handicapped children is to be utilized wisely.

Description and Classification of Excess Behavior.

The behaviors we have been loosely referring to as abnormal, maladaptive, or disruptive will be referred to as "excess" behaviors throughout the remainder of this discussion. By this is meant that they are behaviors which occur with excessive frequency, intensity, or duration and which need to be reduced or extinguished in order to meet common societal norms for appropriate child behavior. The term "abnormal" implies that these behaviors are qualitatively different from those found in non-handicapped infants and children, and the term "maladaptive" implies that the behaviors are not functional for the child. Neither of these assumptions has been demonstrated to be true and so will be avoided.

Although the number of such excess behaviors is potentially infinite, in practice it seems possible to identify a limited number of behaviors which are quite commonly observed. However it must be realized that the labelling of many of these behaviors is an arbitrary function of the way the observer selects units from the "stream of behavior", so that the "inattention to visual stimulation" might be the same behavior as "fixed staring at hands", and "repetitive manipulation of objects" might be the reciprocal of "lack of toy play" or a cause thereof. Similarly when we group these behaviors into categories we may do so on the assumption that they serve a similar function, e.g., "self-stimulatory activities", or on the basis of common topographies, e.g., "hand and finger mannerisms".

Forehand and Baumeister (1976) proposed categorizing these behaviors into sterectyped acts, antisocial behavior, and classroom disruption. Another scheme, based on functional categories, has been proposed by Teaching Research who grouped excess behaviors into four categories based upon experience with the severely handicapped in the State of Oregon (Baldwin, personal communication):

1. Self-indulgent, attention-seeking ("Me first"): this includes behaviors such as crying, whining, "dependent" demands for teacher attention, etc.

2. Noncompliant ("I don't want to"): refusals to perform a task, etc.

3. Aggressive ("I'll get you"): purposeful hitting, pushing, biting, taking objects from others, etc.

4. Self-stimulation ("Doing my own thing"): refracting light, masturbation, finger flicking, rocking, etc.

A possible problem with this schema is the implied intentionality which may not be psychologically real. Some behaviors could functionally

fit into more than one category; e.g., self-injurious behavior can at various times be described as self-stimulatory or as attention-seeking in different children, or even in the same child in different situations, across time, etc. Excess behaviors need not, of course, have any function, and may emerge as components of neurological dysfunctions, such as seizure-induced behavior, tremors, and athetoid movements.

It may be more useful for purposes of intervention to conceptualize excess behaviors into just two gross categories: (1) ritualisms, and; (2) manipulatives. A third set, self-injurious behavior, could appropriately be located in either of these categories, dependent upon the function which the behavior demonstrates.

Ritualisms are those behaviors which appear to be non-goal oriented but engage the child's total attention during the time in which they occur. These behaviors have been variously described as bizarres, stereotypes, rituals, perseveration, self-stimulation, etc. Some writers include at least some types of stages of self-injurious behavior in this category (Williams, 1974), though there is also evidence that self-abuse may be utilized functionally by the child (Carr, Newsom & Binkoff, 1976; Gaylord-Ross & Weeks, 1978). A feature which differentiates ritualisms from those behaviors in the category of manipulatives is that the former behaviors are thought to be self-reinforcing. That is, they function as their own reward and thus require no external consequences from the environment to maintain and may even be unresponsive to any external consequences intended to decrease their level of occurrence (Wolery, 1978). Ritualisms occur at high rates, in the absence (or may prevent the initiation) of social interaction, and are either highly resistent to change efforts by intervention agents or are highly likely to reappear at a later date when specific programming efforts cease (Brannigan & Humphries, 1972; Hutt & Hutt, 1970; Koegel, Firestone, Kramme & Dunlap, 1974; Smith & Connolly, 1972). Examples of ritualisms include hand clapping, finger flicking, rocking, spinning objects, head weaving, repetitive vocalitations, hand/arm flapping, etc.

Manipulatives are behaviors which appear to be goal-directed and under environmental control, perhaps occurring specifically in the context of social interaction or designed to initiate or end a social interaction. In particular, such behaviors: (1) seem intended to function as aversive stimuli or consequences to others. By refusing to respond to environmental demands or intrusion efforts by others, the child may display certain behaviors which appear particularly well-suited to function as punishment to the caregiver, thus reducing those caregiver instruction efforts (e.g., temper tantrums, crying, etc.), or; (2) occur whenever reinforcement for another behavior--previously reinforced--is not forthcoming (e.g., attention seeking whining, pushing, shouting, etc. See also Lovaas, Freitag, Gold & Kossorla, 1965). Examples of manipulatives include temper tantrums, noncompliance, crying/screaming, hitting others, running away (bolting), throwing objects, etc.

No clear consensus exists as to whether self-injurious behaviors (SIB) are either self-stimulatory or manipulative in nature. It appears that individual SIB occurrences can be included in either category for different

children, different situations, and/or different times (Neapolitan, 1978). It should be possible for a clinician to determine whether a particular SIB is functioning as a self-stimulatory or a manipulative behavior, and thus incorporate this knowledge of the behavior's function into any intervention plan (Carr, 1977; Williams, 1974; Frankel & Simmons, 1976; Bachman, 1972). Examples of SIB include face slapping, head banging, scratching/picking skin, self-biting, eye-gouging, etc.

There is clearly a need for a more adequate schema--firstly for identifying these excess behaviors and secondly for grouping or classifying them according to empirical criteria--since no two children are going to exhibit exactly identical behaviors. To investigate the former problem it would seem necessary to gather examples of excess behavior in as neutral a fashion as possible, for instance by ethologically-oriented descriptions of behavior from an extensive sample of severely handicapped preschool children. As these behaviors presumably alter their topographies over time and in conjunction with developmental changes and the acquisition of new skills, such a survey should include the longitudinal survey of selected excess behaviors. Furthermore, the identification of excess behaviors from the total behavioral output of the child involves a judgement process by the observer. There is a need for studies of how the adults responsible for teaching and parenting the severely handicapped child isolate, respond to, and interpret these various behaviors.

The problem of classifying or grouping behaviors according to some logical criteria, such as their functional significance for the child, is an issue which is closely related to theoretical conceptions regarding the organization of these behaviors and their interactions with other behaviors and with environmental events. Unverstanding the functional organization of excess behavior is the key to determining treatment priorities in a rational, empirically-based fashion. The next section summarizes some of the organizational concepts that have already been proposed in the literature.

Review of Origins and Organization of Excess Behaviors.

(a) Lack of adaptive skills and alternative competing responses. In normal children, socialization skills appear to develop sequentially in a manner analogous to developmental progressions mapped for language, sensorimotor, cognition, etc. Such a sequence begins with variations in crying and smiling behaviors in the infant as a function of caregiver behaviors, and moves through "stages" of solitary, parallel and cooperative play with peers in the older preschool child. The social repertoire becomes increasingly more complex throughout the school-age years, culminating in a well-adjusted adult who knows even the "residual rules" governing social interaction (Scheff, 1966). Strain, Cooke and Apolloni (1976) have documented the importance of these social interactions as the context for acquiring much information in all other areas of development, in addition to their intrinsic value of adaptation to the social environment.

This sequence of socialization is absent in severely handicapped children, and may in fact be prevented from occurring by the environmental response

to a "deviant" child (Buium, Rynders & Turnure, 1974). A serious delay in cognitive and motor development limits naturally occurring responses which might then be reinforced in a social context by caregivers. An ecosystem of adaptive interactions with the environment would fail to develop under such conditions, and the handicapped child might be described as an organism in isolation from his surroundings and the information therein. It may be that all excess behaviors can be redefined in terms of the absence of adaptive or skill behaviors, including social ones. Excess behaviors might be prevented from occurring in a population for whom adaptive behaviors in each response class are carefully programmed, and could presumably be replaced when they do occur by teaching approriate incompatible skill behaviors as the child is developmentally and physically ready. Horner, Holvoet and Rinne (1976) suggest in particular that bizarre stereotyped behaviors may be a consequence of the lack of cognitive and motor repertoires which might promote the development of play and similar social interactive skills with which to occupy free time. Schuler (1982) argues that excess behavior often serves a communicative function for children with limited communication skills, and Carr, Newsom and Binkoff (1980) presented evidence that aggression was utilized in certain situations by severely retarded individuals as a strategy of saying "no" to tasks they considered aversive. Similarly, behaviors such as screaming, crying and hitting in the context of social interaction may actually reflect "tactual defensiveness", i.e., the lack of appropriate association of touch with pleasant outcomes (Bradtke, Kirkpatrick & Rosenblatt, 1972; Siegel, 1972).

The concept that at least certain excess behaviors are a function of absence of more appropriate skills implies that the excess behavior has the same function, or is maintained by the same reinforcer, as the skill. Thus stereotyped behaviors may be thought to generate important stimulus input in the way that more appropriate play would serve the more skilled child. There have been some attempts to document similarity of function between appropriate and inappropriate behaviors (Kissel & Whitman, 1977; Voeltz & Wuerch, 1981), although successful reduction of the excess behavior by teaching the skill does not necessarily establish this common functional element. In Flavell's (1973) study of the reduction of stereotypes by the reinforcement of toy play, for instance, the excess behavior might have been directly influenced by the training proceudres, or might simply have been incompatible, in the physical or topographical sense, with the play activities. This issue is also pertinent to the report by Kissel and Whitman (1977) of the effects of play-training on the self-stimulation of a profoundly retarded boy. The amount of self-stimulation seemed closely related to the amount of play, so that in situations where play was limited, such as the ward, self-stimulation reamined at high baseline levels despite the acquisition of the skill itself. If toy behavior is only topographically incompatible with self-stimulation, then one could expect reduction in hand stereotypes, but not, for example, oral self-stimulatory responses such as blowing or teeth grinding.

The importance of understanding the relationship between excess behavior and some more adaptive social skills is underscored by the recent reports of the relative failure of direct modification techniques, such as overcorrection, to eliminate stereotypic behaviors. These failures have mostly been in terms

of the increase of other self-stimulatory behaviors following the decrease in the target (Epstein, Doke, Sajwaj, Sorrell, & Rummer, 1974; Foxx & Azrin, 1973), through Becker, Turner & Sajwaj (1978) reported only a temporary increase. When Rollings, Baumeister and Baumeister (1977) reported that the suppression of body-rocking in one subject was associated with an increase in other stereotypic behaviors, they recommended that training programs designed to suppress such behaviors should attempt to develop more desirable behaviors in their place. Ironically, a review of self-injurious intervention efforts revealed that only 18% of published studies included a controlled training program for positive behaviors (Johnson & Baumeister, 1978). Similarly, Schroeder, Mulick and Schroeder (1978) emphasize that a major deficiency in past research is a failure to even report those simultaneous skill building efforts which do occur. Because of this and other problems noted earlier, no systematic guidelines exist to aid the choice of effective alternative behaviors for educational programming.

(b) Response Classes and Clusters. There is considerable data as well as theoretical support for the expectation that any manipulation of one target behavior will result in predictable changes in other behaviors which were not specifically manipulated. There are numerous constructs noted in the literature which clearly suggest an observable phenomenon; Table 1 provides a listing of these constructs.

Skinner (1953) has conceptualized the notion of response classes, defined as sets of discriminable behaviors which would vary together and in a predictable pattern. This would imply that the deliberate manipulation of any one behavior should also result in measurable changes in other members of that response class in a predictable direction. At the theoretical level, only Staats (1975) has taken seriously the interdependence of behaviors and discussed the functional consequences of responses as independent variables for other responses. At the applied level, there has long been evidence suggesting the reality of behavioral inter-relationships with considerable consequence for efforts to intervene with children's behavior. In one of the earliest and most well-known behavioral intervention studies with a severely handicapped child, Risley (1968) described the appearance of a topographically similar and equally undesirable side effect (chair climbing) when a negative target behavior (wall climbing) was punished--through the new behavior was also extinguished with no further undesirable behaviors which are closely related to the punished behavior (e.g., an effort to extinguish echolalia might extinguish all speech, both echolalic and spontaneous). Unfortunately, with the notable exception of Wahler's work (discussed below), behavior modifications efforts with children during the past 20 years have focused almost exclusively upon single dimensions of behavior, in which single intrusions are conducted in piecemeal fashion.

Willems (1974) has called for a serious investigation of the way in which behaviors and environments might be interrelated, stating that:

> The question of larger and unintended effects within interpersonal and environmental contexts and over long period of time beg for evaluation and research, because lessions learned in other areas suggest that we should always be sensitive to "other" effects of single-dimensional instrusions (p. 346).

He argues that a system-wide ecological outlook is more re-

TABLE 1.1

Constructs Suggesting the Existence of Behavioral Interrelationships[a]

A. Phenomena observed in intervention studies.

Side effects (unintended effects): unplanned or unexpected changes (usually negative) in behaviors which were not directly altered.

Sympton substitution: the appearance of another (undesirable) behavior along with the successful reduction of the intended target, assumed to be "taking the place" of the reduced behavior due to some remaining need which has not been addressed by the intervention.

Collateral/Concomitant effects: changes in behavior (generally considered positive or neutral) not specifically programmed during intervention.

Response generalization: the appearance of responses which were not specifically trained but are similar to the target behavior such that they are presumed to belong to the same response class or have the same function.

Generalized benefits: broad benefits (e.g., positive social interactions, enhanced learning) which occur because the learner has acquired a certain "crucial" behavior or skill important to his/her more general functioning.

B. Theoretical (behavioral) constructs:

Adventitious reinforcement: modification of an unintended response due to the chance temporal contiguity between that response and the one actually being modified through some reinforcement contingency.

Response interdependency: the occurrence of one response is dependent upon the prior of simultaneous occurrence of another, i.e., one response functions as the independent variable for another.

Functional equivalence: two response which may differ topographically nevertheless serve the same function and thus may be interchanged.

Functional incompatibility: behavioral states or responses which prevent or interfere with alternative ones (e.g., attentional states).

Setting events: environmental events which have a broad influence on an individual's behavior, usually considered to be temporary (e.g., a death in the family, a move).

Developmental prerequisites: behavior considered to be an essential component for the development of later, presumeably more complex, behaviors and stages (e.g., object permanence is viewed as a prerequisite to the use of symbols as required for language development).

Mediating responses: a general response presumed to have utility for the development of more complex behavior (e.g., problem-solving skills, learning-to-learn, self-control strategies).

Habit family hierarchy: response which occur according to relative probabilities given particular situations and a common stimulus response.

Displacement (adjunctive behavior): irregular, stereotyped and active movements occurring when another response pattern is prevented or interrupted.

C. Terms used to describe reponse organization in the literature.

Response class: topographically different behaviors which share the minimal conditions of a response needed to obtain reinforcement.

Response chain: contiguous responses where one serves as the discriminative stimulus for the other, or as the conditioned reinforcer for the preceding response.

Response hierarchy: different probabilities of occurrence of alternative responses in given situations dependent upon individual learning history.

Keystone behavior (pivotal skill): a response which is necessary and sometimes sufficient for the appearance of other, usually positive responses.

Concurrent behavior: the occurrence of two or more responses in rapid alternation, produced by the control of two or more schedules of reinforcement which are simultaneously in effect.

---

[a] Condensed from Table 1 in Voeltz and Evans (1982).

presentative of the actual sequences of events, and that an idiosyncratic perspective of capricious or randomly appearing individual behaviors in the repertoire of children with problems is not supported by empirical evidence of multiple effects and environmental functions. Though Baer (1974) did state that some response chains and classes in children might well be capricious, he also supported the possibility that individuals will share similar chains or classes such that a typical pattern, which would provide useful intervention information, might emerge through careful study.

Voeltz and Evans (1982) have provided a systematic evaluation of the clinical evidence supporting the existance of behavioral interrelations. The four major sources of published empirical data relevant to this issue are summarized briefly here as follows:

(1) Intervention Studies with Negative Targets.

Following an exhaustic review of behavioral intervention studies, only 29 studies were located (through summer 1980) in which a successful effort to reduce an excess target behavior included information on concomitant nontargeted behavior charge. The quality of the evidence varied greatly, including anecdotal reports by the child's teacher, parents, or the authors themselves, potentially verifiable written records (e.g. movement to a less restrictive educational placement), and systematic behavioral observations of the non-targeted behaviors reported to have changed. The latter information would provide the strongest support for behavioral interrelationships, and 20 of the studies provided such data. However, few of those studies monitored the behaviors through all phases of the experimental design, and the more recent work indicates an increased use of subjective clinical impressions of generalized improvement, etc., rather than actual data to validate concurrent behavior changes. Findings of specific behavior changes included both positive and negative unintended effects, with no clear patterns which might imply behavioral interrelationships across children and/or behaviors. In general, then, these studies provide little basis for generalization regarding collateral effects or response interrelationships.

Nevertheless, the absence of clear patterns may simply reflect an absence of data. Our review of interventions with children revealed more than 90 discrete, topographically distinct excess behaviors identified as targets, and the small number of studies which reported collateral effects dealt with only a sub-sample of these behaviors. There are few examples of either direct or systematic replication across subjects where evidence of behavioral interrelationships has been documented (cf. Zlutnick, Mayville, & Moffat, 1975). In general, then, the absence of clear patterns may simply reflect an absence of data. This situation can only be ameliorated if future studies monitor multiple responses, including both the intended target behaviors and other behaviors which are of some concern and which might also evidence change.

(2) Intervention Studies with Positive Targets.

There has been a traditional interest in behavioral interrelationships in intervention efforts to increase children adaptive behavior, includ-

ing: (a) appropriate toy-play as a substitute for self-stimulation (Kissel & Whitman, 1977; Voeltz & Wuerch, 1981); (b) independent toy play as an alternative to oppositional behavior (Wahler & Fox, 1980); (c) communication skills as a functional alternative to disruptive behavior (Casey, 1978; Schuler, 1982); and (d) social skills in lieu of agression (Bornstein, Bellack & Hersen, 1980). This work has attempted to demonstrate a reciprocal relationships between the two classes of behavior, such that the acquisition of a positive skill will result in a decrease of the negative behavior. Thus, Gaylord-Ross (1980) recommends a curricular approach to the remediation of excess behaviors whenever possible, i.e., the teacher should identify and program an incompatible skill to replace a negative behavior and include more direct modifications of the latter only if a curricular approach fails.

(3) Learning States and Developmental Prerequisites

Work in developmental psychology on learning states and learning prerequisites also suggest the existence of the notion of response-response interrelationships. In special education, much remedial effort has traditionally been devoted to establishing attentive behaviors assumed to be necessary for learning, motoric prerequisites for subsequent complex motor development, etc. Risley (1968) proposed the term "functional incompatibility" to describe the possibility that certain excess behaviors may reflect or induce behavioral states which displace a learning state, and thus make it impossible for the child to learn while s/he is engaging in the excess behavior. He considered the stereotyped behavior of handicapped children as functionally incompatible with the establishment of new socially productive behaviors, for example, and this presumed relationship between certain excess behaviors and the development of skills has dominated both clinical psychology and special education. Yet, there are few systematic investigations of this issue. We could locate only two which provide direct evidence with regard to children, Koegel and Covert (1972) and Lovaas, Litrownik, and Mann (1971). Contrary evidence is provided by Rincover, Cook, Peoples, and Packard (1979) and Wolery (1978) who reported increases in skill acquisition by autistic and retarded children when they were allowed to self-stimulate briefly as contingent reinforcement for correct responses during instruction.

Evidence on developmental prerequisites is equally limited. Where instruction in isolated skills taught out of context has failed to result in behavior change which generalizes and maintains, this may be evidence of the negative effects associated with ignoring learning prerequisites, etc. There is evidence that training of particular complex skills can result in generalized improvements (Hart & Risley, 1980) and effects on nontrained skills can serve as prerequisites by providing specific patterns of stimuli for others in certain directions only.

(4) Correlational Studies

An alternative strategy to identify behavioral interrelationships would be utliziation of multivariate statistical procedures to examine patterns among responses derived from systematic observation of more exhaustive categories of behavior. There are several examples of this approach, most of them investigating behavioral covariation in naturalistic settings without

any direct attempt to modify any of the responses but instead relying on situational variations to determine cluster patterns and their stability (Harris, 1980 ; Lichstein & Wahler, 1976; Strain & Ezzell, 1978; Voeltz & Evans, 1979). Only two studies have gone beyond this descriptive approach. Kara and Wahler (1977) and Wahler (1975) utilized multivariate analysis of baseline data to select a target behavior and subsequently attempted to demonstrate that successful intervention of that behavior would also involve optimal positive concomitant effects upon other nontarget behaviors in a response cluster. To date, Wahler's work provides strongest support for the feasibility of ultizing multivariate data on response interrelationships to plan for maximum intervention effects.

(c) Physiological Correlates of Excess Behaviors. Most of the research discussed thus far has been concerned with practical issues regarding the reduction of specific excess behaviors in specific children with little attempt to understand the functional significance of these behaviors for the child. Although, as discussed above, there is growing interest in the inter-relationships between excess behaviors and others in the child's repertoire, the dominant conceptualization is still that of excess behaviors being operants, under the control of the external environment (Forehand & Baumeister, 1976). It is equally plausible that certain excess behaviors are associated with psychophysiological variables.

The bulk of the earlier research involving physiological measurement of multiply-handicapped children revolved around debates as to whether autistic children where chronically over-aroused or under-aroused. Results of such investigations were generally inconclusive and in some cases contradictory. More recently attention has turned to the notion that it is the regulation of arousal level that seems to present special difficulty for autistic children. MacCulloch and Williams (1971), for instance, noted that autistic children appeared to have difficulty in the homostatic mechanisms which regulate heart rate, and most clinical descriptions of the syndrome point to the striking tendency of autistic children to show over-arousal to some forms of stimulation and a lack of responsiveness to others. In a study more adequately controlling general activity level Hutt, Forrest, and Richer (1975) confirmed MacCulloch and Williams' (1971) finding, noting that the greatest variance of beat-to-beat heart rate in autistic children coincided with their stereotyped behaviors.

The relationship between the stereotyped, manneristic, repetitive behaviors of such children and general arousal level has also been noted. Hutt, Hutt, Lee and Ounsted (1965) reported that autistic children engaged in more arm and hand flapping, finger twisting, and other repetitive movements in a complex environment than in a simple one, and that these behaviors were correlated with activated EEG patterns. In line with the above findings it has become common to interpret these behaviors as being involved in the regulation of sensory input and thus arousal level. Other authors (Ornitz, 1976; Goldfarb, 1963) have noted the similarity between some autistic behaviors and those repetitive behaviors of deaf-blind children which appear to be related to increasing sensory stimulation. In accordance with concepts of the previous paragraph, it may be that these excess behaviors serve to modulate

arousal level rather than simply to increase it through sensory input. For example Bernal and Miller (1970) reported one child whose self-stimulatory behaviors ceased when presented with a series of tones, but increased with photic stimulation by 6-per-second light flashes.

The clinical and diagnostic significance of physiological measures of arousal was brought into sharp focus by Sroufe, Stuecher and Stutzer (1973) in a pioneering study. These investigators monitored psychophysiological parameters (heart rate, respiration, skin resistance, and muscle tension), overt behavior (facial expressions and autistic mannerisms) and task-related cooperative behaviors in one autistic child. Self-stimulatory finger flicking was associated with predictable increases and decreases in level of arousal. It should be noted that a rather different interpretation of hand and finger flapping has been suggested by Ornitz (1976). He has argued that these hypermotility patterns are compensatory activities that generate kinesthetic feedback used by the child in learning situations. Although different in detail to Sroufe et al's (1973) contention, the implication of Ornitz's theory is still that the excess behaviors serve an important role for the autistic child.

Clearly if excess behaviors are functionally related to psychophysiological integration rather than being random and purposeless, the implications for treatment goals are considerable. Psychophysiological monitoring would therefore appear to be a significant technique for the detailed understanding of the significance of excess behaviors for the handicaoped child. There also seems to be a barely explored potential for psychophysiological monitoring of children's physiological state during learning tasks which might allow for the early detection of phases related to excess behavior before such excesses become manifest in overt behavior. For example, Schroeder, Peterson, Soloman, and Artley (1977) report that self-injurious behavior in two severely retarded children was always preceded by specific EMG patterns of muscle tension; relaxation training might then interrupt the typical sequence. In fact, relaxation training has come to be incorporated into various programs for autistic children despite properly contrasted evidence that prior relaxation instruction does not necessarily reduce subsequent disruptive and stereotypic behavior (Marholin et al., 1978).

Project Goals and Research Components

Excess behavior in severely handicapped young children can present a major hindrance to the educational process, either by limiting the ability of the teacher to implement skill training programs or by more directly affecting the child's ability to learn from standard educational experiences. Because of this, numerous efforts have been made in the past to modify or reduce these excess behaviors, often with limited success or to the detriment of continued development of positive, adaptive social and academic skills. Piecemeal attempts to directly modify individual excess behaviors in individual children have resulted in no systematic guidelines to assist in establishing treatment priorities. Furthermore the common claim that excess behaviors are best remedied by teaching adaptive alternatives fails to recognize the possibility that the excess behaviors may be interfering with such teaching

or, more seriously, with the child's ability to learn. The purpose of the proposed research was to be able to clarify the nature of relationships of excess behaviors and/or skills by investigating a number of children in educational programs over a relatively long period of time and as a function of specific intervention efforts.

Research Components

The project involved three major components, which can be summarized as follows:

Study One: Ethological description of a set of "excess" behaviors in severely handicapped preschool children, taken as a group and additionally with reference to each individual child. A subcomponent will address the issue of caregiver responses to and categorization of these behaviors, as determinants of current typical intervention procedures.

Study Two: Analysis of covariation between teacher-child task related activities, child excess behaviors, and child skill level to formulate testable hypotheses of functional response class relationships for educational programming purposes.

Study Three: A series of (multielement, multiple baseline, reversal) individual intervention studies to test the hypotheses of response-response relationships and the issue of indirect behavior control over excess child behaviors.

Upon completion of Study Three, information was available to compile a systematic treatment package for behavioral categories in severely handicapped preschool children. This package included comprehensive guidelines for optimal prioritizing and sequencing of excess behavior programming in conjunction with skill level and specific excess behavioral repertories.

Study One: Description and Classification of Excess Behaviors

The purpose of this component was to obtain descriptions of the range of excess behaviors as they occur in the natural environments of severely handicapped preschool children. The goals were to set out an empirically-based descriptive listing of such behaviors so that consistent use of terms can occur across the various professionals working with handicapped children, and thus be more directly related to the appropriate recommended intervention procedures.

Part 1: A statewide sample of preschool children with severe developmental delays and excess behavior was identified and observed in educational programs.

Although the method of observation was ethological--meaning that the behaviors of interest were described as neutrally as possible without reference to functional assumptions such as "mannerism", "self-stimulatory", etc.--there was continuous interplay at this stage of the research between the naturalistic observation and the predetermined examples of behavior described above. Thus we developed a checklist of excess behaviors to facilitate the identification of children.

At this stage the emphasis was on description of the excess behaviors, not such factors as frequency, intensity and duration. However, once several individual cases were available it was possible to relate the type and number of excess behaviors displayed to the age level of the child, the kind of diagnosis given, the severity of the developmental delays (as measured by such brief assessment instruments as the TARC) and the manner in which the excess behaviors were described in the child's case history. We were also able to provide a frequency count of the number of different types of excess behaviors observed in our sample. In general, this is the first time that a comprehensive attempt has been made to provide basic demographic and descriptive data on the occurrence of excess behaviors outside of the context of institutional settings (cf. Balthazar & Phillips, 1977). These initial data were used to construct the observation system and the appropriate "header" codes for subsequent use in studies two and three.

While gathering observational data on the natural occurrence of these excess behaviors, we conducted a comprehensive review of all published descriptions of excess behaviors in which some sort of intervention procedure has been attempted. The review therefore summarized methods of intervention that have already been investigated and also classified each specific excess behavior reported. This is the first time such a review has been attempted and it demonstrated which excess behaviors have typically been identified for purposes of formal intervention research. The review listed, excess behavior by excess behavior, each intervention strategy attempted, the reference to the study reporting the technique, the results obtained by the authors, our own appraisal of the internal or external validity of the findings (adequacy of design, degree of generalization observed), and any evidence of inter-behavioral relationships either noted by the authors or apparent from their data. Figure 1 provides an example of this abstract-in-formation format. The major generalizations that can be drawn from this review were incorporated into the treatment-decision package.

Part 2: Those children identified as having a range of specific excess behaviors were selected as subjects for the longitudinal observations of excess behaviors. This study was considered necessary as an attempt to observe changes in excess behaviors which appear to be due largely to alterations in developmental level of the child. This is an important element for the later investigation of the effects of programming skill development for the children, since modifications in the excess behavior identified may appear

BEHAVIORAL SYSTEMS INTERVENTION PROJECT

Literature Review Abstract

EXCESS BEHAVIORS: (62) Putting others' hats
(45) Punching others

INTERVENTION STRATEGIES: (1.1) Positive reinforcement

REFERENCE: Carr, E. G., Newsom, C. D., & Binkoff, J. A. Escape as a factor in the aggressive behavior of two retarded children. Journal of Applied Behavior Analysis, 1980, 13, 101-117. Study II.

DESCRIPTION OF EXCESS BEHAVIORS: Aggressive behaviors--pinching (squeezing with a pincer grasp); hair pulling (grasping or pulling experimenter's hair); scratching (digging in the fingernails and dragging them across experimenter's skin or clothing).

DESCRIPTION OF INTERVENTION STRATEGIES ATTEMPTED: Demands condition- Subject handed a buttoning board and once every 10 seconds was instructed to button it with physical prompts provided when necessary. Each correct response reinforced by brief verbal praise. Removal of board signaled end of session (safety signal). No demands condition--Subject handed a buttoning board but made no demands. Demands + toys and food condition- Experimenter dispensed one of several toy or food reinforcers + verbal praise contingent upon a correct response.

SUBJECTS (AGE AND DIAGNOSIS): 9 year-old boy diagnosed as mentally retarded with autistic features

RESEARCH DESIGN: Reversal design (ABAC)

RESULTS OBTAINED BY AUTHOR: During the first and second demands condition, X frequency of aggressive acts was 61.0 and 71.1 respectively. Demands + food + toys resulted in X frequency = 12.1 and 9.8 during final condition.

MAINTENANCE: no

GENERALIZATION: n/a

ANY EVIDENCE OF INTERBEHAVIORAL RELATIONSHIPS EITHER NOTED BY THE AUTHORS OR APPARENT FROM THEIR DATA: Intensity of subject's aggression during the demands plus toys and food condition was much less than during demands condition.

Figure 1.1

to be occurring only because it is common for that particular type of behavior to change in some way as a function of maturation. The single subject methodology described later presupposes that the behaviors of interest will be relatively consistent over time, but it is unknown whether this is true or not for excess behaviors in this age range.

The methodology for this component involved videotaping the behavior of interest at regular time intervals. On the basis of the videotapes, the behaviors are described in terms similar to but more detailed than in the first study. Eventually attention could be paid to changes in the topography of the response; for example, does self-destructive behavior emerge from a background of self-stimulatory behavior, or do hand and finger mannerisms show much variation over time, and if so, can they be separated from truly stereotyped behavior?

Part 3: One of the ways in which excess behaviors are identified is by the implicit categories used by adults working with children. Furthermore, the adult's emotional response to the excess behavior is almost certainly a major factor in the manner in which priorities for treatment are established, as opposed to a procedure of decision-making based upon empirically established child needs. It is a common observation of behavior modification programs that caregivers of handicapped children are most likely to first focus upon the supposedly undesirable excess behaviors displayed by the children, and the majority of single case studies of behavioral techniques have been concerned with the reduction of undesirable behaviors often employing judgments if children's problem behavior tend to focus on aggressive, acting-out types of behavior that interfere with instructional procedures (e.g., Griffiths, 1952), but that teachers are generally able to identify emotionally handicapped children (Bower, 1961). To establish an empirically-based system of educational priorities for such children, it is important to investigate the factors involved in this initial adult decision-making process, both in terms of categorizing the excess behaviors and in terms of how adult attention is selectively focused upon problem behaviors (cf. Hawkins, 1975). This attitudinal information was taken into account as the final treatment/decision package was developed.

The methodology for this study was to select two children from among the experimental subjects and to make a series of videotapes of each child in which a variety of behaviors were displayed. These videotapes were shown to teachers, special education graduate trainees, educational assistants and other professionals and paraprofessionals involved in the treatment process for handicapped children. The adult subjects viewed the videotapes, categorized the behaviors observed, indicated the kinds of experiences they had had with these behaviors, and set priorities for intervention. By manipulating the behavioral content of the tapes, a typical judgment task was presented (Rapoport & Wallsten, 1972) from which we could isolate those cues which adults use to rank the significance and "change-worthiness" of the behavior involved. As the impact of the excess behavior on the adult is a major factor in treatment priority-setting, it was necessary to examine the

generality of our findings by alternative methods of investigation, such as an interview study in which teacher's priorities were investigated in the context of the specific goals they had established for specific children showing severe excess behaviors.

Study Two: Functional Analysis of Excess Behavior

Study One enabled us to design a comprehensive observation system which monitored an exhaustive list of excess behaviors, on-task and other instructionally-related responses, environmental variables, teacher and peer presence and behavior, and any other classification or situational variables which might be relevant for the functional analysis. In Study Two, a smaller sample of severely handicapped preschool children who displayed multiple excess behaviors were systematically monitored over a period of at least two years. The children were observed periodically in their typical educational setting at the time, and these data were analysed over time. The primary focus of this research phase was to develop methodological and statistical procedures which would be suitable for analyzing and describing behavioral clusters across time. This involved the elaboration of the coding system so that it was suitable for real-time observations using electronic data recording devices (MORE), writing programs which would generate meaningful summary statistics for MORE-described data, and looking at different methods for determining response covariation and then interrelationships in clusters to be found in extended baselines. Phase baselines provided the background data for these experimental subjects who were investigated in greater detail in Study Three.

Study Three: Intervention Series to Test Hypotheses

There were several aspects of our approach which mitigated against the utilization of a group comparison design in the intervention studies: (1) components of the treatment procedures have not been advanced to the state of a well-developed treatment package (i.e., the use of behavioral intervention procedures with preschool severely handicapped) but exist only as piecemeal intervention reports with single behaviors taken one at a time; (2) the population of interest (i.e., severely handicapped preschool children with multiple behavior problems) is of extremely low incidence and expected to be a heterogeneous group with respect to diagnosis (autistic, psychomotor retardation, Down's Syndrome, cerebral palsied retarded, etc.), developmental skill level and excess behavioral repertoire. This expectation was supported by the actual characteristics of those children referred to specialized services designed for children with multiple problems during the previous three years (e.g., the Diagnostic Observation Center which was originally used as a research site and source of referral of potential subjects). Generally speaking, the only unifying characteristics of these young children was that their behaviors were significantly more extreme than those of other severely handicapped children in their respective settings.

(3) Finally, the present research assumed that shared treatment needs are based upon the existing repertoire of functional and excess behaviors, rather than diagnosis or age per se. This repertoire is likely to show wide variation across children, but it is reasonable to assume that some direct and systematic replication will be possible. As emphasized by Hersen & Barlow (1976), neither direct replication series nor nonfactorial designs with nontreatment controls deal with the issue of generality across teachers, settings, or different children. Both factorial designs and systematic replication designs address the issue of generality of findings. The intent of this research was to develop a generalized design package for designing interventions with excess behaviors. Such a package should be based upon the systematic documentation of the interaction of skill-excess and excess-excess behavior relationships under varying treatment conditions, which involved the recording and analysis of multiple behaviors in order to identify clusters, response classes, or relationships between behaviors. Then the present research would lay the groundwork for group comparison designs utilizing the procedural guidelines we developed, which might then ethically justify the consumption of significant program time for a low incidence severely handicapped population. However, such group intervention approaches may become increasingly difficult to implement in view of the mandate for Individualized Education Programs. Nevertheless, as Hersen and Barlow (1976) emphasize, such group designs are not defensible until after the treatment package has been thoroughly tested through design procedures which admit for greater flexibility in the initial stages of research.

The exact nature of the individual intervention component (Study Three) therefore depended upon the characteristics of the children and the findings from Study One and Study Two. However, the basic procedures were as follows:

1. Study One established excess category descriptors;

2. Study Two established initial behavioral covariation patterns (including excess behaviors) for given children;

3. Hypotheses regarding specific and multiple intervention outcomes provided the basis for the experimental investigations conducted with individual children.

All individualized intervention decisions and procedures occurred under criterion conditions--in the child's actual (special education classroom) environment, determined by teacher and parent as being consistent with each child's IEP, and implemented by regular professional and paraprofessional

special education services personnel in the various Department of Education (or Department of Health) settings. Thus, our research efforts were fieldbased and reflected the actual settings and contingencies of public school and community settings. This was essential if our procedural recommendations are to be considered replicable and our findings generalizable to other educational programs.

Educational Implications

Previous behavioral research on interventions with excess behavior in handicapped children has thus far not advanced significantly beyond single-intrusion efforts, i.e., modifying and monitoring change in a single targeted response within a limited time frame. The vast literature on the problems of maintenance and generalization of such behavioral changes reflect the limitations of this state of affairs (e.g., Carr, 1980; Cone, 1973; Rincover & Koegel, 1975; Lovaas, Koegel, Simmons & Long, 1973; Walker & Buckley, 1972).

This project varies significantly from that approach while adhering in principle to the basic behavioral paradigm, by attempting to modify and monitor all behaviors requiring intervention through the careful (and documented) selection of key behaviors. A demonstration of predictable covariation of excess and skill behaviors in the repertoire of the severely handicapped could radically alter behavioral theory and the application of behavioral intervention procedures with handicapped children. In particular, the process of target behavior selection is currently guided primarily by caregiver preference (an issue of social validity) in conjunction with professional expertise (presumably based on empirical validity). If children's behaviors are interrelated, then decisions which result in the modification of intended target behaviors can be expected to have broader effects upon children's repertoires. Thus, decisions to intervene must consider more than the "change worthiness" of a particular target, and must be made with reference to possible effects upon other behaviors which might be altered in various ways. Ultimate treatment or educational validity is affected by such programming decisions, and is not simply an issue of demonstrating effective changes in designated targets. What is needed is a process whereby the decisions which are made result in the largest positive effects on the child, across time, and with reference to eventual outcomes. As long as research and intervention efforts focus upon monitoring only a single troublesome behavior, or anything less than a major portion of the child's total repertoire, "unintended" effects cannot be mapped and the possibility of maximumally efficient behavior change cannot be investigated or demonstrated.

The ultimate goal of this project was to develop systematic guidelines which teachers and other clinicians can utilize to select programming

priorities on behalf of severely handicapped children with multiple behavior problems. These guidelines were based upon: (1) an analysis of the intervention literature with such children; (2) a synthesis of existing ethical and professional recommendations regarding how such decisions ought to be made; and (3) empirical and social validity data regarding both decisions and possible multiple outcomes based upon the findings which emerged from the research components outlined here.

CHAPTER 2

The Research Setting: Method

Subjects[1]

Initial Screening and Identification of Potential Subjects. The first task of the project was to identify the statewide sample of severely handicapped children with multiple behavior problems, ages birth through eight years, not limited to a specific ethnic, socioeconomic or geographic group. Identification began by contacting all State and private agencies serving handicapped preschool age children and requesting an opportunity to observe their children to look for excess behaviors of the kind described earlier in this proposal; this included all Infant Stimulation Programs, the Child Development Center, the Diamond Head Child Development Clinic (including the Diagnostic Observation Center) responsible for most of the initial assessments of children with developmental delays both on Oahu and the Neighbor Islands, and all public school special education services enrolling children in this age range. All programs were visited early in each project year to observe children who might fit within the subject selection criteria to be described below.

In addition to surveying all children currently enrolled in known programs, additional case finding efforts were initiated by the project. The present system in Hawaii is successful in identifying those children with developmental delays who also display obvious indications such as physical stigmata, epilepsy and sensory and orthopedic impairments, but children with significant behavioral problems and developmental delays who do not show physical signs are more likely to not be identified through existing procedures. Such children form a significant percentage of the severely handicapped population and are generally initially referred subsequent to age three (Meier, 1975; Stein & Susser, 1975). Additional search procedures included:

- Contact with 40 additional programs with direct or indirect contact with this population, via letter, brochure, and/or phone.
- A massive literature mailing of the project brochure, a letter containing specific information directed toward professional staff explaining the project and children of interest, including list of "Behaviors of Concern" (see Table 2.1).
- Informational poster Watch Me Grow (designed to help parents detect delays in their children ages 1 month - 6 years) distributed to all private agencies, psychologists, and child psychiatrists.
- Public service announcements on the four major television networks, aired during November 1979.

[1]This chapter will focus upon the research components involving children. Specific information on sample and setting, method, etc., for the decision studies (involving teachers and other service delivery personnel) can be found in Chapter 4.

- A radio talk show interview explaining the project and search in Fall 1979.
- Agency Newsletter information spots (HSAC, SCFC, HARC, Commission on the Handicapped, etc.).
- Presentations at local professional meetings and conferences.

During fall 1980 (year 2), a new series of site visits was conducted to those newly established preschool classes serving severely handicapped children to identify additional potential subjects. No additional search procedures were conducted during year 3 on the assumption that our extensive communication with all programs would (and did) insure than any "new" children would be brought to our attention by school and agency staff.

Subject Selection. Subjects identified by the methods outlined above exhibited the following general characteristics:

1. Be of preschool age, i.e., from birth through eight years.
2. Exhibited one or more of the excess behaviors described in Chapter 1 such as stereotyped mannerisms, self-stimulatory behavior, self-injurious behavior, resistive or non-compliant behavior, etc.
3. Showed a significant developmental delay in at least two of the following areas: language, socio-emotional (affective) development, gross and fine motor skills, and self-help activities (toileting, feeding, etc.) as measured by assessments such as the TARC (Sailor & Mix, 1975) and UPAS (University of Washington, 1978).
4. Had been diagnosed as moderately to profoundly mentally retarded, autistic, deaf-blind, or severely multiply handicapped (in the case of children already diagnosed and receiving services).

Within the framework of the AAMD Adaptive Behavior criteria, these children would be located within the moderately, severely and profoundly handicapped ranges.

Subject Groupings. For all the children identified as fitting our criteria, permission was obtained from the parents to collect basic demographic and clinical data from the child's case records and to observe and record excess behaviors as they occur in the natural environment. This was the initial stage of the descriptive component of our research and allowed for a detailed listing of excess behaviors across a relatively large group of children (see Chapter 5).

From this group of children, two subject populations were selected. The "comparison" group of children displayed representative examples of excess behaviors and were observed (on video-tape) repeatedly over a two to three year period. Although the intention of this component of the research was to obtain naturalistic observation of excess behaviors over time, the observation of these behaviors did not in any way interfere with

TABLE 2.1

BEHAVIORS OF CONCERN

0-5 months	- Child does not "mold" to your body when picked up, does not anticipate (extend arms towards you, look at you) when you are going to pick him/her up.
	- Child cries excessively for long periods of time, and cannot be comforted.
	- Child seems to prefer being left alone for long periods of time, seems quiet and content for hours without any interaction with you, and may rock in his crib for a half hour or more.
	- Child does not explore objects or play with infant toys appropriately.
8-18 months	- Child is not crawling (for ages below 15 months) or walking independently (15-18 months).
	- Child has no "words" or sounds which seem to stand for words.
	- Child does not seem to recognize his/her name or respond when she/he is called.
12 months and above	- Child spends a considerable amount of time waving arms up and down, spinning self and/or moving fingers in ritualistic patterns.
	- Child will not look at you and seems to prefer being left alone.
18 months and above	- Child throws frequent and very disruptive temper tantrums (thrashes arms and legs, cries, screams, etc.).
24-36 months	- Child is not making any progress towards toilet training or actively resists toileting.
	- Child does not have any two word sentences, but uses only single words or perhaps does not talk at all.
30 months and above	- Child's sentences are primarily one to two words in length.

Remember: Children develop at different rates. Evidence of any of the above behaviors may not be cause for alarm, but should be reported to your physician.

the efforts of the programs enrolling these children to modify these excess behaviors. The comparison group essentially provided a control group whose excess behaviors were programmed via current procedures by their teachers and other staff on an ad hoc basis, as opposed to the experimental group whose excess behaviors were more systematically investigated throughout our intervention efforts. However, if a child in the control sample displayed harmful self-injurious behaviors, etc., which only our observers noticed, this information was relayed to those responsible for the child's treatment program and consultation was provided. Thus the requirements of unobtrusive observation was always secondary to the needs of the child, although, again, this consultation is representative of services which should ordinarily be available to handicapped children.

The second subject population were those children designated as the experimental group. A subject sample of approximately twenty children was identified at an early enough point of time to be included in Studies 2 and 3.

After a determination had been made that a child presented the behavioral criteria relevant to the research project, the parents were informed of the project activities and a request was made to include their child in the sample. Informed consent from the parents of all children in the research group was obtained, emphasizing that enrollment and thus services in an educational program were not contingent upon participation in the research project. Across the three years of the project, there were no refusals to participate in research efforts with one exception. During year three, a transferred special education teached refused to allow us to continue observations of an experimental child enrolled in his class. We thus observed this child in other settings (i.e., at home and in a simulated instructional setting at the University).

Subject Protection and Confidentiality. Videotape recordings made of the children in the naturalistic observation group were identified by code numbers only and kept in locked cabinets in the project offices at the University of Hawaii. Clinical case folders for all the research children in educational placements were kept in closed files according to current Department of Education and Department of Health practices; only professional staff involved in the education of the children had access to these folders. Research data in the form of behavioral observations, coded frequencies of behavior, and psychophysiological tracings were coded by number on computer records rather than by the child's name. All project staff were introduced to their ethical and clinical responsibilities toward the children in introductory training sessions conducted by the principal investigators.

Settings and Staffing for Experimental Subjects. The primary settings for the project were special education services for the handicapped children, available to all children ages birth through school-age at no cost to parents. By fall 1980, the State of Hawaii implemented full services for preschool handicapped children, ages two and one-half and above, in neighborhood elementary schools close to children's homes. Prior to that date,

handicapped children younger than age 5 (for whom services were mandated by state law) were served primarily by the Department of Health (DOH) which continues to provide special education and related services for developmentally disabled children from birth. During year one (1979-80) prior to full implementation of preschool services, many of our subject children were enrolled at the Diagnostic Observations Center (DOC), a small, intensive, research-oriented program operated by the Children's Health Services Division of the DOH. The DOC was established in 1974 as a long-term assessment placement for preschool children (ages one through five) from Oahu primarily, with multiple behavioral handicaps (including autistic characteristics) and/or who presented as "diagnostic enigmas". Long-term assessment involved detailed multi-disciplinary assessments in conjunction with treatment and the investigation of individually appropriate instructional strategies through "trial" educational programming efforts. The intent of the DOC program was to identify the child's handicapping condition and level of performance, negotiate an appropriate placement, and prepare a "field-tested" IEP-type educational program to accompany the child to that program. Although this setting could accomodate only a maximum of eight children, most of these children were appropriate subjects for the research project and the DOC therefore, offered us the most "concentrated" pool of subjects close to the University of Hawaii. However, with the establishment of preschool services in Department of Education public schools in year two and following (1980+), the movement of most DOC children into those schools, and our own preference to conduct research in more "typical" educational placements, the DOC was not utilized as a research site during years two and three.

Additional experimental subjects included in year one efforts were located at two elementary schools in Honolulu District (both serving autistic children) and one elementary school in the Leeward Oahu district (serving severely to profoundly retarded and severely multiply handicapped children). Year two (1980-81) sites included these three public schools and three additional elementary schools in Honolulu, Leeward, and Central Oahu Districts; one of these additional schools contained a preschool class for autistic children, another a preschool class for severely multiply handicapped (SMH) children, and the third a lower elementary class for primarily TMR (through enrolling our "autistic-like" subject) children. In year three (1981-82), sites attended by experiental subjects included a preschool autistic class in Honolulu District; a preschool SMH class in Leeward Oahu District; a somewhat heterogeneous primarily TMR, lower elementary class in Leeward Oahu; both SMR and SMH classes in another Leeward Oahu school, and a SMR class in Honolulu District. All were located on public school (DOE) campuses serving primarily regular education children.

Instructional staff in these services were considerably above average in the quantity and quality of their professional training and experience. DOE special education teachers of severely handicapped children were typically recent graduates of (or currently completing) the University of Hawaii's competency-based, two-year master's degree program to prepare teachers of severely handicapped learners. Most of these serving younger children has completed a comprehensive, special preschool training sequence,

and all teachers of preschool handicapped classes were required to have done so. Most of these special education teachers maintained close professional and consultative contact with UH special education faculty (e.g., Voeltz), such that we were well-acquainted with both the teachers and the children in our sample throughout the research efforts. Ancillary professional (particularly OT, PT, Speech) and paraprofessional (EA) staff were also typically involved in providing services to our subjects, though with the exception of the EAs we had less contact with these staff members. By and large, the special education teachers who generally instructed the children were observed throughout our research for all instructional conditions, and were the primary programmers who planned and conducted the intervention studies with us. In some cases, an EA would "monitor" the free play observation condition or would be present in the group condition.

Although all the above settings enrolling experimental subjects were located on the island of Oahu, some were at great distance (more than a half-hour's drive) from the University. This was unfortunately inconvenient for the observations described in the next section, ultimately proved expensive, and complicated our efforts considerably. However, it was also unavoidable, given the nature of the diagnostically heterogeneously, low incidence combinations of educational needs which were the focus of our work. The establishment of special education services for these severely handicapped young children close to their homes is, of course, the intent of P.L. 94-142. Organized research efforts such as ours would be at odds with (recommended) public school programs were we to suggest centralized groupings or placements for our convenience. It was our judgement that meaningful intervention research on children's programs and educational needs had to be conducted in actual educational environments. Since the vast majority of our subjects were appropriately placed (in our professional opinion) in programs designed to meet their educational needs, those programs became our research settings.

Setting and Staffing for Comparison Subjects. The larger group (more than 50 in total across the three years of the project) of "comparison" subjects who were monitored less frequently, though longitudinally for two years or more, were located in a variety of programs throughout the State of Hawaii. These settings included dozens of classes in DOE public elementary schools, DOH Infant Stimulation and Child Development Center Programs, and private agency programs generally contracted by the DOE and/or DOH (e.g., Easter Seals, the Special Education Center of Oahu, United Cerebral Palsy, Salvation Army's Kula Kokua Program, etc.). These programs were scattered throughout the State on both Oahu and the neighbor islands, since, of course, our subjects were again generally receiving services close to home.

Rather than send observers to the programs to code in vivo, these subjects were regularly videotaped in situations with usual instructional personnel and classmates by a University staff member with media experience following a carefully prescribed videotaping protocol. For every taping, either one of the graduate research assistants or a key project staff member (the Coordinator, Evans, or Voeltz) was present to assist with the taping and to consult with program staff. As described in previous reports, we met with children's parents and teachers whenever requested to describe the information we had collected previously and, again if requested, to

provide specific consultation regarding each child's educational needs and program. Our rationale for this approach was that such consultation would typically be available to the students as a function of teacher and parent interest, and thus the "comparison" sample would be distorted were we to suddenly cease to provide it. Although we obviously attempted to provide clinically-sound suggestions, it is doubtful that such episodic advice would have an overwhelming impact upon child outcomes in comparison to the day-to-day influence of each child's program. In fact, most programs did not request such consultation (though many did) and there were no examples of a dramatic change in an IEP, etc., as a function of our consultation. This is, unfortunately, probably true of any "one-shot" intervention by an "expert", which emphasizes again the necessity of preparing teachers to become better decision-makers and to develop skills in using consultation constructively.

Just as comparison subjects were enrolled on a great variety of programs, staff in their settings reflected a wide range of training and experience. In general, the DOE teachers and staff were (like those teaching experimental subjects) recently and well-trained in teaching severely handicapped young children. This included completion of the newly designed, specialized University of Hawaii, Department of Special Education preschool sequence instructed by faculty with expertise in this area (e.g., McCormick, Voeltz). The program model utilized by DOE staff reflected the state's Program Standard Guidelines for handicapped children, which were at the time developmental-behavioral and heavily data-based in orientation. DOH staff were more likely to lack specialized training, though most were experienced in preschool services; DOH staff also tended to be less educationally oriented than DOE teachers. Private agency school staff were generally the least qualified in terms of formal training, although they were experienced. Staff turnover in these services was high, probably due to the considerably lower salaries and the fact that qualified teachers often moved into DOE positions after a brief period of employment in the private agency. In many cases, these staff were more likely than DOE or DOH staff to reflect a strong bias toward a particular program model or philosophy which differed from the state guidelines (e.g., one private preschool was heavily psychoanalytic; another private school adhered to the Judevine Model, etc.).

The Observation Conditions

Observational data were collected for experimental and comparison subjects under three discrete but typical classroom conditions: (1) Situation 1: Free Play; (2) Situation 2: One-to-One Instruction; and (3) Situation 3: Group Instruction. For comparison subjects, the three situations were videotaped, as described earlier, for coding at a later time by the observers in our University laboratory. For experimental subjects, data were coded *in vivo* by the observer pairs who positioned themselves approximately five to ten feet from the child and relevant instructional personnel, so that each observer had a good view of the subject/s of his/her coding system. Each experimental child was observed five times (five separate days) every two weeks. As noted in the previous section, all children were observed in their usual classroom environment, under the actual instructional

and activity conditions which occurred in that setting. For the two instructional conditions, we observed the child with his/her special education teacher.[2] Table 2.2 provides an overview description of the three conditions.

One-to-One Instruction. We found that one-to-one instruction was the typical program delivery format utilized by all the special education teachers for one or more of the instructional programs being conducted with subject children. Thus, we generally had a range of programs which might be the focus of the Situation 2 observation. We asked that either the identical or similar programs be observed, and whenever possible, that this be a table top activity or one in which the child remained seated on the floor or a chair so that our observers would be able to keep the child in full view as much as possible. Thus, we avoided one-to-one instructional situations which involved gross motor movement which would require the child to move from one place to another or otherwise move out-of-view of the observers or the camera. Most teachers had a variety of such relatively "stationary" programs in place, so that this instructional condition was already occurring.

Group Instruction. In a few cases, there were no group instructional programs in place for subject children, i.e., all programs were one-to-one. However, where this was the case, we assisted the teacher in designing an appropriate group instructional session consistent with the IEP objectives of the subject pupil and one or more of his/her classmates. Whenever possible we again attempted to design a group session with the children seated and not engaged in a great deal of physical movement. The group session then became a regular part of the student's education program along with the various one-to-one sessions, with the exception of an unknown number of comparison subject observations, where the group condition might be arranged only for our twice-a-year videotaping; however, even these pupils did regularly participate in a variety of other less structured group activities (snack, field trips, etc.).

Free Play. Situation 1, Free Play, was a more structured version of typical "waiting" periods in these special education classrooms. Typically, since more programs were instructed in one-to-one format by the teacher, the EA, or ancillary staff (OT, PT, speech, etc.), students spent a considerable amount of time engaged in a variety of "holding" activities in-between their turns for that one-to-one attention. In some classrooms, this waiting time was structured as time spent at an activity station staffed by an aide, but in most cases, the child was given access to an appropriate play or sensory stimulation event and loosely supervised for brief periods of time. For our free play observation condition, we arranged with each teacher that the child would be located alone in an enclosed area of the classroom which was nevertheless in full view of the teacher or whoever was responsible for monitoring the child's behavior. This area was generally an already identified "free play" portion of the room, and was typically

[2]There were two exceptions to this: (a) if another staff member (OT, EA, etc.) conducted that particular program, we would observe that instruction; and (b) if the teacher were absent but another staff member also regularly conducted the program, we would observe the latter individual.

TABLE 2.2

The Three Observation Situations

Situation 1: Free play

- Child alone
- Location: defined area, blocked off from the remainder of the classroom.
- Equipment present:
  a. gross motor toy
  b. busy box type toy
  c. small rubber ball
  d. raggedy ann type doll
  e. stacking rings and stack
  f. pull toy with string attached

Teacher/trainer behavior: teacher/aid to minimally supervise child, should interrupt/intervene only if child exhibits dangerous behavior.

Situation 2: Discrete trial, one-to-one instruction

- Child alone with teacher
- Location: child seated kiddy-corner from teacher (or directly across from teacher) at table top surface.
- Equipment present:
  a. table and chairs
  b. materials necessary for task, e.g., (1) fine/sensory motor (puzzle, form board, stacking, etc.), or; (2) cognitive/language (discrimination task, verbal/sign labeling, matching task, etc.)
  c. reinforcers as typical/usual for each child

Teacher/trainer behavior: teacher conducts usual instructional trials, following usual procedures/strategies for that child including use of reinforcers, etc.

Situation 3: Group instruction

- Child, one or more peers, plus teacher
- Location: teacher and children seated in semi-circle across from teacher in chairs, at table or on floor. Target child to be seated to one end of peer group (so that she/he has a peer on one side only)
- Equipment present:
  a. possible table/chairs (or may be floor task)
  b. task items
  c. music

Teacher/trainer behavior: see B above.

Each situation to last approximately eight minutes per child, with brief transition breaks occurring between situations. Three situations consecutively occur for each child, though the order will vary on a regular basis. Data would be collected for one child, and then for the second (and possibly third) child.

partitioned off from the rest of the classroom by waist-high shelving, bolster cushions, etc. A limited number of activity materials (see Table 2.2) were placed within reach of the child; these materials were selected to provide a variety of play options within the child's play capabilities, and were somewhat standardized across children though certain adaptions were regularly made (for example, for older children the materials would be age-appropriate, for certain children a toy which was in "maintenance" phase of instruction might be included, etc.). In several cases where our subjects had extremely limited mobility skills, the child and toys were carefully positioned so that access to each was feasible; in rare cases, this involved seating the child in an adaptive chair with the toys arranged on a large tray attached to the chair. Whoever supervised these free play sessions was instructed not to engage in any interactions with the subject unless this was absolutely necessary, e.g., the child actually ran from the area, engaged in self-abuse, began to destroy materials, etc.

Most of the settings for severely handicapped and/or preschool handicapped children in Hawaii regularly utilized vide taping equipment of their own for purposes of data collection, teacher training, etc., so that the presence of our equipment was minimally disruptive to usual procedures. For the experimental children who were observed in vivo, our observers appeared to "blend" into the classroom environment within a short period of time; they were almost never attended to by the children being observed, with the exception of during the free play condition where a small number of children occasionally attempted to manipulate the MORE and/or interact with the observer/s. Our first observation in the classroom -- when both teachers and children might be most likely to react to their presence -- were actually observation training sessions for the observers so that these data were not included in any of the analyses. In vivo training, in fact, generally extended for a period of approximately three weeks. No doubt the staffing and service delivery structure of the classrooms themselves made it easier for our observations to seem less intrusive than they might otherwise have been. Each of these classrooms typically served no more than four-eight children, and employed a variety of instructional personnal; the special education teacher and the educational assistant were present daily, the Occupational Therapist, Physical Therapist, and Speech Therapist were present once or twice each week in the classroom, and each class was also frequently utilized as a training site so that regularly scheduled university teacher trainees were present. In addition, many of the programs employed a variety of volunteers and other paid staff members, such as Foster Grandparents, parents, and nonhandicapped peer tutors, or "Special Friends". Thus, the presence of our observers became a typical event within the context of a variety of teachers, participants, and other visitors.

Each subject child was observed in all three situations, where the order of the three situations was randomly determined. The randomization was conducted on a regular schedule; e.g., once a month, all ten observation sequences for that month would be generated. This was done in advance so that these "schedules" would be known to the teachers to help them make the necessary arrangements with us, since, for example, each observation

condition required organization of staff, room space, other children, etc. Occasionally, the "random" order was supplanted by unique circumstances and the teacher would than arrange an alternative sequence, and occasionally we would not be successful in observing all three situations (e.g., a major seizure would end an observation early). Each situation observation was eight minutes in duration, which was selected as a time period which reasonable reflected the typical length of an instructional session, etc. Where a session was not exactly eight minutes in duration, of course, our software program which generated percentage durations for each behavior would adjust the denominator function accordingly. Finally, there were brief "breaks" between the situation observations; these were usually only two-three minutes in length, allowing sufficient time to arrange the next situation, etc.

Where observers coded experimental children in vivo, each observer pair typically observed two and possibly three subjects on any given day of observation. The observers generally "dumped" (see Chapter 3) their data for each individual child after an observation and before they began an observation of the next child. This "dumping" time undoubtedly performed a dual function of providing the observers with a break and a change of pace in between these very intensive observation codings. Such changes in pace would be essential for maintaining observer accuracy, since the level of attention required by an observation system as demanding as the one described in the next section is highly vulnerable to observer fatigue.

The Observation System

The Behavior Systems Observation System was designed specifically to collect all potentially relevant data for the research effort (Voeltz & Evans, 1979b). These data include all possible excess behaviors which might be exhibited by the child as well as information on the child's interactions with persons and materials in the environment. Finally, certain "header" information was also recorded for each individual observation session, including basic demographic and descriptive information regarding that child-day's data (e.g., which child was observed, the date, the teacher, the school, the situation order, the intervention phase, etc.). The observation system itself and all procedures for observer training, observing in vivo, and the various data management/analysis programs have been described in sufficient detail and are available so that other researchers might utilize these research tools (Brennan & Freedland, 1980; Evans, Freedland, Lipton, & Voeltz, 1980; Freedland & Brennan, 1982; Voeltz & Evans, 1979b).

Procedures for Collecting Data: Using the MOREs. In the BSIP research effort, two observers watched each child simultaneously and coded his/her behavior according to one of two observational systems known as Observer 1 (coding the excess behavior of the child) and Observer 2 (coding the interactions between the child and his/her environment.) Appendix A contains the complete observation system, including behavioral definitions for each code developed over the three years of the research project. Codes for the child's behaviors and interactions with the environment were recorded on a small portable, solid-state device called a MORE

(Microprocessor Observation Recording Equipment). The MORE has a keyboard with keys for entering data and for controlling certain modes of operation and functions performed by the MORE; up to 9500 characters of data can be stored in memory, and an entire session of data can be regularly "dumped" onto audio tape in a matter of a few seconds so that even this large amount of data capacity does not limit the amount of observation data which could be collected within a single day. Although the MORE is capable of recording data collected through other modes (interval, etc.), the mode used by our system was the Time Event Mode in which the observer enters events as they occur. In this mode, the MORE automatically records that event in sequence and records its duration in seconds until the behavior ceases to occur, as would be signaled to the MORE by the entry of a new "word" which does not contain that behavior. In the Observer 1 System one can enter varying lengths of data up to 10 digits (called "words") which describe what the child was doing at a particular moment in time. If for example the child was happy (code 1), looking at the teacher (code 4) and swinging his legs (code 07), the observer would enter 1407 on the MORE keyboard. If the child continues these behaviors but also begins to talk (code 69), the observer would enter the new word 1407,69. If the observer entered 1469 on the keyboard, this would mean the child has continued the other behaviors but stopped swinging his legs. The Observer 2 System always requires coding in each of five categories and thus each "word" length is always five digits. (More detail on the two observation systems will be presented later in this chapter.)

Because the data from Observer 1 and 2 are eventually combined by computer program for analysis, the observers must synchronize their observations by starting together. To do this, the two MORE's were connected by a connecting cable containing "command switches" to start and end each observation session simultaneously.

Software to accompany the MORE generates a total percentage duration for each individual behavior for each conditions, and also tells how many discrete occurrences of that behavior were entered. Both kinds of information are important, given that duration data is most relevant for some behaviors (e.g., for how many minutes was a pupil "off-task"?) and frequency data is crucial for others (e.g., how many times did a child hit his peer?).

Use of any observation system obviously requries training of the observers in the use of the various codes recorded by the system. This was the case for our research, which involved continuously monitoring an extremely large number of potential behavioral variables. Use of the MOREs was in fact, necessitated by the size and complexity of our observation system since no paper-and-pencil measure would have the capacity of monitoring the number of variables involved in real time as the MOREs enable us to do. However, use of the MOREs required additional observer training procedures in becoming facile in use of the keyboard for the various behaviors as well as in the disposition of the MORE data as it was collected. Training in use of the MORE and the BSIP observation system was organized into five phases which were largely self-paced and required considerable self-instruction; different observers required anywhere from 40-90 hours to become competent in use of the sytem, and it appeared that there was a negative correlation between the number of training hours required and the degree of observer accuracy. Details of the observer training are provided in Evans et al. (1979) but can be summarized as follows:

PHASE 1 - Introductory Meeting

A. The Co-Director provides a brief history and review of project and research.

B. The Coordinator reviews the general requirements for the observer position.

1) training sequence
2) time sheets and training logs
3) schedules
4) professional responsibilities

BREAK

C. The Coordinator reviews observational system, (1st memorization assignment).

D. The project's statistician models use of the MORE

E. The Head Observer leads a tour of the lab.

PHASE II

A. Observers begin to memorize codes and definitions. The system is divided into 5 sections for both Observer 1 and Observer 2. Each section is memorized to criterion prior to beginning the next. The sections are cumulative with a built-in maintenance test.

B. Observers begin to practice using the MORE. A series of three audio tapes play a sequence of numbers at an increasingly faster rate until the speed of coding numbers closely replicates in vivo observations.

PHASE III

A. Observers maintian their ability to recall code numbers and definitions by weekly reviewing and testing.

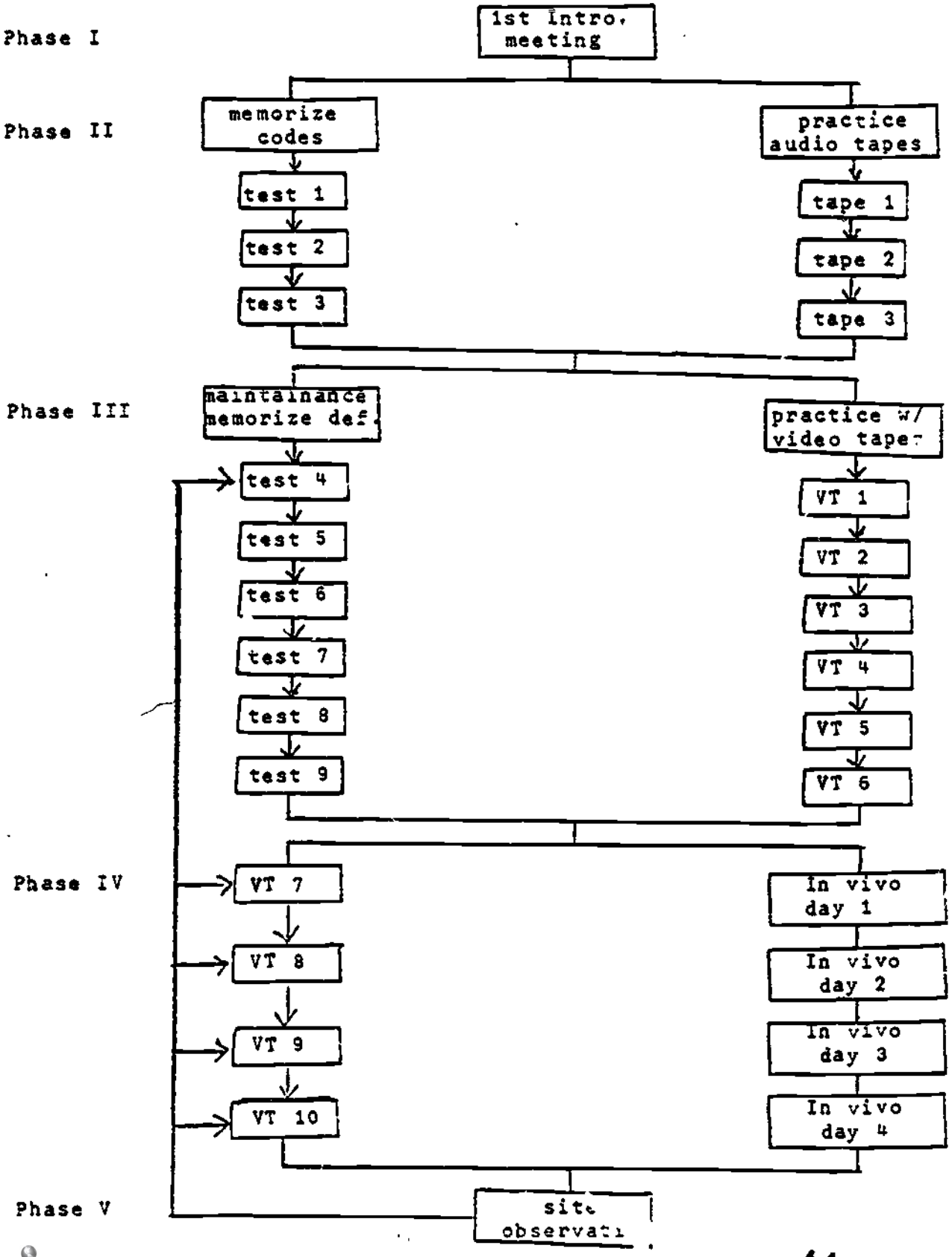
B. A series of videotapes portraying various observation situations in an easy-to-hard sequence are observed and coded. The coding of each tape is reviewed with the observer trainer, the printout is examined, feedback provided and criteria met prior to moving on to the next tape sequence.

Figure 2.1 displays the training sequence graphically. Complete details on observer maintainence training procedures and observer reliability procedures are provided in Chapter 4.

Figure 2.1

Behavioral Systems Intervention Project

Observer Training Sequence



The Observation System

Appendix A contains the Observer 1 and Observer 2 observation systems. The Observer 1 system is designed to monitor pupil affect, attention, and all possible excess behaviors. Both the affect and attention categories reflect exhaustive and mutally exclusive sets of codes, i.e.: (a) affect must be coded for each new entry in one of four possible categories (neutral, happy/excitement, sad/distress, or anger); and (b) visual regard/attention must be coded for each new entry in one of 8 possible categories (space, distant object, proximal object, self, teacher, non-teacher adult, peer/s eyes closed). The reminder of the Observer 1 systems involves monitoring the occurrence of any of 96 individual excess behaviors (including several categories of language behavior). Each individual excess behavior is represented on the system by a two digit number ranging from 01 through 96.

The list of 96 excess behaviors represents an exhaustive set of topographically-defined potential "problem" behaviors. The list was generated over a two-year period based upon all behaviors described in the literature and known to us clinically, and indeed, we have to-date been able to name the code for any behavior observed or reported in clinical studies within this system. The codes are not mutually exclusive, i.e., it is conceivable that one child could exhibit any combination of the behaviors simultaneously (limited only by possible motor limitations) and many of our subjects did so. A few children in our sample actually displayed nearly half of the behaviors on the list within a single day's observation. However, with the exception of certain individually-relevant codes, we conducted subsequent analyses only on those behaviors appearing for at least 4% of the session's duration. Most of our subjects exhibited a range of from ten to twenty or so of the behaviors within each day's data.

The numerical codes (and thus the numerical sequence) of the excess behaviors were for the most part roughly organized into "conceptual" groupings which were inteded to help the observers memorize the Observer 1 system. Thus, the first ten behaviors (01-09) involved gross body and leg movements, the next ten (10-19) involved the hands; however, a few codes not originally included in the schema were added later and thus do not follow this pattern. The Observer 1 system also included two-digit numbers to signifiy the starting and stopping points of any teacher "restorant" procedure which could effectively prevent an excess from occurring; (e.g., the teacher holds the child's hands on his lap briefly); these procedures were specified for individual coding whatever relevant.

The Observer 2 system is designed to monitor general pupil behaviors in environmental context. Briefly, the Observer 2 system includes an exhaustive set of mutually exclusive codes for each of the following:

1. Position in space (lying, sitting, walking, etc.);
2. Child Response (neutral, excess, excess plus, on-task, off-task appropriate, etc.);
3. Objects in vicinity (none, program materials, toys, food, etc.);
4. Teacher response (not present, neutral, neutral plus contact, approval, etc.);
5. Peer Response (not present, neutral, approach, avoid, etc).

These codes represented variables which might be functionally related to the child's excess behavior, and several of the categories also provided us with "treatment integrity" information whereby we could monitor systematically a teacher's implementation of an intervention. For example, we were surprised to find what the teacher positive affect codes for these preschool children averaged only approximately 5% duration across many settings, and we identified only one teacher who consistently scored high (30% or more) on this category. For one child whose behavior appeared to be related to teacher affect, the teacher's intervention consisted of producing high percentage durations of positive affect. The 4th digit code in the Observer 2 system gave us systematic data as to whether the intervention was faithful to this intention. In addition, this general "excess" coding performed by Observer 2 gave us a rough observer agreement check on the behaviors being coded by Observer 1.

CHAPTER 3

Data Gathering and Analysis

In this chapter we shall describe the data collection and analysis procedures developed for this research project. As the standard techniques of direct observational research using either interval or point sampling coding on a paper-and-pencil measure were inadequate to deal with the complex questions of response interrelationships, it was necessary to develop relatively novel and innovative methods. As was explained in the Introduction, the research on interventions with excess behavior has been shaped and is now rather limited by relying on traditional single-subject, single-response sampling procedures. In the same way, decisions we made regarding methods for gathering and analyzing data had a profound impact on the kinds of empirical questions which could be posed and answered. This complex relationship between types of conclusions and types of data-gathering methods is illustrated in Table 3.1. At each stage of the procedure of establishing a technology for data gathering, decisions made have an impact on the overall methodology that is possible. In this chapter, therefore, we will try to evaluate the methods used, indicate clearly the strengths, weaknesses and costs of various procedures, and thereby provide some guidance for future users of these techniques. Thus, the purpose of this chapter is not only to explain the decisions made and methods used by our research project, but also to do so in a way which might assist future researchers attempting to accomplish similarly complex data collection and analysis tasks.

Data Collection and Management

BSIP sent pairs of observers out into each classroom to code the behavior of individual children. Each observer was equipped with his/her own MORE and each used a different coding system. These were called the Observer 1 and Observer 2 systems and have been described in the previous chapter. The clocks in the two MOREs were synchronized so that the two sources of data could eventually be combined on a second-by-second basis. After an observation session, each observer dumped the data for one child for one day onto an audio cassette tape called a file. Later these files were reloaded into the MORE and dumped onto the IBM 360 mainframe computer of the University of Hawaii via a terminal located in the Psychology Department.

As noted in Chapter 2, each child was observed under three different conditions. The data for all three conditions coded by one observer on one day for one child were dumped into a single file called the ODIN file. The conditions (called situations) were demarcated by separate header entries. The dump itself was controlled by a program called JBODIN which saved one file on TSO disk for each observer's data on one child for one day of data collection and produced a hard-copy print-out for the observer to edit. The JBODIN program performed a variety of error checks on the data and a second program, JBINCHK, produced a second printout complete with error checking and warning messages. The observers were responsible for examining these print-outs within 24 hours, correcting any errors and correcting errors which were "legal" errors, such as incorrect codes that they would recall had never occurred for that child on that day. After editing the hard copy print-outs, the changes in the stored files had to be made by

Table 3.1

Methodological Problems in Assessing Behavioral Interrelationships through Systematic Observation and Statistical Analysis of Observational Data

Methodological Issue	Difficulty with Standard Practice	Solution or Alternative Strategy	Remaining Problem/s
1.0 Multiple Response Recording in Real Time 1.1 Simultaneous recording of multiple responses 1.2 Rapid coding of changing multiple responses	Observers can code limited number of behaviors with paper and pencil systems in vivo. Multiple uses of videotaped sessions expands observation capability, but clarity lost on videotape and many behaviors cannot be coded accurately from video. Interval coding systems artificially break up stream of behavior, may not represent reality.	In vivo (use of both of following): Multiple observers, with non-overlapping responsibilities Mechanical data collection devices (e.g., MOREs, Datamyte, etc.) which can code many behaviors in real time	In vivo using multiple observers and MOREs too intensive for many situations/clients. Multiple observers present differing response latencies, so "synching" observation for simultaneity becomes a problem. Mechanical devices present special data loss problems.
2.0 Observation System Design 2.1 Molar vs. molecular; topographical vs. functional definitions, etc.	Use of molar and functionally defined behavior categories requires observer inferences regarding interrelationships prior to data entry. Use of molecular and topographically defined codes ignores clear differences in identically-appearing responses.	Given simultaneous coding of environmental events, molecular categories of child behavior based upon topographical definitions should allow inductive determination of larger categories (including functional), while converse not true.	
2.2 Establishing validity of observational data	Current procedures--based upon single (or few) response monitoring--rely upon observer agreement information, which in turn is evaluated by inter-observer interval-by-interval agreement on behavioral occurrence. Such checks would require perfectly synched observers with multiple MOREs. Furthermore, checks on single response agreement would need to be supplemented by checks on (agreement regarding) clusters and sequences in multiple response monitoring.	Purpose of conducting observer agreement checks is to establish validity of data. Alternative strategies should be explored which more directly address issue. 1. Observer competence retraining and checks; 2. Independent, interobserver agreement across alternating data points (e.g., same conclusions based upon two sets of data?), 3. Across-session interobserver agreement for same session on sequences (Guttman, 1980), not interval-by-interval	Whether such alternative procedures (which violate current convention regarding observer agreement checks as major strategy to establish accuracy) will support validity as well as or better than existing practice, requires empirical investigation. Both logical (Yarrow and Waxler, 1981; Frick and Semmel, 1978) and mathematical (Guttman, 1980) support for alternative approaches must be followed up by comparative studies.
3.0 Data Management Systems for Analysis and Clinical Use 3.1 Storage 3.2 Data Accessibility	Electronic recording services generate large sets of data; the computer storage of which as well as the generation of meaningful summary statistics, etc. very expensive. There are logistical delays in scheduling and transportation between observation and feedback to investigators, teacher-teachers, or other caregivers. Timesharing systems in University computers are heavily utilized, causing delay.	Efficient and cost effective management systems is crucial. A small computer system that can interface with a computer (of University system would save money in initial input and managing of data.	Regardless of solution considerable expertise and money will be required by projects using such systems. Data regarding response frequencies or duration will be more readily accessible than information on clusters.
4.0 Statistical Analyses of Interrelationships 4.1 Appropriate statistical procedures to identify clusters	Variety of parametric and nonparametric procedures exists including cluster analysis, small space analysis, factional, scaling, and factor analysis procedures, principle, oblique, balance or unbalanced solutions, etc. "Acceptable" procedure but clearly established literature shows use of all.	A systematic strategy for searching for interrelationships should begin with a micro factor or small space analysis, followed by selected log analyses based on one's hypotheses	For a number of procedures data recorded in real time has to be broken up again into discrete intervals. Problem of detecting incompatible behavior clusters (factors) results due to characteristics of the statistical technique's algorithm.
4.2 Appropriate input data for factor or cluster analyses (across situations, time)	Statistical analysis procedures such as factor analysis require multiple data points. Though consensus on how many data points is not always clear, the number is large enough to require data from multiple observations (or single subject analyses). Is behavior available across situations, days, etc. to allow for these or other sets of data points as basis for factor analysis?	The answer to this problem actually begs the question of behavioral interrelationships. Issue of behavioral reliability must be addressed by multiple factor analyses for data within and across situations and during different and across all time segments.	Ironically (or perhaps wisely), the accuracy of the combinations of data points may require clinical impression input by teachers and inputs until procedures more clearly established.
4.3 Appropriate shift x time chunks to signal behavioral occurrences for sequential and clustered analyses	Usual procedures of interval coding creates artificial chunks which may create data loss and distort sequences. Dividing "real time" data into artificial time chunks (e.g., 10 second intervals) parallels this problem.	Events signalling combined with time analyses (to provide information on frequency durations of behavior) would most accurately reflect the clusters and sequences intended and coded by observers, coding in real time and marking change by event.	Due to different response latencies by observers, a time window "error margin" is necessary even with event signalling.

a research assistant using the terminal. After this editing process the Observer 1 and 2 files were combined into a single file using the JBCMB program and the combined file saved on a SAS batch disc "save" file. These combined files served as input data for the various programs of statistical analysis. Perhaps the most important of these was the JBALL program which computed, saved and printed descriptive statistics which summarized the codes present in the combined files output by JBCMB. Separate tables are printed out for each category in the BSIP coding system, and each table displayed the frequency, duration, mean duration, standard deviation of the duration, precent frequency and percent duration of each code found in the data. Separate tables for each situation could be obtained and it was these tables which have been used in subsequent analyses presented in this report. A sample JBALL printout is provided here as Table 3.2 (see page 42).

After these steps were completed, a management program (JBMGT) was used for a variety of data steps: The combined files were copied onto a "data" tape and deleted from the disk file, and a backup copy of the data tape was made. This procedure was complicated by the fact that we had to manage several different sources of observational data, from experimental and comparison subjects, from training observers, and from special projects such as reliability studies. An extensive management system was set up to insure that project staff and observers dumped, edited, filed, saved, and backed-up files of data in a timely and orderly fashion. Examples of the forms used to keep track of files and where they were in the sequence are available in a technical manual which describes in detail all aspects of data management and the various programs that had to be developed to control such complex data files (Brennan & Freedland, 1981; Freedland & Brennan, 1982).

Although some of these programs and procedures were modification of the canned data management and software package sold to accompany the MOREs by Observational Systems, Inc., extensive programming was required by our project staff since the available software could not accommodate our needs. It took fully the first year to develop such software and two years to have a reliable management and analysis system in place, which in turn limited our ability to conduct the intervention studies described for the second year until the last year of the project. There were a number of reasons why extensive individual adaptation was necessary: (1) adaptation to the specifications of the local computer environment; (2) the complexity of our coding system exceeded the capacities of the commercially available software; (3) to capture the complexity of response relationships, we required non-rectangular data matrices in which an observer can enter variable numbers of behavioral categories, depending on what the child subject is exhibiting at any one time; (4) the idiosyncratic elements of our coding system had to be represented so that JBALL printouts produced the format illustrated in Table 3.2; (5) the programs necessary to reformat, merge, organize, and maintain the data base were not commercially available because of the wide differences in requirements across projects and computers. The highly sophisticated programming skills these adaptations called for would effectively exclude from practical use microprocessors such as the MORE; it would be most unlikely that a school system, public agency and even many academic environments would have the capability of adapting the MORE software unless their data volume was so small that a MOREs capacities would not really be required in the first place; that is, a paper-and-pencil system would function quite adequately for a system which did not require the kinds

TABLE 3.2. Sample of JBALL Printout, Situations 1 & 2, Child Excess Behavior

I0610021 EXCESS

CODE	FREQ	DUR	MDUR	SD-D	PF	PD	FOR S=1
0	4	17	4.3	2.5	0.036	0.035	(00)NO EXCESS BEHAVIOR
7	16	87	5.4	3.7	0.144	0.178	(07)LEG/FEET SWINGING
10	15	58	3.9	1.8	0.135	0.119	(10)HAND CLAPPING
13	1	3	3.0	.	0.009	0.006	(13)FINGER (HAND) FLICKING
21	9	59	6.6	3.9	0.081	0.121	(21)RUBBING EYES HANDS
23	12	83	6.9	4.0	0.106	0.170	(23)RUBBING FACE NOSE MOUTH
25	15	115	7.7	4.4	0.135	0.236	(25)MOUTHING BODY PARTS
32	3	11	3.7	0.6	0.027	0.023	(32)BITING SELF
65	7	29	4.1	1.5	0.063	0.059	(65)VOCALIZATION SUBGLOTTAL
73	2	8	4.0	.	0.018	0.016	(73)SALIVA SWISHING (DROOLING)
76	10	50	5.0	4.6	0.090	0.102	(76)HEAD WEAVING/SHAKING
86	11	56	5.1	3.4	0.099	0.115	(86)TONGUE MOVEMENTS
89	6	25	4.2	3.0	0.054	0.051	(89)POUNDING ON OBJECT/SURFACE
TOTAL	111	486					

CODE	FREQ	DUR	MDUR	SD-D	PF	PD	FOR S=2
0	2	3	1.5	.	0.012	0.006	(00)NO EXCESS BEHAVIOR
1	3	10	3.3	1.2	0.018	0.021	(01)BODY ROCKING (SITTING)
10	8	27	3.4	1.8	0.048	0.056	(10)HAND CLAPPING
21	2	4	2.0	.	0.012	0.008	(21)RUBBING EYES HANDS
23	5	10	2.0	0.7	0.030	0.021	(23)RUBBING FACE NOSE MOUTH
25	4	15	3.8	2.2	0.024	0.031	(25)MOUTHING BODY PARTS
28	2	8	4.0	.	0.012	0.017	(28)SCRATCHING/PICKING SKIN
32	3	21	7.0	5.0	0.018	0.043	(32)BITING SELF
42	1	16	16.0	.	0.006	0.033	(42)PULLING OTHERS HAIR
62	4	17	4.3	1.0	0.024	0.035	(62)SHRIEK/SCREAM
65	3	12	4.0	1.0	0.018	0.025	(65)VOCALIZATION SUBGLOTTAL
67	12	39	3.3	1.6	0.071	0.081	(67)CLICKING VOCALIZATION
73	13	56	4.3	3.2	0.077	0.116	(73)SALIVA SWISHING (DROOLING)
78	25	105	4.2	2.1	0.149	0.217	(78)HEAD WEAVING/SHAKING
79	13	51	3.9	2.9	0.077	0.106	(79)HEAD DROPPING
86	30	183	6.1	3.9	0.179	0.379	(86)TONGUE MOVEMENTS
89	29	140	4.8	2.6	0.173	0.290	(89)POUNDING ON OBJECT/SURFACE
91	9	31	3.4	0.9	0.054	0.064	(91)JUMPING
TOTAL	166	483					

of modifications we needed to make. This sophistication in initial set-up of a system is not made very clear in the advertising and promotional descriptions of such microprocessor data collection devices.

In addition to the time and effort involved in programming, serious considerations should be given in future projects to the cost of managing such large data sets. A number of programs were re-written to make them more efficient; however the costs for moving from data in the MOREs to permanent tape storage of useable data files was $100 per data point (one observation session) per month. As we were collecting one to two data points per week for approximately 8 experimental children at a given time, this cost ranged from $800 to $1200 per month. The monthly costs of storing data, backing it up, and carrying out management tasks was $300 per month. Statistical analyses were, of course, of average cost, with the most expensive being about $25 to conduct one factor analysis on a number of large data files. At any rate, the cost of more than $1500 per month would obviously be prohibitive to many investigators and was greatly in excess of the amount we originally budgeted for; our costs were covered by departmental allocations of computer time to the principal investigators as faculty members active in research at the University of Hawaii, i.e., we could not have conducted the data collection, storage and analyses required by our project with the computer budget considered to be typical of previous research projects funded by the U.S. Department of Education.

Finally, another major expense with respect to gathering data in the field (and some of our school sites were 30 miles from central Honolulu) was loss of data or an inability to record data due to failure or malfunctioning of the MOREs themselves. At one stage, we lost data on nearly half (46%) of all observation visits made to school classrooms. Reasons for failure to complete an observation included the following:

(1) Logistical problems in the classroom. Some severely handicapped children (not all) have high levels of school absenteeism due to illness, medical consultations, and perhaps home stresses. Children were sometimes present but asleep following a major seizure. Occasionally children were not in the classroom, as arranged, because of field trips, special school activities, assessment sessions, and so forth, but generally those absences which could be anticipated were monitored conscientiously by the teachers who consistently made a major effort to notify observers in advance. Quite often, however, teachers were unable to complete a full session of recording once it had started. This might occur because of a visit to the classroom which required the teacher's attention, a crisis with another child in the room which required the teacher's intervention, and various interruptions involving the target child him/herself, including the occurrence of a major seizure or a major tantrum or disruptive episode which ended the activity. Eighteen percent of lost sessions were for child-related reasons, and 22 percent were for teacher-related reasons.

(2) Problems in coordinating two observers. Because two observers were necessary for each file of data gathered, illness or transportation problems in one observer forced cancellation of the entire observation. Many of our observers were students, who tended to cancel at times when papers were due, exams were given, etc. (Other researchers have commented on this problem which suggests a solution: hire non-student observers who live close to the school where observations are conducted. Landesman-Dwyer reports good results with this observer-hire strategy, and Voeltz used these

criteria to select observers in another research project with excellent attendance results. However, these involved paper-and-pencil systems, and our system required daily travel to the University to dump and edit data; hence, the problem is again complicated by the hardware and software involved).

(3) Observer errors in set-up and dumping. There were a number of places where observers could make errors that would effectively eliminate an entire file of data. These included incorrect header information, switching off the batteries before the data were dumped, and not following dumping procedures correctly. The observers had to carry out this procedure in a secluded area of the classrooms and sometimes tape recorders and power sources were missing or inoperative. Those observers who had been with the project longest made the fewest errors so that some observers, through training and practice, became highly proficient at dumping orocedures. Fourteen per cent of lost data sessions could be attributed to observer error.

(3) MORE failures. Constant use in both field and laboratory settings placed considerable strain on the MOREs. A common problem was wear and eventually breakage of the connecting wire from the battery to the MORE; this we could repair ourselves. During the course of the study three MOREs had to be returned for repairs to basic circuitry. One MORE persistently gave an erratic performance which generated streams of meaningless data; this MORE was returned to the manufacturer for repairs twice. Another common problem was that the MORE would not verify that a dump onto tape had been successful; although the dumps were most often correct, without the verification the observer was forced to keep trying to re-dump until assured that the dump was successful. MORE failures were responsible for most of the lost data files and were particularly costly as they often occurred after the observers had completed the observation itself; this meant not only a loss in terms of travel and payment of observer time, but the loss of teacher effort as the observed session needed to be replaced by another at a subsequent time. Forty-six percent of lost data was due to the MORE failures.

Observer Reliability

A great deal has been written in the behavioral assessment literature regarding observer reliability, but not all of the discussion is relevant to the unique problems created by real-time observation, multiple-observers, and multiple-response recording using microprocessor recording devices. First, the term "reliability", derived from psychometric theory, is a misnomer. When reporting coefficients of agreement between two observers we prefer to use the less ambiguous term of "observer agreement." As has been pointed out many times before, two observers may agree but both be wrong; thus the most desirable quality is observer veridicality or observer accuracy. To measure observer accuracy, one must have an absolute criterion, standard, or master, against which comparisons can be made. We achieved this in two ways: (1) a set of "staged" videotape protocols were constructed, using child actors who depicted the set of behaviors in the observation systems as both single and clusters of behaviors. These protocols not only contained known behaviors, then, but were coded numerous times until a standardized protocol was obtained. Observers were required to meet a certain criterion performance level on each of this series of increasingly more complex protocols during training; and (2) a "master" observer was ultimately identified for each of the two observation systems. This observer was judged to be particularly skilled in all aspects of the observation process, according

to past performance in the laboratory, in field comparisons, etc., and was subsequently used as the criterion observer in field checks on observer agreement as well as in the regularly scheduled "re-training" sessions in the laboratory. We considered these calibration sessions to be crucial to maintaining observer accuracy in the field, and structured them as follows: Approximately once every two to three weeks, each pair of observers[1] assigned to particular children in the field were required to view a "novel" videotape collected for their subject child. They coded this tape exactly as they would in the field, the tapes were dumped immediately, and the two observers viewed the videotape again with their hard copies (print-outs) of the session in hand. They then compared their coding with one another's while reviewing the taping. Evans, Hanashiro and/or Voeltz (at least one and generally two of us) were present at these sessions, and any areas of disagreement were resolved by discussion among key project staff and the observers.

Some investigators report observer reliabilities (agreement coefficients) derived from various assessments throughout the study. While useful for communicating to other researchers the general level of agreement which was achieved, this strategy does not help the investigator during the course of the study. That is, meeting some level of "acceptable" observer agreement such as .80 may satisify other researchers who would otherwise doubt the validity of the results, but a far more useful function to collecting continuous records of observer agreement is to use the information to "correct" observer behavior in future observations and avoid crucial losses of data relevant to the research question. Our major concern throughout the studies reported here was to document observer accuracy in recording single behaviors and clusters of behavior during training and in the field in a manner which would also be used formatively to increase accuracy immediately whenever problems such as observer drift were noted.

Some minimum standard of competence of the observers is required at each stage of training before systematic data gathering should proceed. We also included a period of data collection in the field under actual "criterion" observation session conditions as part of training, i.e., the initial data gathering sessions in the field were actually extensions of training and were not considered "real" data included in the analyses. To determine observer competence, we attempted to analyze the components of the task that was required of them. Observers had to reach criterion of these tasks, so that, for example, their ability to operate the MORE without removing their attention from the target child was measured; their knowledge of the code numbers was determinined by repeated quizzes in which the expected performance was 100% correct (a similar test was used for knowledge of the definitions of the behaviors represented by each code). In addition, observers coded videotapes of the experimental children and could not proceed through training until they had met a minimal requirement of .70 Kappa agreement on all codes. This procedure included provision of printouts of the specific disagreements regarding the application of the various definitions to be resolved. This latter concern is, of course, one of the most significant contributors to

---

[1]These would be both Observer 1s or 2s, i.e., each child was observed by four separate observers. This enabled us to also examine our data for an individual child and determine if child behavior variability might actually be a function of the particular observer-team conducting the coding session.

observer disagreement and knowledge of problems with identifying a particular child's behavior is critical to the investigator at all phases of the study. We refined a number of our definitions over the course of the study, often by adding criteria in which an action had to be repeated a certain number of times before it constituted an excess behavior rather than a random movement. In the case of the child Kathy to be described in Chapter 6) her particular mannerism could be described by three codes: hand clapping, finger flicking, and finger rubbing. Through discussion it was finally decided that one of these would suffice to describe the behavior. As the excess behavior codes were a mixture of functionality (e.g., hitting others, which could be done in various ways) and topographical characterisitics (e.g., head weaving) of the behavior, ambiguity was minimal and there were not many discrete actions that could fit a global definition as is the case if one is coding motor behaviors such as "aggression," "self-injurious behavior," etc. In the case of the Observer 2 system, however, the observer had to make much more complex indentifications of affect and judgments of the functional intent of behaviors; was the child's behavior "on task" or appropriate to the situation? These decisions would vary with the requirements imposed by the teacher, so that less agreement between observers would be expected under these conditions. Finally, looking at agreement on videotapes does not capture one important feature of the task: many of the behaviors included in both systems were subtle responses not readily identifiable on tape due to the inevitable loss of acuity in even excellent recordings, e.g., tongue movements, teeth grinding, etc. Thus agreement checks were regularly carried out in the field at regular intervals with a minimal level of .70 (Kappa coefficients) required on all codes.

Two important considerations must be raised. One is to be concerned with the cognitive nature of the task required of the observers and the other is to come to grips with the issue of what constitutes acceptable or adequate agreement (or better yet, accuracy) coefficients in order to have some faith in the quality of the data. The first of these questions is the easiest to address. The MORE does not operate like an event recorder in which response durations are determined by the onset and offset of a marker (switch, push button, or whatever). In the real time mode, the MORE's internal clock times the duration of an event from the moment of its entry on the keyboard until a second entry describes a new set of behaviors by the subject, so that ongoing behaviors must be re-entered on the record entry along with something new or the ongoing behavior would be considered terminated at that point. Thus, if the observer notices a new behavior beginning, s/he must enter a new "word" which includes that new behavior as well as the codes of all the other behaviors previously entered which are still ongoing. Similarly, if a behavior stops occurring, the observer must enter a new line listing all other behaviors still ongoing and omitting the terminated behavior. Clearly, the observer must not only be vigilant and notice occurrences of new behaviors (and whether they match particular codes, etc.) but must also notice that ongoing behaviors have ceased, what other behaviors were previously entered so that they will be repeated in the new entry if ongoing, etc. Our observation system is quite different from existing systems, then, in requiring continuous monitoring of the child's total repertoire as well as certain environmental occurrences, and the demands made upon and the judgments required by observers are both multiple and complex.

In order to see whether the observers were equally good at noticing behaviors which were being added to or deleted from the ongoing repertoire of perhaps two or three other behaviors, a brief study was conducted with

the observers. A videotape was made of two child actors who were instructed to perform a series of excess behaviors simultaneously and then either add or drop one or two behaviors from this ongoing repertoire. Observers coded this videotape and errors were compared against the scripted excess master. The results of this trial showed that although all observers made some error in the exact moment of time that they noted onsets and offsets of behavior (an observer "latency" issue), they were equally observant of either type of change in the behavior stream and were able to yield overall durations of each behavior that were very close to the actual durations.

The second question posed earlier in this discussion was how one determines an adequate degree of agreement among observers; this question must be expanded to include the interval at which point-by-point similarity is being sought, and the complexity of the codes used. Obviously, if two observers are coding 50 different categories which are analyzed in the 1-second interval chunking systems--the resolving interval of the MORE--point-by-point similarity is likely to be less than if the pair were coding three behaviors at 10 second intervals. This is not just because the former task is cognitively more complicated, but because the longer interval allows a larger margin of error. For example, if one observer consistently notices behavior change 1 second later than another, they may both be entirely accurate in observing the behavioral occurences and total duration of occurrence, but their "agreement" will be seriously affected on a second-by-second analysis; this difference in response latency by observers is unlikely to be a problem in a 10-second interval system. Thus, minor variations in latency of observers' coding responses will prooduce considerable discrepancies unless agreement is based on a "moving window" in which the comparison observer's record is systematically shifted forward or backward, interval by interval. Instead of using this strategy, we estimated agreement on data streams which were re-chunked into five-second intervals as this was the interval at which we were carrying out factor analyses of the data.

This procedure was based on the logic expressed--quite superbly--by Yarrow and Waxler (1979). They argued that the level of analysis of observer accuracy (agreement) should be at the level at which the dependent variables are being analyzed in order to reveal agreement on the phenomenon of interest. Or, put another way, the issue is the degree to which the findings and conclusions based upon the data source are similar to findings and conclusions based upon another data source, including data collected during the same time period (an observer agreement issue) and for the same child at different time periods (a more traditional, data reliability issue). Gottman (1980) made a similar argument with respect to sequential analysis: Since the goal of sequential analysis is to detect sequential relationships, the criterion of observer reliability should not be based on a point-by-point agreement on the occurrence of individual behaviors, but rather on the extent to which two independent observers produce data that yield similar sequential structures. In order to plan for this general strategy, we arranged for pairs of observers to alternate their data gathering sessions, so that the data gathered (or at least the summary conclusions reached) by one pair could be compared to the data gathered by the other pair for each individual child. The disadvantage of this procedure was that when slight systematic differences between observers did exist, these differences were not dispersed randomly through that data; sometimes, patterns of response durations over time would thus seem to represent differences between observers in the application of a code's definition. These patterns would not

necessarily be problematic in a research project as long as they are consistent across phases of the investigation: They become a problem when data within a phase contains variability of applications of the code definitions. The other disadvantage of this procedure was that the summary findings (factor analyses, in this case) were much less stable than we had originally assumed they would be; as this seems to represent the large variations in daily occurrences of many excess behaviors, it means that trying to establish observer agreement by comparing alternating pairs of observers is confounded by the intrinsic lack of reliability ("test-retest reliability" for a child) in the occurrence of the behavior.

Brief mention must be made of the actual statistic of agreement that should be reported. We routinely calculated conventional agreement indices (agreements divided by agreements plus disagreements) and Kappa coefficients, which generally tended to be lower as they take chance agreement into account. However, in a "Monte Carlo" study of Kappa, we ascertained that even when actual agreement is known (say 50%), the Kappa coefficient could vary from .00 to .44 depending on the number of behavioral categories and the duration of the session. Also, if the observers are always cognizant of some 96 possible excess behaviors, their agreement regarding the non-occurrence of many of these is of importance. But the reality is that observers become accustomed to a child's predominant set of behaviors and are likely to miss the occasional occurrence of rare behaviors. What we really need to know is the probability that the observer did not miss any occurrences of a behavior or did not report any false alarms.

Issues in Data Analysis

As we discussed generally in our review paper on response interrelationships (Voeltz & Evans, 1982), the concept of response organization covers a number of models of actual response relationships. Clusters, for example, may be thought of as concurrent events or sequential events (including distant sequences, such as lags of some number), or possibly "anti-clusters" where the occurrence of one behavior reduces the probability of a second behavior occurring regardless of whether that behavior is physically incompatible). Obviously, whether two events are treated as sequential or concurrent depends on the interval within which they are coded and analyzed and the respective durations of the behavioral events. If only two or three behaviors are being considered, the clustering issue can be stated in terms of conditional probabilities--given the occurrence of Behavior A, what is the likelihood of Behavior B co-occurring (lag 0) or occurring immediately after (lag 1 event) or occurring within a certain time after (within a particular lag-range in seconds)? However, when a number of behaviors are considered simultaneously and no one behavior is known to be the primary event, and when co-occurrence is only within longer time intervals, a method of describing general clusters seems necessary to describe response relationships and how they might change over time.

Until recently, the typical method for determining such relationships was to examine the covariation in behaviors over fairly long periods of time. For example, Voeltz and Evans (1979) examined the correlations between behaviors over a number of observation periods, specifically across approximately 100 successive days of data collection over a period of four months for one child. This method establishes that on occasions (particular daily observation periods) where one behavior is relatively more (or less) frequent than usual, another behavior will also be more (or less) frequent

than usual. There are certain limitations to this procedure, the most obvious being that variations in two behaviors may be due to unmeasured third events (variations in general activity level cause increases in both "running" and "finger flicking", though the two behaviors are not at all related to one another). Thus, apparent "relationships" between two behaviors do not, in fact, reflect any internal organizational structure. There is also a mathematical limitation in which two behaviors which have quite different patterns of variation but which have a common trend (such as both decreasing over time) will correlate; thus, it is necessary to de-trend each of the time series and analyzes them according to concomitant time series analyses. Further discussion of previous research as a function of whether across session and time total percentage occurrences vs. point-by point occurrences were used as the basis for analyzing of response inter-relationships can be found in Voeltz and Evans (1982).

We assumed in our original proposal that no descriptive statistical technique for depicting inter-relationships among behaviors would suffice to reveal internal response organization. We therefore planned a series of studies in which experimental interventions would be used to determine the true nature of the relationships, such as manipulating one element of the behavioral cluster and seeing what happended to response inter-relationships (as well as the individual behaviors) thereafter. As will be seen later when sample reports of studies with the children are presented, we were only partially successful in implementing this experimental strategy. One reason for this is that a descriptive model of response relationships that reveals reasonably stable clusters needs to precede any such analysis. In other words, a dependent variable must be established which reflect in a reliable way the phenomenon of interest--in this case, some kind of cluster unit which can then be subjected to further quantitative analyses.

A number of possible clustering techniques were explored, and the results reported in summary form by Evans, Voeltz, Freedland, and Brennan (1981); Tables 3.3 through 3.6 and Figures 3.1 through 3.3 display the results obtained for one of our experimental children's behavioral repertoire as a function of the alternative analysis strategies used. The issue requires that a series of decisions be made early on in data analysis, and our results indicate that these decisions significantly impact the observed results. Our real-time observational and data recording techniques provided us with information on response durations and their onset and offset points relative to one another. Any kind of correlation-based cluster technique requires some metric for each response across some common time interval. If we wish to obtain the cluster within observation sessions, this time interval must be less than the duration of the session and have sufficient occasions to generate stable correlations. The ideal interval would appear to be the smallest detectable, in this case the one-second intervals, and the response metric within that interval must then be dichotomous--the behavior did or did not occur during the interval. If, however, a large interval is used, the metric could be duration within that interval, although in practice if a short enough interval is chosen to provide sufficient number of occasions per session, the variations within intervals are not large and essentially one again has dichotomous data. Thus, once we decided on the appropriate interval, data were re-formatted into a dichotomous form so that interval by interval each response was shown as occurring or not occurring.

Another decision to be made in this respect is whether the factor

Table 3.3

Excess Behaviors for Child 01

#	Descriptor	Definition
08	Bear Walking	feet and open palm on surface, bent at waist, propelling self using feet and hands, 2 or more cycles, or standing in position for 2 seconds or more.
10	Hand clapping	pounding or clapping of fists or hands together.
13	Finger flicking	repetitive finger (1 hand) movements, one or both hands, child may/may not watch hand/s.
23	Rubbing face, nose, mouth	rubbing fingers/hands across facial area (from ear to ear, and top of forehead to chin) in more than one cycle; or holding hand/finger/s in contact with face for more than 2 seconds.
25	Mouthing body parts	moving lips on and/or over a body part, inside or against mouth; body part must be visible or body part covered by clothing (e.g., foot in sock, shoulder in shift, etc.).
36	Head banging	banging of head against an object or person.
54	Object flicking	manipulating an object rapidly in "fluttering" motion, movement in wrists/elbows/finger joints.
64	Blowing	blowing air out of mouth with/without saliva and/or tongue protrusion.
72	Staring	holding a fixed, glassy-eyed look for more than 3 seconds.
79	Headdropping	head drops abruptly forward or backwards, chin toward chest or neck stretched, facing ceiling.

Table 3.4

Three Factoring Solutions for Child 01 Behavior:
Free Play Condition

Factor	Unbalanced Oblique 10 second intervals 3 days		H-Corr Oblique 10 second intervals 3 days		H-Corr Oblique Total Session % Duration 40 days	
	Behavior	Loading	Behavior	Loading	Behavior	Loading
1	Finger flicking	.74	Finger flicking	.97	Head drop	.95
	Mouth body parts	.56	Mouth body parts	.74	Staring	.84
	Blowing	.32	Blowing	.69	Hand clapping	.82
	Rub face/nose/mouth	.30	Object flicking	.50	Cry (tears)	.73
	-Head drop	-.43	Rub face/nose/mouth	.48	Drooling	.55
	-Subglottal voc	-.43			Head banging	.33
2	-Head drop	-.35	Hand clapping	.90	Object flicking	.88
	Hand clapping	.79	Pounding objects	.51	Blowing	.81
	Hand/arm flap	.42	Subglottal voc	.51	Rub face/nose/mouth	.79
			Hand/arm flapping	.49	-Cry (no tears)	-.33
3	Head banging	.73	Head banging	.74	Finger flicking	.68
	Cry (no tears)	.60	Cry (no tears)	.62	Bear walk	.53
					Mouth body parts	.50
					Subglottal voc	.43
					Hand clapping	.30
					-Head banging	-.50
					-Cry (no tears)	-.46
					-Scratching	-.37
4	Bear walk	.35	Object dropping	.76	Hand/arm flap	.72
	-Smelling objects	-.71	Bear walk	.44	Smelling objects	.63
					Finger tapping	.46
					Drooling	.42
					-Object flicking	-.29
5	No excess	.50	Head drop	.71		
	Supraglottal voc	.48	Subglottal voc	.35		
6	Bear walk	.52	No excess	.61		
	Object dropping	.49	Supraglottal voc	.46		
	Smelling objects	.40				
	-Object flicking	-.45				
	-Blowing	-.37				
	-Head drop	-.36				

Table 3.5

Three Factoring Solutions for Child 01 Behavior:
One-to-One Instruction Condition

Factor	Unbalanced Oblique 10 second intervals 3 days Behavior	Loading	H-Corr Oblique 10 second intervals 3 days Behavior	Loading	H-Corr Oblique Total Session % Duration 40 days Behavior	Loading
1	Finger rub	.67	Finger rub	.85	Blowing	.98
	Hand/arm flap	.55	Hand/arm flap	.55	Rub face/nose/mouth	.85
	Leg/feet swing	.57	Leg/feet swing	.46	Finger flicking	.81
					Mouth body parts	.75
					Hand clapping	.67
2	Pounding objects	.74	No excess	.93	-Hand clapping	-.31
	Leg/feet swing	.49	Object dropping	.39	Subglottal voc	.77
	Finger flicking	.34			Hand/arm flap	.74
	-No excess	-.62			Object flicking	.65
					Smelling objects	.60
3	Head drop	.77	Head drop	.85	Cry (tears)	.90
	Object flicking	.39	Object flicking	.37	Head banging	.59
					Head drop	.45
					Object flicking	.38
					Cry (no tears)	.37
					Bear walk	.33
4	Mouth body parts	.37	Rub face/nose/mouth	.82	Finger tapping	.73
	Head drop	.37			Cry (no tears)	.61
	Finger flicking	.37			Scratching	.38
	Object flicking	.36				
5	Mouth body parts	.44	Pounding objects	.69		
	Blowing	.40	Finger flicking	.60		
	Subglottal voc	.35	Mouth body parts	.51		
	Hand/arm flap	.34	Hand clapping	.46		
6	Hand clapping	.57	Blowing	.62		
			Hand/arm flap	.33		
			Finger tapping	.33		
			Mouth body parts	.31		

Table 3.6

Three Factoring Solutions for Child 01 Behavior:
Group Instruction Condition

	Unbalanced Oblique 10 second intervals 3 days		H-Corr Oblique 10 second intervals 3 days		H-Corr Oblique Total session % duration 40 days	
Factor	Behavior	Loading	Behavior	Loading	Behavior	Loading
1	Rub face/nose/mouth	.89	No excess	.37	Finger tapping	.95
	Blowing	.83	Finger tapping	.35	Hand/arm flapping	.81
	Mouth body parts	.74	Pounding objects	.33	Mouth body parts	.36
	Object flicking	.53				
	-No excess	-.40				
2	Finger flicking	.97	Mouth body parts	.77	Drooling	.88
	Mouth body parts	.31	Object flicking	.69	Head drop	.77
	-Subglottal vocal	-.35	Rub face/nose/mouth	.62	Subglottal vocal	.57
			Finger flicking	.53	Cry (no tears)	.43
			Blowing	.49		
3	Hand clapping	.89	Hand clapping	.83	Scratching	.91
			Leg/feet swing	.56	Smelling objects	.82
			Finger flicking	.39	Rub face/nose/mouth	.35
4	Pounding objects	.68	Blowing	.60	Blowing	.97
	Hand/arm flap	.66	Rub face/nose/mouth	.55	Rub face/nose/mouth	.67
	-No excess	-.60	Hand clapping	.37	Object flicking	.47
					Finger flicking	.38
					Subglottal vocal	.34
					-Smelling objects	-.30
5	Blowing	.51	Finger rub	.90	Bear walk	.86
	Rub face/nose/mouth	.45			Head banging	.78
	Pounding objects	.40				
6	Finger tapping	.70	Hand/arm flap	.96	Cry (no tears)	.78
			Pounding objects	.74	Staring	.69
					Cry (tears)	.43
					-Hand clapping	-.45
					-Finger flicking	-.44

Figure 3.1

Two-Dimension, Nonparametric Small Space Analysis
for Child 01 Behavior (10 second intervals)
During Free Play Condition

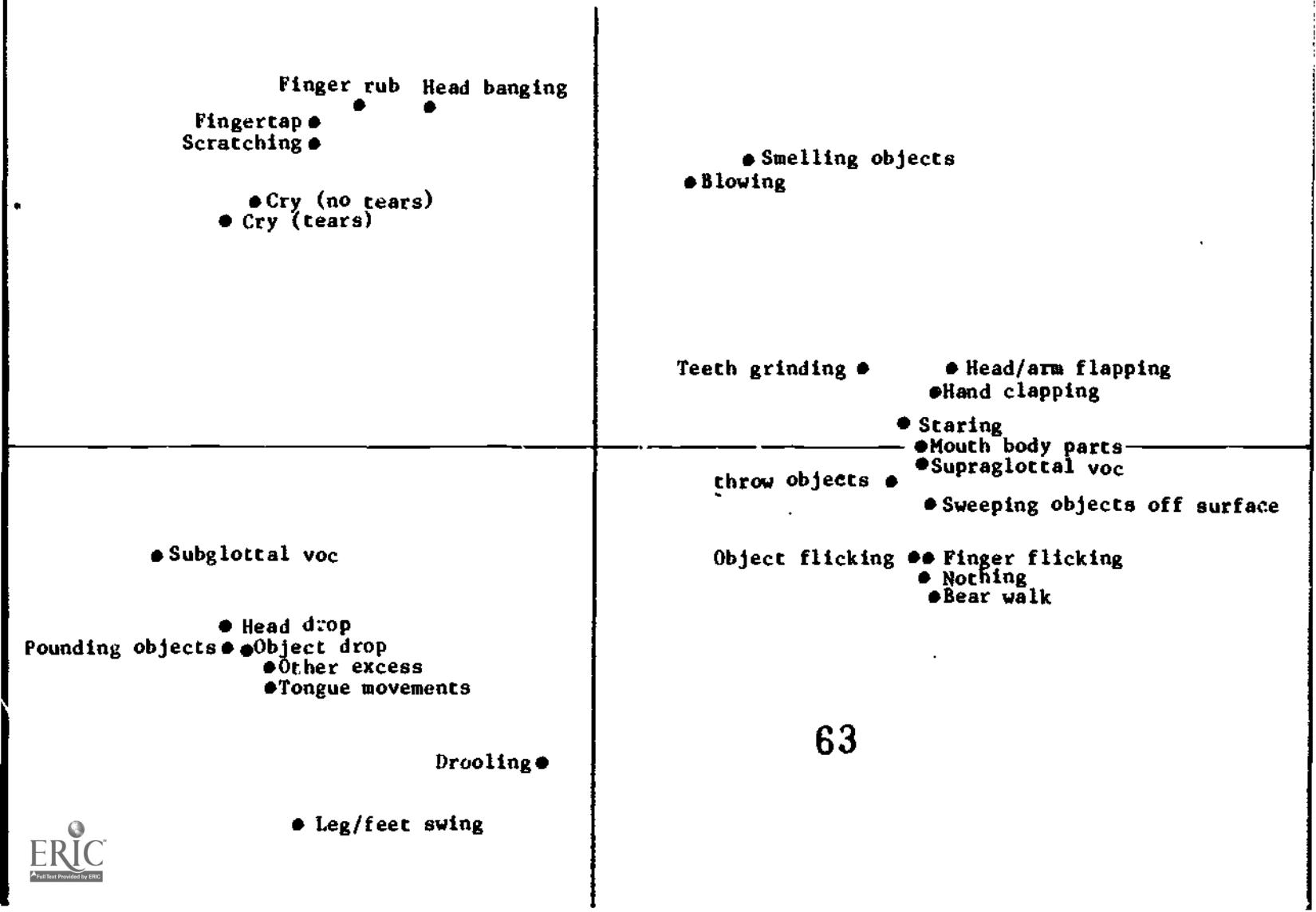


Figure 3.2

Two-Dimension, Nonparametric Small Space Analysis
for Child 01 Behavior (10 second intervals)
During One-to-One Instruction

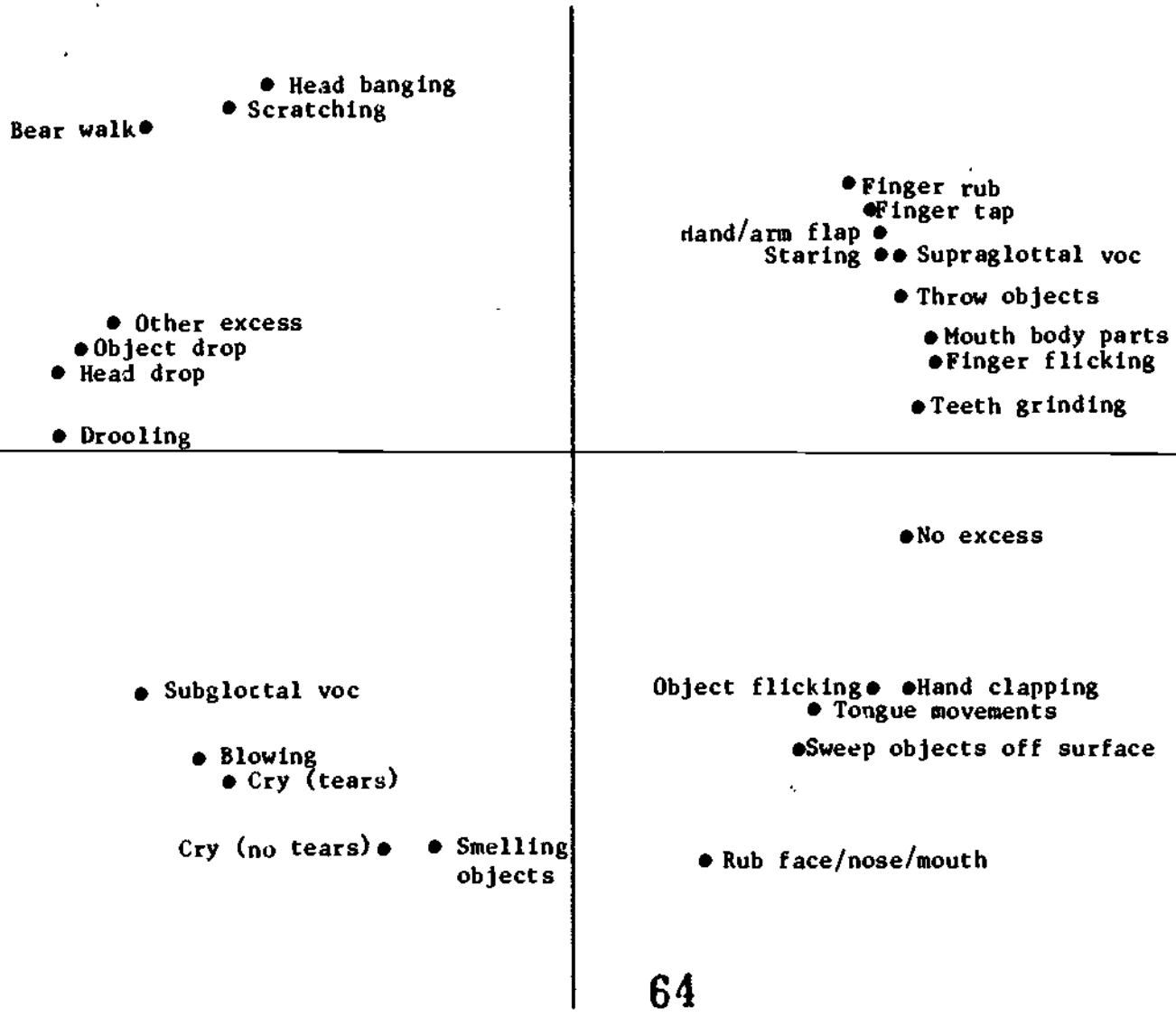
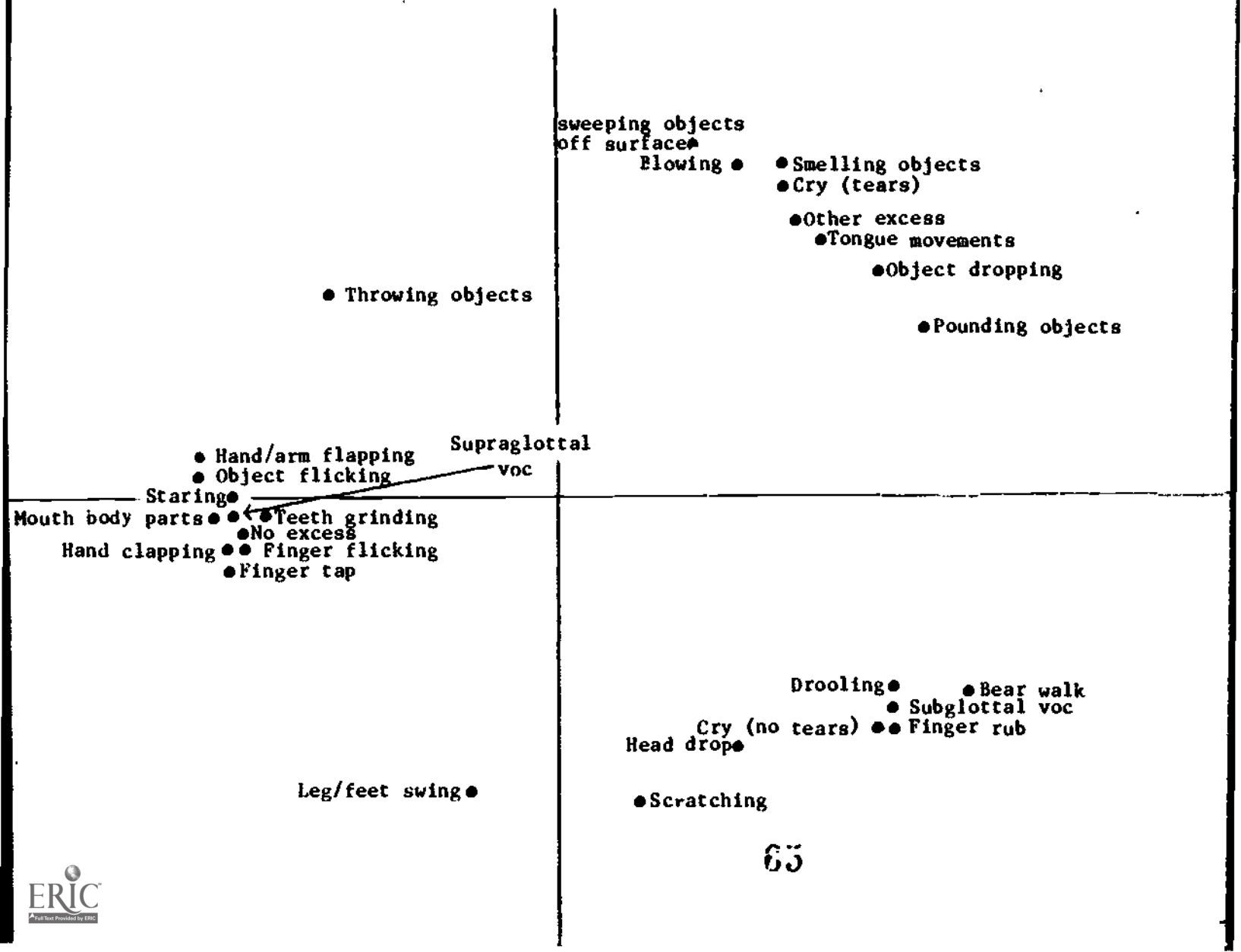


Figure 3.3

Two-Dimension, Nonparametric Small Space Analysis
for Child 01 Behavior (10 second intervals)
During Small Group Instruction



solution should be derived from oblique or orthogonal rotations. We generally have assumed that clusters should be considered to relate to each other. But oblique rotation may tend to produce higher-order factors reflecting relatively uninteresting dimensions such as levels of activity. In the end we decided on certain compromises: choosing 5-second intervals and looking at oblique rotations: using a balancing technique called H-Corr in which a mirror-image of the dichotomized variable is set up and included in the overall analysis. We also aggregated three observation sessions per factor solution, so that if each session was 8 minutes, we based the analyses on 24 continuous minutes or 288 5-second intervals; in several of the intervention studies, we collected "pairs" of three-session data points, computed factor analyses for each of the sets of sessions, and used a factor analysis congruence testing procedure to determine whether the factors were reliable within an intervention phase.

The problematic feature of these various decisions is, as can be seen from Tables 3.4-3.6 and Figures 3.1-3.3, that rather different solutions do emerge when seemingly minor changes are made in method or the interval of analysis.

Conclusions and Summary

The emergence of micro-computers and microprocessors and the opportunities these afford for multiple response recording in real-time modes has outstripped the available methodology for summarizing and analyzing these large data sets. Methods derived from the operant conditioning laboratory where the frequency of only one or two responses was plotted over time are clearly quite inadequate for the 1980s, both in terms of their limiting effects on clinical and educational research designs (see also Voeltz & Evans, in press) and in terms of the basic issues of veridical representation of behavior. Thus issues of observer reliability, type of coding system, molar-molecular level of analyses. selection of which behaviors to record, units for oresenting data, and modes of analysis are all outdated in the methodological discussions that are still common in behavior modification and behavioral assessment.

With the value of hindsight, we now see how much more background work was required in this area than originally conceptualized in the research proposal. New methodologies open new possiblities for discovery, but also create their own set of novel pitfalls for the investigator. Many of these problems were overcome or at least reasonably defensible decisions were made, as can be seen from the quality and richness of data presented later in this report. On the other hand, many basic questions remain. We feel that a productive and worthwhile technological advance has been made in this field which will have considerable impact on the work conducted by other investigators. Conversely, we did not develop methods which were able to support all the investigations planned in the original proposal. By outlining the difficulties involved and some of the solutions utilized in our work, we intend to provide subsequent efforts with a realistic and informed perspective which will allow work to proceed without addressing identical elementary concerns without the benefit of past experiences.

CHAPTER 4

Decision Making

Caregivers, psychologists, teachers, and other professionals and advocates must assume the responsibility of making decisions with far-reaching consequences on behalf of severely handicapped children. These include decisions to intervene, decisions how to intervene, and the implied decisions not to intervene whenever potential target behaviors are tolerated, ignored, and/or monitored while other target behaviors are programmed. These various decisions occur informally on a day-to-day basis, and formally on a year-to-year basis in conjunction with a child's Individualized Education Program (IEP), yet the consequences of the choices made have a cumulative impact. Severely handicapped children exercise limited control over their environment, and thus are more dependent than their nonhandicapped peers upon the choices made on their behalf by their parents and the professionals who work with them throughout their school years. At the same time, severely handicapped children present these caregivers with an extremely complex task: selecting priority goals and intervention strategies for instruction from among the many potential goals and strategies which could be utilized. The task increases further in complexity because some degree of formal intervention is needed for each skill acquired (unlike nonhandicapped persons who acquire many skills outside the context of a formal educational program) and because empirical information regarding which goals should be highest priority and which interventions will be most effective is not yet available (see Chapter 1).

Ideally, then, teachers' and caregivers' decisions on behalf of severely handicapped children will be empirically based on the accumulation of knowledge regarding the immediate, indirect, and long-term effects of specific interventions and programs. Some of the needed empirical information is already available: Voeltz and Evans (1982) have, for example, summarized the available evidence of response interrelationships which might allow teachers to eventually identify one goal rather than another because of the known benefits vs. costs associated with either choice. Much information is nevertheless missing because intervention researchers have generally failed to consider the issue of response interrelationships seriously and have not monitored multiple effects systematically. Derer and Hanashiro (1982) have abstracted over 235 studies and specifically attempted to identify particular strategies which might be clearly associated with successful outcomes across children for particular behaviors. Again, results are inconsistent throughout the intervention literature and no such clear patterns emerge. Butler and Stenmark (1981) rightly emphasized that the mandate for IEPs should now be associated with efforts to document children's outcomes as a function of various program components. Such efforts remain relatively rare (cf. Maher, 1982), but we could logically expect our knowledge of beneficial intervention-outcome relationships to increase over the next few years, so that eventually teachers' decisions can be truly "data-based."

In the interim, however, teachers and other interventionists must continually make decisions and implement particular programs where empirically-based data are lacking or inappropriate to an individual child's problems. Thus, there do exist examples of rational and ethically based guidelines for teachers to use in selecting goals and intervention strategies (e.g., Brown, Branston, Hamre-Nietupski, Pumpian, Certo, & Gruenewald, 1979; Gaylord-Ross, 1980; Heads, 1978; Nelson & Hayes, 1979; Stolz, 1978). Yet each of these guidelines reflect a "Delphi" approach to the issue, wherein a single expert

or groups of experts have generated their recommendations based primarily upon a personal/professional interpretation as to how such decisions should be made, rather than upon a model of how such decisions are actually made by interventionists or empirical evidence that the proposed model will result in the most beneficial outcome.

The decision model proposed by Voeltz, Evans, Derer, and Hanashiro (1982) and detailed in our training manual (Evans, Derer, Voeltz, & Hanashiro, 1982) is an effort to incorporate relevant expert opinion and available empirical information into a process which realisitically reflects the way in which teachers and other caregivers can and do make such decisions. We have already summarized the currently recommended approaches and the limited empirical input (see Chapter 1; also Voeltz & Evans, in press), and Chapter 6 will present additional information obtained from several of our intervention studies. This chapter presents the results of several studies designed to determine: (1) to what extent teachers and other decision makers currently utilize available guidelines; (2) if not, why not; and (3) what the actual decison processes exhibited by teachers does look like. This information was then synthesized as the framework for a model which teachers could apply to make optimal choices--based upon "best practices," available empirical information, and information on how teachers actually make decisions--when planning and conducting children's programs.

Study 1: A Comparison of Decision-Input Rules Utilized by Various Professionals and Laypersons

Subjects

Four groups of individuals were subjects in the first study. The first group consisted of 36 special education teachers and graduate level teacher trainees enrolled in an advanced behavior modification methods class; all were certified teachers with one or more years of special education teaching experience. A second group consisted of 101 educational assistants (EAs) employed in Hawaii public schools and attending a one-day paraprofessional training workshop on behavior management; this group actually comprised the majority of EAs on the island of Oahu, Hawaii, serving children with a variety of educational needs, from remedial classes for mildly handicapped to self-contained classes for severely handicapped youngsters. A nonspecialist third group consisted of 61 undergraduate psychology students who had had no course work or formal training in childhood handicaps, behavior modification, or special education. 21 clinical psychology trainees comprised the fourth group; this group included all those in the clinical psychology doctoral program who had had internship or practicum experiences with behavior problem children but, in most cases, little direct experience with young handicapped children. All student participants were attending the University of Hawaii, and subjects were representative of the heterogeneous ethnical and cultural backgrounds characteristic of the population of the State of Hawaii.

Materials and Procedures

A list of 17 items (see Table 4.1) was constructed, representing a compendium of those reasons or criteria most often offered clinically and in the literature (Heads, 1978; Nelson & Hayes, 1979) as justifications for selecting intervention target behaviors. This list was intended to include all

nontrivial considerations that teachers might weigh in making decisions; the participants were encouraged to list at the end any "additional reasons or criteria which may be missing from the list, but which nevertheless seem important to you;" the very few additions represented variations of reasons already included in the list.

Each item was followed by a continuous line, subdivided into 20 segments for rating on a scale from 1 to 20. Five major rating categories were specified beneath the scale at equidistant intervals: very important consideration, important consideration, undecided, not an important consideration, not a consideration. A sample item and rating were provided to clarify the task, and participants completed the ratings in groups. Written instruction (plus verbal instructions to the EA group) indicated that: (a) the various reasons had been suggested as good criteria to use in identifying a target behavior for intervention in educational programs for children; (b) the reasons would have different applicability to different kinds of behaviors (i.e., some applying to curriculum objectives and others to negative or problem behaviors); and (c) participants should read over the entire list prior to rating individual items, since their individual ratings should reflect the relative value attached to each reason. All participants also completed a checklist of professional or practical experiences, so that numerical estimates could be made of the extent of their relevant experience with handicapped children.

Results and Discussion

Major results are summarized in Tables 4.1 & 4.2 , providing the most detail regarding the responses of the teacher group as those individuals most likely to be currently involved in making programming decisions. Table 4.1 thus includes the mean ratings and standard deviations for the teacher group only, and the ranks of the ratings for the items by all four groups, while Table 4.2 displays the factor analysis for the teacher ratings. Additional results and a more detailed discussion of these data--with the exception of the EA ratings--can be found in Voeltz, Evans, Freedland, and Donellon (1982).

As can be seen from Table 4.1, there was considerable agreement across the four groups in their relative rankings of the various items, with several exceptions. If a difference of four ranks represents a significant relative discrepancy in judged importance, the educational assistants differed from the other groups--particularly the teachers--on items 14 (dangerous to child) and 11 (increase in community acceptance). They rated the item dealing with behavior leading to increased community acceptance highest, while this item was ranked considerably lower by teachers (8th) and somewhat lower by undergraduates (5th) and psychology trainees (6th). On the other hand, while all other groups ranked the item concerning a behavior dangerous to the child as either highest or second highest in importance, the EAs ranked this item 5th. These discrepancies would be explained by the fact that a majority of the EAs were employed in classrooms serving mildly handicapped children, where the problems are more likely to consist of "typical" childhood deviance/acceptance behaviors and self-abuse by a child may not even occur. EAs differed markedly from the undergrads on item 8 (replace existing negative behaviors), which the latter group ranked highest in importance while the EAs ranked this item 7th.

The responses of the undergrads were expected to be somewhat representative of laypersons or parents, yet this nonspecialist group differed from the teachers and the other groups on only one item, item 12 (behavior is danger-

TABLE 4.1

Summary of Ratings of Importance of Criteria for Selecting Target Behaviors

Item #/decision criteria items	Teachers' mean ratings(SD) (N = 36)	Teachers' rank order (N = 36)	Ed. Assts. rank order (N = 101)[a]	Undergrads' rank order (N = 61)	Clinical Psychology rank order (N = 21)
14. The behavior is dangerous to the child.	19.86(.53)	1	5	2	1
12. The behavior is dangerous to others in the child's environment.	19.42(1.44)	2	3	8	2
6. The behavior may interfere with learning unless it is modified.	18.50(1.63)	3	2	3	7
8. The behavior would replace existing negative behaviors with a positive alternative.	17.97(2.31)	4	7	1	4
9. Attainment of the behavior would increase the child's independence.	17.92(2.85)	5	4	6	9
1. The behavior would be immediately functional for the child.	17.58(3.17)	6	9	9	8
17. The behavior is a prerequisite to learning other adaptive behaviors.	17.42(2.29)	7	8	4	5
11. The behavior would increase acceptance of the child by parents, teachers and peers.	17.31(2.61)	8	1	5	6
15. The behavior is one which would broadly affect the child's repertoire, i.e. positive collateral or side effects are likely to occur in more than one area after intervention.	16.50(2.56)	9	10	10	3
5. The behavior is a major concern for the child's parents/caregivers.	16.19(2.41)	10	6	7	10
7. The behavior is an appropriate activity which the child would probably enjoy being able to do.	15.81(2.74)	11	12	12	14

(TABLE 41 Continued)

Item #/decision criteria items	Teachers' mean ratings(SD) (N = 36)	Teachers' rank order (N = 36)	Ed. Assts. rank order (N = 10)[a]	Undergrads' rank order (N = 61)	Clinical Psychology rank oder (N = 21)
10. The currently available staff (and/or parent) time, materials, and physical facilities are adequate to conduct the necessary intervention.	14.53(3.39)	12	11	11	12
2. The behavior is damaging to materials, etc., in the child's environment.	14.36(4.16)	13	13	13	13
13. The behavior is age-appropriate and thus consistent with normalization concerns.	12.81(4.27)	14	15	17	17
3. The behavior is markedly deficient in comparison to the child's level in other areas (i.e., it is a weakness).	12.60(3.63)	15	17	14	16
16. The behavior is developmentally appropriate given the child's functioning level.	12.56(4.24)	16	14	15	15
4. Given an otherwise equal need, this behavior will probably be easier to modify than another.	11.06(4.74)	17	16	16	11

[a]The N for items 2 and 7 was 100 since one person each failed to rate these particular items.

TABLE 4.2

Varimax Factor Solution of Teacher Ratings
(N = 36)

Factor/Items (including item #s)		Loading (above .30) on specific factor
Factor 1:	Child Adjustment	
9.	Increase child's independence	.85
8.	Replace existing negative behaviors	.83
7.	Child would enjoy	.71
11.	Increase community acceptance	.61
13.	Age-appropriate	.49
5.	Parent concern	.42
6.	Interfere with learning	.32
Factor 2:	Instructional Utility	
10.	Resources to intervene are available	.78
16.	Developmentally-appropriate	.76
2.	Damaging to environment	-.32
Factor 3:	Behavior Change Efficiency	
4.	Easy to modify	.98
1.	Functional for child	.48
15.	Positive collateral effects	.41
Factor 4:	Positive Child Repertoire	
13.	Age-appropriate	.44
15.	Positive collateral effects	.44
17.	Prerequisite to learning	.79
14.	Dangerous to child	-.31
Factor 5:	Concern for Others	
2.	Damaging to environment	.49
12.	Dangerous to others	.87
Factor 6:	Urgent Child Needs	
14.	Dangerous to child	.61
1.	Functional for child	.47
7.	Child would enjoy	.34
3.	Deficit in child's repertoire	-.37

ous to others), where the lower rank (8th as compared to 2nd or 3rd) assigned to this item by the students presumably reflects their lack of experience in dealing with groups of children. The teachers differed from the clinical psychology students on the relative importance of four items: They considered item 6 (interfere with learning) and item 9 (increase independence) more important, and item 4 (easy to modify) and item 15 (will broadly affect the repertoire) less important than did the clinical trainees. These differences reflect predictable disciplinary emphases: That is, teachers would be expected to be more concerned with learning and see themselves as preparing children for independent functioning, while the psychology trainees see themselves as primarily consultants. Thus, while teachers seem to have rejected "easy to modify" as a legitimate reason for selecting a target behavior, the clinical students ranked this time higher in importance (11th, as compared to lowest rank for the teachers) which is consistent with their consulting experiences wherein the likelihood of achieving success is suggested as an important criterion (Tharp & Wetzel, 1969).

Aside from item 4 for the clinical trainees, there was considerable agreement across all groups on the five lowest ranked items 2 (damaging to environment), 13 (age-appropriate), 3 (deficit in child's repertoire), 16 (developmentally appropriate), and 4 (easy to modify). On the basis of face validity, currently recommended "educational best practices" would support the lesser importance of these items with the one notable exception of item 13. Judging activities, materials and even identifying programs and placements according to the criterion of age-appropriateness has become a major concern of educators working with severely handicapped persons (cf. Wilcox & Bellamy, 1982), yet our group's responses do not reflect the importance of this issue. However, the teachers did assign a considerably higher absolute rating--as opposed to ranking--on this item than either the clinical or undergraduate students, and the EA rating was slightly higher still ($\bar{X}$ = 13.60, SD = 4.06) which is consistent with their highest ranked response, item 11, relating to increasing community acceptance.

Generally, then, the teachers and the other groups were more alike than different in their opinions regarding the relative importance of the lowest ranked reasons. This suggests that these professional judgments have either been well-known and/or they actually reflect "ordinary knowledge" criteria for making important decisions. Whatever the reason for these results, they do provide support for the use of certain criteria hierarchically in making future decisions.

Teachers do seem to reflect higher priorities which differ from those identified by laypersons, paraprofessionals who work for them, and professional trainees in another discipline. The teachers appeared to respond according to educationally meaningful dimensions of decision criteria reflecting their professional training and/or classroom experience as special educators and not concensual, "ordinary knowledge" judgments. This is supported also by a factor analysis of the teachers ratings: The results are provided in Table 4. 2. Factor 1 (Child Adjustment) is concerned with the development of a behavioral repertoire to increase independence and facilitate community adjustment; this factor is oriented toward the child's individual needs. Factor 2 (Instructional Utility) appears to represent criteria that are instructionally "easy," and also includes a lack of concern for whether the child's behavior might be damaging to materials in the environment; this factor could be interpreted as an educationally oriented utility dimension that is not motivated by simply eliminating behavior disruptive to the classroom. Factor 3 (Behavior-

Change Efficiency) includes items considered important in technical discussions of behavior modification relating to acquisition, maintenance and generalization of behavior. Factor 4 (Positive Child Repertoire) is educationally oriented, including items relating to acquisition of skills and positive behaviors; this concern was accompanied by a lack of concern for whether a behavior might be dangerous to the child, although this negative loading is only marginal (-.31). Note that the item concerned with positive collateral effects--which was of special interest to our research--loaded on Factors 3 and 4 which suggests that teachers are most likely to be sensitive to this issue when they attribute importance to other criteria of sound behavior modification principles and skill acquisition concerns. Factor 5 (Concern for Others) suggests a concern for behaviors which might jeopardize other children and materials in the environment, and Factor 6 (Urgent Child Needs) includes items that indicate urgent child needs and a lack of concern for whether the behavior reflects a developmental deficit. Teachers consider important concerns for behavior dangerous to the child (item 14) and those skills that are immediately functional for (item 1) and likely to be enjoyed by (item 7) the child, but were not particularly concerned about relative skill deficits. Unlike the clusters derived from the naive undergraduate ratings (see Voeltz, Evans, Freedland, & Donellon, 1982, for more information), the teacher dimensions are clearly patterned according to general educational goals, such as the child's increased well-being and social integration, a concern for the principles of effective behavioral intervention, and an instructional utility orientation different from simply preventing the disruption of classroom routines.

Of course, providing ratings on general decision criteria which are not specific to actual behaviors, children, and program practices may not accurately reflect what teachers do in practice. Studies 2, 3 and 4 were designed to investigate how such criteria might be reflected in specific judgments made in actual situations.

Study 2: Effects of Developmental Delay vs. Excess Behavior on the IEP

Subjects

A separate sample of 40 subjects participated in Study 2. These participants were experienced special education teachers enrolled during the summer in graduate-level course work in the area of early childhood handicaps; most had already completed several core special education graduate courses as prerequisites to the early childhood training sequence. More detail on subject assignment to groups is available in Voeltz et al. (1982).

Procedures

Complete details on administration and scoring procedures are available in Voeltz et al. (1982) but will be summarized here. Briefly, subjects were provided comprehensive though typical child diagnostic and assessment information and asked to generate a hypothetical IEP which would reflect their priority goals for the child, behavior or instructional management concerns, and information regarding their personal attitudes toward the experience with a number of negative behaviors.

Design

Each subject was exposed to moderate or severe developmental delay accompanied by few or many excess behaviors in the child's repertoire for one of two target children, resulting in a 2 X 2 X 2 factorial design. The different levels of the variables were reflected in all written and videotaped information. The developmental delay and excess behavior variations were of theoretical interest, and by using one of two target children within each of the four main conditions, effects due to the unique contribution of the individual child could be investigated. The two handicapped children seen on videotape were both Caucasian and similar in appearance, including physical size; both were described as being 4 years of age. Severe vs. moderate developmental delay and many vs. few excess behaviors were controlled in the information received on the DSRS and by editing of the videotapes showing each child performing tasks of varying difficulty level and exhibiting various excess behaviors which might be generally described as "autistic." Overall, a large number of potential curriculum goals and target behaviors far exceeding the number typically listed on children's IEPs was presented to the teachers in each condition.

The major dependent variables of interest were the specific goals and target behaviors identified and their relative importance as a function of the three independent variables of developmental level, excess behavior, and, potentially, the specific child observed. The open-ended format for subject responses on the IEP was designed to provide qualitative insights into the way teachers made decisions and conceptualized their task while also imposing some restrictions on quantification. Qualitative findings were used to design the subsequent investigations, and the next section will discuss only the quantifiable outcomes.

Results and Discussion

As might be expected for an educational plan, the majority of the four annual goals mentioned by the teachers were exclusively curriculum oriented (132 of 160 goals, or 82.5%), i.e., they did not reflect obvious behavior management intentions. Most of these were categorized as language (e.g., "increase expressive language") or adaptive behavior (e.g., "develop self-help skills"); of the 132 curriculum goals, 31.1% were language and 31.1% adaptive behavior concerns. Motor goals (e.g., "develop gross/fine motor skills") and cognitive skills (e.g., "increase readiness for academic programs") were mentioned less frequently, accounting for 18.9% and 17.4% of the total respectively. Finally, only two goals (1.5%) reflected socioemotional concerns (e.g., "increase independent play" and "develop adequate social skills"). The bias in the direction of language could be a result of the fact that more than half the teacher sample was enrolled in a language development course. But the children did exhibit autistic-like behavior, so that the teachers' concern for language and adaptive behavior--rather than motor or cognitive development--could reflect their awareness of the needs of autistic children. On the other hand, the lack of concern for socio-emotional needs is inconsistent, unless these teachers viewed behavior management objectives (see below) as socio-emotional ones (e.g., "lack of eye contact" or "increase attention to task" were categorized here as excess behavior concerns, not curriculum goals). They listed more than twice as many cognitive goals for the severely delayed condition in comparison to the moderately delayed condition, and twice as many adaptive behavior goals for the few excess condition in comparison with the many excess condition. No other major differences were apparent.

Excess orientation. Many of the annual goals were focused upon modifying excess behaviors rather than skill instruction. An "excess orientation" score was calculated for each teacher's simulated IEP by assigning increasingly higher numerical values to goals primarily oriented toward behavior management or the reduction of excess behaviors. This excess orientation score was then employed as the dependent measure in a three-way analysis of variance, with delay (2 levels), excess (2 levels), and child (2 levels) as independent variables.

The main effect for excess was highly significant ($F[1,38] = 30,3$, $p < .001$), with teachers exposed to the many excess condition having ahigher excess orientation score. None of the other main effects or interactions was significant. Thus the degree to which excess behaviors were present in a child's repertoire affected the nature of the priority annual goals listed by that teacher as part of a program of educational remediation, irrespective of the severity of the child's developmental delay.

Excess behavioral targets. Regardless of whether excess behaviors were included in their annual goal choices, the respondents were then asked to list all target behaviors (as opposed to curriculum/skill needs) that they might target for intervention, in order of seriousness. If a respondent had already included any excess behavioral target as a priority annual goal, this was considered to be highest priority and was also included in this analysis. The five major groupings mentioned in the target responses were: (a) attention (e.g., establish eye-contact, increase attending), (b) tantrum behavior (e.g., reduce temper tantrums), (c) self-stimulation (e.g., reduce hand flapping, decrease rocking), (d) aggression (e.g., hitting others), (e) compliance (e.g., follow directions, obey commands), and (f) self-injury. Five separate analyses of variance were then performed on each of the first five behaviors in order to determine the effects of the independent variables on respondents' selections. For attention, there was a significant main effect due to child ($F[1,32]= 12.8$, $p < .001$) and to excess ($F[1,32] = 7.9$, $p < .01$), with the many excess conditions resulting in much higher priority given to attention. The only other effect which was significant was degree of delay upon working independently ($F[1,32]= 5.7$, $p < .05$); this target was mentioned by 30% of the subjects in the moderately delayed conditions and by only 5% of the teachers viewing the severely delayed conditions.

Although nearly half the teacher subjects mentioned attention as the first target behavior choice, one-fourth of them did not mention this behavioral objective at all. This level of agreement--characteristic of all target behaviors ranked--together with the general lack of effect due to the manipulated variables of child, degree of delay, and number of excess behaviors, suggest that the teachers were selecting behavioral targets on a more personal, idiosyncratic basis. Two major sources of individual influence might be the teachers' attitudes about and their own personal experiences with children exhibiting various behavioral excesses. Both these aspects were measured in the two follow-up questionnaires, and although the sample size was too small to relate individual attitudes to individual choices of IEP targets, the general group findings from both questionnaires is reported in detail in Voeltz et al. (1982).

The most general conclusion to be drawn from these results is that there are wide variations in teachers' familiarity with the excess behaviors, both directly and vicariously, which presumably affects both educational judgments about behavioral goals as well as teachers' receptiveness to recommendations by outside consultants.

Although we attempted to structure a realistic IEP judgment task, there may be large differences in involvement between constructing a hypothetical IEP and planning an actual educational program for a real child for whom the teacher will in fact be responsible. Thus, Studies 3 and 4 were designed to further investigate teacher decisions made on behalf of children actually enrolled in the respondents' classrooms.

Study 3: A Cost-Benefit Analysis of Child Gain vs. Regression

Introduction

The concept of considering costs and benefits of behavioral interventions in educational settings emerged early in our thinking about response interrelationships. If, as we demonstrated in our review of the literature (Voeltz & Evans, 1982), some reported interventions have produced negative side effects, and if behaviors are organized into clusters in which both positive and negative behaviors may covary, then the circumstances exist in which certain interventions(and certain chioces of targets) could have costs as well as benefits for severely handicapped learners. In a more recent, wide ranging paper, Kazdin (1982) came to a somewhat similar conclusion: That what had been called "symptom substitution" in psychoanalytic psychotherapy is a phenomenon that has also been observed in behavior therapy and is most probably due to the prior response relationships.

In psychotherapy outcome research a recurrent issue--first articulated by Bergin (1966)--is whether there is a "deterioration effect"; whether psychotherapy is influential enough to be detrimental to some clients. No exactly comparable issue has ever been raised in behavioral research with severely handicapped children, although severe criticism of "readiness" and other developmentally-oriented approaches as essentially wasting many pupils' limited educational time on trivial educational goals (Brown, 1982) shows that the question of harm as opposed to just varying levels of benefit will become a more major issue in evaluation studies. That treatment methods per se may carry risk of harm has long played a central role in medical decision making and has certainly been considered in behavior therapy as an ethical issue when using aversive stimuli and other invasive procedures (e.g., Stoltz, 1978).

Proponents of aversive procedures with children will continue to argue the ethical guidelines which permit their use, despite the fact that the side-effects of physically aversive interventions are really not well known. We, however, support the TASH resolution on intrusive interventions (1981) and have built into the decision model (Evans, Derer, Voeltz, & Hanashiro, 1982) very explicit criteria for the use of <u>any</u> punishment contingency--which can be justified only in the case of Level I excess behaviors which are those that are life-threatening or likely to cause irreversible phsycial harm to the child. Another interpretation of cost-benefit which we have explicitly argued against is the economic concept of cost--the attempt by some recent commentators to place a dollar value on educational and treatment services for severely handicapped children. We have expressed our serious concern for this attitude among service providers and evaluators (Voeltz & Evans, in press; Evans & Voeltz, 1982) and will not repeat it here. However, it should be re-emphasized that our study was based on concerns regarding costs to the handicapped child regarding the possible outcomes of an educational decision and was not concerned with the financial cost of services nor with the side effects of intrusive interventions.

After excluding economic costs, harmful consequences of intervention procedures, and costs to the child in terms of efficiency and relative value of the educational program, the costs of an intervention pare down to negative collateral effects, namely the increase in some other excess behavior or the decrease in some skill which is already present. As has been mentioned throughout this report, our original hope had been to be be able to isolate common response clusters empirically and thus make specific data-based recommendations regarding expected collaborative effects, both negative and positive. With this goal not being realized (and probably unrealistic in terms of general prescriptions) we approached the issue from the point of view of the teacher's best estimate of what behaviors might increase or decrease as the result of intervening with specific target behaviors. In the decision manual, suggestions are made as to how to make these intelligent guesses; how to weigh the costs and benefits to reach a deicison was the focus of this study.

Method

Subjects

The subjects for this study were originally all of the approximately 200 teachers who appeared on the Hawaii State Department of Education's listing of teachers serving moderately to severely handicapped and seriously emotionally disturbed children; this listing included teachers of moderately to profoundly retarded, severely multiply handicapped, deaf-blind, and autistic children. A questionnaire and a cover letter was sent to this group, and 53 were returned by the dealine. Of these, 17 were judged unusable in some way or were incomplete, leaving a final subject pool of 36 respondents. The relatively low rate of usable returns can largely be attributed to the complexity of the questionnaire which therefore required considerable time to complete near the end of the school year (several teachers indicated verbally that other responsibilities at that time precluded a response).

Materials

Each teacher was mailed a packet containing an introductory letter, the BSIP list of excess behavior definitions (see Appendix A), a questionnaire, and a stamped addressed envelop for return of the questionnaire. In order to eliminate the bias that might result from pre-selecting behaviors for the teachers to consider, we asked them to select one of the children in their own classroom and essentially to construct their own questionnaire. Thus teachers were asked to select the student "who displays the greatest number or more serious excess behaviors." They then had to list from this student's IEP the four major goals and the first priority objectives for each goal; to help us identify the type of goal, they also classified each objective according to 10 domains--language, motor, leisure, etc.--that were defined for them. Finally they rank ordered each objective and entered them in the blank spaces provided in two rating scales so that increases or decreases in the four skills could be rated according to how much better off they might consider the child to be as a result of such changes. (The actual identities of the children was not requested and hence confidentiality of information on individual children was not involved.)

The second task for the subjects was to check all excess behaviors on the list of definitions which were exhibited by their selected students. They then had to select the four most serious of these, rank them in order

of their seriousness, and enter them in the blank spaces provided so that increases or decreases in these excess behaviors could be rated according to the same dimensions as the four skills. In this way, each teacher was rating skills and excess behaviors that were meaningful priorities for an actual child in their own classroom.

The rating scale used was an 8-point Likert-type scale anchored by the statements "No better off" or 'Substantially better off" (or "No worse off" and "Substantially worse off" when rating a decrease in a skill or an increase in an excess). What do the terms "better off" or "worse off" mean when judging the lives of handicapped children? Obviously these dimensions will mean different things to different teachers (as they have been shown to do with parents, administrators, policy makers, and so on) and will depend on the context assumed for the question--better off financially? worse off in terms of physical comfort? better off in terms of social adjustment? We presumed that teachers would be able to combine such criteria along a dimension that accords with current professional judgment regarding desirable goals: maximum independence and participation in current and future least restrictive environments. To show that this was the dimension of ultimate concern we used examples from everyday work situations, not because we consider work to be the highest human goal but because work situations impose more stringent requirements or appropriate adaptive behavior than social, leisure, and other equally important contexts. The examples, while slightly facetious, were designed to show that "better" or "worse" off should be construed in the context of adequate <u>functioning</u> in least restrictive environments. They were as follows:

> "What do we mean by better or worse off? We mean according to the child's needs at the present <u>and</u> in the future. For example, if you worked on a farm you would be better off if you learned how to drive a tractor, and probably only a little bit worse off if you acquired a tendency to make finger postures. If you were an insurance salesman you would be very much worse off if your verbal skills deteriorated, and perhaps somewhat better off if you learned to reduce some mannerism such as rubbing face or mouth. As a graduate student you might be better off in the future if you acquire the tendency to read current journals, and you might be immediately worse off if your tendency to stare off into space became significantly greater.
>
> When you do your ratings try to use a wide range of the scale: If you enjoy playing the piano and you start to develop a finger mannerism you are going to be <u>substantially</u> worse off than if you enjoyed jogging and developed a hand flapping responses, although in both cases you would be slightly "worse off." If you learn to read words you are going to be much better off than if you just learn to recognize letters, although both may be very significant accomplishments for a handicapped student.
>
> With these considerations in mind, please now rate your four skill behaviors and four excess behaviors in the following scales:"

<u>Design and Procedures</u>

The design of the study was contained in the questionnaire. Subjects rated their four priority goals twice (once in terms of decreases, once in

terms of increases in behaviors) and they rated their four most serious excess behaviors twice also. This produced a balanced two by four (type of behavior by direction of change by order of priority) within-subject design. The dependent variable was the ratings provided by the teacher subjects, and their responses were examined in more detail as described below.

Results and Discussion

Ratings

The ratings given by each respondent were subjected to an analysis of variance. Of the main effects, only order of priority was significant ($F[3,464] = 3.24$, $p < .05$) which really indicates little more than a confirmation that the subjects followed instructions to the point that changes, whether improvements or deteriorations, in priority behaviors were not seen as significant as changes in higher priority behaviors. No other main effects or interactions were significant, which indicates that the subjects did not discriminate significantly between changes in skills or changes in excess behaviors, not did they perceive improvement in behavior as relatively more positive for the child than deterioration in behavior was negative. Although not significant, the teachers did give the highest ratings (6.9 on the 8-point scale) to improvement in skills--indicating their perception of importance to the child--followed by worsening of excess behaviors (6.7).

Qualitative Analysis of Ratings

One reason why differences did not emerge is that the ratings--as can be seen from the above figures--were uniformly high, despite efforts to encourage subjects to use the full range of the eight-point scale. One response style seen in a number of the teachers was to give maximum ratings (8) to all behaviors in all conditions, or perhaps a rating of 7 to the lowest priority excess or skill. We feel that this indicates many special education teachers' deep commitment to the need for their students to show active progress and that any response not acquired or excess getting worse is an extremely serious matter. Given the relatively small gains that severely handicapped students often make and our failure to measure success in terms of quality of life variables, the intense emphasis placed on behavior change by the teacher might help create unrealistic expectations of their pupils and unrealisitically negative approaisals of their own efficacy, with the accompanying risk of "burn out."

Some teachers gave ratings that we would consider sophisticated, namely their ratings agreed with their rankings, first of all, and then they tended to rate skill acquisition as more important than improvement in excess behavior and skill loss as more serious than a worsening of excess behaviors. This conforms to the priorities we have expressed in the decision manual and is probably in close accord with current ethical assumptions of leading professionals in the field. Only 5 teachers clearly indicated this pattern of response; however, only two gave the opposite, namely rating changes in excess behaviors as more important than changes in skills. In our earlier teacher decision making study we identified what we called an "excess orientation" and these two teachers seemed to reveal it here. The present materials could be used in a training situation as one way of sensitizing teacher trainees to the effects that their attitudes might have on their priorities for seeing change in their pupils.

Skill and Excess Behavior Priorities

Because of the low return rate, this sample cannot be considered truly representative of teachers of severely handicapped pupils in the State of Hawaii. However, as the teachers were asked to select their most difficult student in terms of excess behavior and as they were asked to report on the actual IEP goals they had established, their responses are of some interest in terms of revealing trends in actual IEPs.

The ages of the children selected ranged from 3 through 20 (Mean= 10.5). When the overall frequency of each domain was tallied, the following results were obtained:

	Total # of	Rank Order
Language (e.g., signing, speech sounds)	18	2
Cognitive (e.g., discrimination, matching)	8	6
Adaptive behavior/self-help (e.g., dressing)	27	1
Motor (e.g., walking, range of motion)	15	5
Social/emotional (e.g., play, turn-taking)	17	3*
Leisure/recreation (e.g., toy play, use of playground equipment)	4	9
Community integration (e.g., going to a store, a restaurant)	3	10
Domestic living (e.g., cooking, cleaning)	5	8
Prevocational/vocational (e.g., object manipulation, assembly)	6	7
Preacademic/academic (e.g., coloring, sight words)	17	3*

The most popular domains were adaptive behavior, language, academic, social, and motor. Interestingly, if broken down by order of priority, adaptive behavior, language and social goals dominate the first and second priority positions, but as third and fourth priorities are considered, academic and then motor goals become more frequently mentioned.

Excess behaviors were varied: Among the first ranked excess behaviors only 4 behaviors were mentioned by more than one person, so that 26 behaviors were listed by the subjects; similar diversity was found for the 2nd, 3rd, and 4th ranked behaviors. As the teachers were asked to check off on our list of excess behavior definitions all behaviors exhibited by their target student, it was possible to see which behaviors were noted most frequently to be occuring among these children.

*tied rank.

Of the 94 behaviors on our list only 4 were mentioned by none of the subjects teachers--04 spinning self; 46 choking other; 50 shadow play; 93 finger/hand posture--which seems to provide support for the realisitically comprehensiveness of the list. The following behaviors (in rank order of frequency of mentions) were noted in 25% or more of the children:

01: Body rocking (15 mentions)
26: Mouthing objects (14)
25: Mouthing body parts (11)
28: Scratching/picking skin (11)
40: Grabbing at others (10)
41: Grabbing at objects (10)
60: Cry, no tears (10)
11: Hand/arm flapping (9)
52: Object banging (9)
29: Genital touch/masturbation (8)
61: Cry, tears (8)
66: Vocalization, supraglottal (8)*

Conclusions

We tried to create a task that was realistic and based on experienced teachers as well as the specific students that they instructed on a daily basis. Under these circumstances, studying teacher judgments seems a defensible component of the overall research plan. Unfortunately, the magnitude of the task of constructing, essentially, their own questionnaires, seemed to result in a low rate of usable returned questionnaires.

Of those that were returned, the dominant response pattern was to give rather extreme ratings of the importance of the child's behaviors improving and not deteriorating. This, we believe, reflect an intense commitment to the importance of severely handicapped learners' making steady progress without loss of skills or emergence of new or more intense excess behaviors. It may also reflect the uniformly high priority given to the four priority goals specified on children's IEPs as well as the high concern expresseo for the "most serious" excess behaviors--this is, teachers (with each child's parents) had already given each child's total repertoire serious consideration and selected for instructional concern those skills and behaviors which were the most crucial, and equally so, for the child's outcome.

A few teachers made more carefully refined judgments in which skill acquisition and loss was related more significantly than excess behaviors' improving or worsening; we have argued elsewhere that this is a desirable orientation in educational contexts. However, when considering the subjects overall, there were no significant main effects of skill or excess or improvement or deterioration in behavior or their interaction. This answers, at least for this group of "expert" judges, our original question regarding the weighing of costs and benefits of interventions: The loss of a skill is not perceived as relatively more serious for a student than the increase in an excess behavior. Thus, when weighing the pros and cons of anticipated negative collateral effect, a simple additive rule can be used and elabor-

*This category is not technically a negative "excess behavior, but can include pre-language vocalizations.

ate relative weights for loss of positives as opposed to gains in negatives are unnecessary. The final judgment, of course, requires that the collateral behavior also be given some kind of importance rating. The apparent tendency for teachers to rate a number of behaviors as equally important for their students could reduce the fidelity of their choices of targets and interventions, assuming that not all excess behaviors or skills can be programmed at once for children who have multiple needs. The materials used to generate these data, however, could be used in training programs to heighten student teachers' awareness of cost-benefit analyses that show some ability to make fine-grained judgments about behaviors when children's needs are many.

Study 4: The Teacher Interview Study

Introduction

In the previous studies reported in this section, the emphasis was on the quantitative analysis of various components that are of importance in teacher decision making with respect to behavioral problems in severely handicapped children. The present study attempted to determine more qualitative information regarding the way teachers typically proceed through a complex behavior management decision. In the medical decision making literature, Elstein, Shulman and Sprafka (1978) labelled this a process-tracing approaching, in which the reasoning processes used by subjects as they solve problems are articulated and described. More specifically, the type of process-tracing study used was that of stimulated recall (Shavelson & Stern, 1981): The teachers were interviewed using a standard protocol, and asked to explain their reasoning behind what they were actually doing with children who exhibited high levels of excess behaviors in their classes.

The purpose of this investigation was two-fold. One was to gain possible insights into actual teacher criteria that may not have been considered in the previous investigations, and the other was to determine whether teachers used a conceptual frame work that would be antithetical to the principles described in the flow diagram of the decision manual. Before suggesting a model decision making procedure, it seemed valuable to have some indication that the procedures would appear valid according to the pre-conceptions and strategies actually used by special education teachers.

Method

Subjects

The subjects were nine special education teachers who were conducting classes in the State of Hawaii Department of Education that contained either one of our experimental children and/or one of our comparison children; all teachers were serving children enrolled in self-contained special education classes on regular campuses serving a larger number of nonhandicapped children. Selection of the teachers was somewhat arbitrary as a major criterion was that they would have to be willing to be interviewed in depth regarding their strategies and be willing to spend the hour to hour and a half that the interviews required. However, from the pool of teachers available, we selected as widely divergent a group as possible. Two were recent graduates of the University of Hawaii Department of Special Education, three were former graduates of that program (prior to implementation of the present training program specifically designed to provide "best practices" training in the area of severely handicapped), and the remaining four were from a variety of academic backgrounds but who had tenure as special education teachers in the public school system. They were not selected because they were the most

skilled of our teachers, but rather because they represented a fairly typical cross-section of "expert" (in terms of experienced) teachers.

Procedure

Each teacher was interviewed according to a standard protocol in an unstructured interview. All interview were conducted by Evans based upon two considerations: (1) considerable expertise and experience in the interview process was essential, thus making it essential that they be conducted by a fully qualified professional; and (2) as a clinical psychologist, Evans both met these qualifications and also was not directly involved in having trained any of the teachers to be interviewed specifically, since most of the teachers had been trained at the preservice and/or inservice training levels by Voeltz, it was felt that the latter could not conduct the interview without biasing teacher responses in some unknown way.

A child in each teacher's class was the focus of each interview; each of these children, as either an experimental or comparison subject, was well-known to the interviewer, usually over a period of several years. The teachers were asked to specify the most serious of the child's excess behaviors in order of seriousness. They were asked to explain why the behavior they identified was changeworthy or what negative implications it had for the child. The subjects were then asked which of those behaviors they were currently attempting to modify, if any, what the intervention plan was, and how formally that plan was expressed (was it written, a part of the IEP, etc.). Information that they supplied was followed up on if it seemed to indicate their attitude toward excess behaviors and difficulties in modifying them. Perception of any behavior cluster was probed, as was awareness of the possibility of negative side-effects or costs of interventions. The teachers were also asked, towards the end of the interview, what frustrations or difficulties they had experienced in producing ideal educational services to the children, and, in order not to sustain a negative tone to the interviews, they were asked to indicate the rewards or satisfactions they obtained from teaching the child who was the focus of the interview.

Interviews were tape-recorded and later transcribed. Rather than performing a quantitative analysis, the transcripts were carefully scrutinized and features which seemed of special interest to the theme of this investigation were summarized and are reported in the next section.

Findings: (1) Management of Excess Behavior in the Classroom

One of the most striking findings from the interviews was the degree to which the teachers are obliged to engage in day-to-day and moment-to-moment decision making regarding how to respond to excess behavior. Excess behavior seem to show substantial fluctuations with intervening holidays, changes in teaching personnel and classroom aides, nature of the task requirements, and other environmental factors. A second very general observation by the teachers was that each child considered had exhibited one or more excess behaviors when first coming into the class which had since been successfully modified

In no case was the modification of an excess behavior listed as an IEP goal although one IEP (not written by the current teacher being interviewed, but by the child's previous teacher) did have a provision that excess behavior would be monitored:

T: "One of his objectives in the IEP is to watch him for 5 minutes and to note the self-stimming--the mouthing of toys, and the blowing on the toys (which we didn't see much of today) and generally throwing the toys."

However, in a couple of cases, the skill acquisition goals in the IEP were designed primarily to reduce an excess behavior, e.g.:

E: "Is the particular problem of crying when exposed to strangers or new situations written in the IEP at all?"

T: "I think it's included as one of her social skill programs as far as being exposed to others and being included in group activities--there is a statement in her IEP which includes the fact that she will participate in the Special Friends Program.*"

E: "So that means, in fact, that the program is in a sort of positive direction--it's not that you're going to reduce her crying, but to increase her ability to respond socially?"

T: "I think her IEP in general is written rather positively."

Another general feature of the replies was that a very wide range of formality of intervention plan was in operation in the various classrooms. The general tendency was for the teachers to have an explicit strategy for dealing with almost all excess behaviors they saw, though some responses revealed certain deliberate choices to ignore selected behaviors:

T: "Sometimes they're just passing through a stage, especially sucking or whatever. I find that I pretty much ignore behaviors unless they're interfering with the whole classroom kind of scene. I guess crying is one, tantrums--those types I would deal with much quicker than these other types of behaviors."

* * *

E: "...some of these behaviors you've chosen to specifically intervene with and some you're aware of but there is no formal programming. How would you say that came about?"

T: "It's probably due to the fact that if we intervene constantly with every behavior that is on his list, his whole day is going to be a real negative environment. I mean, I can deal with my hair being pulled, if it's going to be, rather than keep a whole string of negative interventions going all day long."

* * *

*The Special Friends Program is a school-based program structuring social interactions between severely handicapped children and similar-aged nonhandicapped peers; the interactions are generally dyadic or consist of small group situations during recess and other play periods.

T: "If he's walking around doing his hand thing over in freetime area, he can do it all he wants."

E: "He can do it? Even though that might also be viewed inappropriate by society when he gets older?"

T: "Well, I can't intervene on everything."

Most of the teachers mentioned that they kept data on the behaviors that were the most serious or for which rather exact counts would be specially useful--for instance, one teacher was keeping data on frequency of seizues prior to a reference to a pediatric neurologist. However, there was a tendency to see charting as a superfluous activity:

E: "Have you tried to keep any formal data in terms of counting these things, charting them, or ..."

T: "Oh, I have,, I've counted. But you know, I just--more important to me is being there, and when he does it, tell him, "No, I don't like it when you do tha , we'll do this instead." That's more important, because when you have a class of five kids it's hard to ... well, everybody's got a chart for everything! There are limits to human possibilities.!"

* * *

E: "Was the program formalized in any way, like your aide knew about it and so forth?"

T: "Yes, everyone who worked with him knew about it, but there was nothing on paper."

E: "Did you ever keep any data on them in terms of ..."

T: "No."

E: "How would you judge whether the program was effective or whether it needed changing?"

T: "It just worked. Well, then it decreased...I guess we did make a few changes as it decreased.

Many interventions were just being worked into instructional classroom routines in the manner we have recommended in the decision manual. As one teacher remarked "Well, I try to control and try to remediate all those behaviors, one way or ne other." The most common interventions is this context were verbal feedback with a physical prompt if the child failed to comply, distraction by oresenting incompatible tasks ("We're trying to have her more holding things and doing things with her hands rather than tapping.") and ignoring. The most consistent stratedgy for intervening with excess behaviors was to find an alternative positive behavior, but only one or two of the teachers articulated this as a strategy related to a careful functional analysis of the behavior:

E: "You mentioned with the self-biting you take his hand out of his mouth and tell him you don't like it when he does that--this was an informal intervention, something you just did consistently?"

T: "Yes, but at the same time realizing that he must be doing this for a reason and trying to provide some other alternative, trying to substitute another, more appropriate behavior, something less harmful to him."

(2) Reasons for Selection of Behaviors

Across the nine teachers there was evidence of excellent sensitivity to the reasons for being concerned about target behaviors which have been proposed in the professional literature and which we investigated more formally in the structured studies reported already. Teachers were responsive to demands of future environments ("If it's something that I see now that when I think of him being 18 is going to cause serious problems, then I'm concerned about that also.") and to social acceptance by other children, other teachers, and the community ("The drooling affects the way people relate to her; the regular ed. kids won't come and touch her at all."). However, as one teacher pointed out (with respect to crying), a complaint by a regular education teacher is the next classroom had not bothered her "because there is so much noise in the school anyway--I mean look what goes on in the gym..." Parental concerns were mentioned only once and the teachers seemed comfortable with their advisory role as experts: "I feel like we need to work together with the parents. If that is what they feel is important to them, then I'll see what can be done. But I'll also express my opinion that it can be indirectly dealt with--if we do these other things maybe it'll disappear." Another teac commented, "I'd try to explain to the parents that this is just a stage the child is going through so that they could accept it more."

Probably because few of the behaviors exhibited by the children in the sample were dangerous to the child, this criterion was not the most frequently mentioned. In a variety of ways, however, the teachers revealed their sensitivity to this as a major consideration in their decision making, as can be seen from the following excerpts:

T: "... and for her own safety, that is my first concern, her own safety. She gets just blind with rage and then I've seen her roll and bang into furniture."

* * *

E: "Why is that an inappropriate behavior from your point of view?"

T: "Self-stimming? Because she won't attend to tasks when she has her hand in her mouth."

E: "So it interferes with..."

T: "Programming."

E: "Any other reasons why you might focus on that behavior?"

T: "Yes, health reasons--her fingers are all cracking and have blisters, they are getting raw."

* * *

T: "...something that is real destructive, you know, like when kids try to scratch their eyes out, or, you know, you don't want to have

to give a child a tetanus shot every six months because they're causing themselves to bleed. Things that cause bodily harm, that are dangerous to their own well-being."

E: "So that might be the most important consideration?"

T: "Yes, that's mine."

Undoubtedly the most frequently mentioned reason for a major intervention effort being organized in the classroom was that the behavior was highly disruptive and interfered with the teacher's ability to conduct successful instruction. One example of this has already appeared; other typical comments were:

T: "He'd throw the materials that we were using in his training program. And I felt that that's not going to work. How can he learn if he keeps throwing everything?"

* * *

T: "Well, when she gets up to the point of screaming, that's very disruptive in class. Screaming and crying. Crying is okay, but when it starts to get up to a screaming it's impossible to teach her in that state..."

(3) Major Difficulties in Successfully Implementing Interventions

Although not specifically questioned on these issues, the teachers did not reveal any particular conceptual model of excess behavior. There was a tendency by some to see excess behavior as possibly a consequence of organic causes (which in the case of one of the target children being discussed was a highly probable explanation and the teacher's identification of an allergic reaction was very perceptive and useful for designing an intervention). There was another tendency to see excess behaviors as stages the child was going through--mouthing objects, for instance, representing a stage of exploration. In both cases, the orientation provided useful insights by the teacher into the functions of the behavior; this was particularly true of disruptive behaviors that seemed to be attention-seeking or represent attempts by the child to regulate his/her environment.

Clusters of behavior were recognized by some of the teachers, with the most commonly noted pattern consisting of a sequence of escalating components of a tantrum. Another common assumption of the subjects was that various excess behaviors co-varied on days in which there was some more general reason for the child to be in a bad mood, such as a change in the child's foster home. For these reasons, perhaps, the teachers all identified lack of follow-up of their intervention programs in the home environment as the major barrier to successful modification. A second difficulty they expressed was the rapid turnover in aides and in support personnel (e.g., OT, PT) who, as a result, were less successful than the teacher in keeping disruptive excess behaviors under control. Another problem was the tendency of the children to lose the gains they had made after a vacation period.

Discussion

We were rather impressed by the level of sophistication that the teach-

ers revealed with respect to current professional practices in special education. Without being cued by a formal questionnaire, the teachers articulated a number of criteria for choosing to intervene with excess behavior. While interference with teaching was most frequently mentioned, it was also clear that dangerousness was a primary consideration. The logical sequence of posing these questions in the flow diagram of the decision model should therefore be rather helpful to teachers who recognize a number of factors which must all be taken into consideration. The teachers were also very clearly oriented toward preparing children for integration into the community and in reducing particular excess behaviors which might interfere with peer acceptance. One subject pointed out that unless teachers of elementary-age handicapped children had knowledge of subsequent environments (e.g., what would be appropriate in middle school and high school settings?) they could not make informed judgments regarding the appropriateness of behavior.

Another encouraging feature of the interviews was that the teacher subjects recognized the crucial importance of replacing excess behaviors with functional and appropriate alternatives. They were generally successful in designing interventions that took place within the context of instruction. It would seem, therefore, that the material contained in the flow diagram (see Voeltz, Evans, Derer, & Hanashiro, 1982) will be readily accepted by teachers, although of course, the representativeness of Hawaii special education teachers is an unknown. The interviews provide a rich source of information regarding the process of decision making by teachers and only major implications have been summarized here. Overall, they were an optimistic group who showed their reward and professional satisfaction deriving from pupil progress and responsiveness and whose frustrations were rarely with excess behaviors in the students. Instead, they expressed difficulties primarily with the lack of support from related professionals (OT, PT, psychologists, etc.) participating in the daily classroom procedures, lack of parental follow-up or continuity with out-of-school environments, and the lack of maintenance of behavioral improvements following vacation and summer breaks as well as the lack of generalization across environments and other persons (e.g., classroom aides).

General Discussion and Conclusions

Our results indicate that considerable variation exists in professional perceptions of handicapped children and correspondingly suitable educational goals. That this was true for our relatively homogeneous group of special education teachers--many of whom have had similar training and/or classroom experiences in Hawaii's state-administered educational system--clearly reveals that decision making is influenced by more than simply the variance in children's repertoires. These findings would, of course, require replication with professionals in other regions and educational systems, but it may be even more important to begin investigations into the processes of certain kinds of decision making and the qualitative improvement of those decision based upon new developments--i.e., changing criteria defining what "educational best practices" really are--as well as the specific results of individual intervention efforts. As Page (1980) emphasized, this will require that substantial effort be focused upon longitudinal monitoring of child outcomes as a function of particular decisions. Without such empirical data, children's programs are clearly being shaped by educational practices which may not reflect the accumulation of knowledge from intervention studies as much as they reflect the accumulated professional biases of acknowledged experts.

Perhaps more serious is the implication that the important opportunities for learning are lost as teachers adhere to personal and idiosyncratic interpretations of recommended best practices. Thus, IEPs would not be based upon empirically-supported child needs but upon overgeneralized applications of simple decision rules learned in teacher training programs or the social climate of school settings. As we have emphasized throughout, empirical evidence on actual child outcomes as a function of the various intervention efforts could answer such questions with some degree of confidence that such decisions will represent a child's ultimate best interests. We have already summarized the complexities involved in efforts to demonstrate conclusive empirical relationships among even children's multiple behaviors as well as between specific program strategies and the effects on those behaviors (see Chapter 3). The next chapter presents an overview of several of our efforts to document such relationships, so that teacher decisions might be based upon empirical data. However, this work will require considerable further effort by researchers as well as finding solutions to a number of methodological problems. In the interim, since teachers must and do continuously make decisions affecting children's lives, information on how this occurs could at least insure that such decisions are oriented not toward idiosyncratic professional biases (as was somewhat apparent throughout these studies), but toward planning and modifying children's programs based upon their effects on the child on both a day-to-day and long-term basis. Thus, both the decision/planning model and the educational evaluation process developed by our project (see Chapter 7) require that the teacher follow systematic procedures across children but appropriately individualized for specific effects on a specific child.

CHAPTER 5

Normative Data from the Comparison Children

In the previous chapter we described critical aspects of teacher decision making when presented with various problems of target behavio selection and intervention priorities. These studies implicitly raise the question of just how widespread these behavioral problems might be in school classrooms for the severely handicapped, and this issue is the focus of the present chapter. The topic is of increasing importance. When we wrote our original proposal we recognized a new and fascinating problem for investigators concerned with behavioral excesses in young handicapped children. Most previous studies have been conducted in institutional settings or in laboratory schools, which however excellent technically, nevertheless violate current rights of handicapped children to receive services in least restrictive settings and to associate with nonhandicapped peers. Researchers in the current era must conduct their inquiry in natural environments that afford lowered opportunities for control over conditions or selection of methods based on scientific rather than educational necessities. This is by no means an unwelcome challenge; it requires us to expand methods and measures that are appropriate to the realities of actual classrooms, families and communities. A major feature of these new requirements is to provide some "baseline" or comparison, normative data on the nature and scope of behavioral problems among severely handicapped children who have been identified early in life and have received adequate educational services under legal mandates.

The children who constituted our comparison group were selected according to the same criteria as experimental children; these have been described earlier. Usually comparison children were in schools that were geographically inconvenient for the observers to reach (especially schools located on one of the neighbor islands in the Hawaiian chain), were identified too late in the school year to be included in the experimental studies, or were experimental students who had moved to less accessible schools. That is, the schools were not necessarily located further away from the university than those attended by experimental subjects, but decisions had already been made regarding "clusters" of schools to which our observers would travel on given days.

The original schedule proposed for comparison children was two observations per school year. In practice only five data points were obtained over the three project years, and although a total of 66 children were at some time in the comparison group only a smaller number were able to be observed all five times. Video-tapes were made of the children in the three situations used for the experimental students; these tapes were replayed in the laboratory and coded by the observers off the video monitor. Feedback was provided by all teachers and any parents requesting it. Where senior project staff had carried out the videotaping, particularly on the neighbor islands (since budget limitations precluded more than one person traveling for both data collection and feedback purposes), discussions with teacher and parent usually followed the taping session. Various suggestions regarding teaching method, curriculum content and behavior management strategies would be provided at this time in addition to ex-

plain; , the previously gathered data. Apart from this rather major consultation effort, however, no specific interventions were implemented for comparison children.

General Background

Before presenting the results of this aspect of the project, some commentary on previous studies of the incidence of excess behavior in handicapped individuals provides a certain perspective. In a fairly recent British survey (where more severely retarded children are in residential institutions), 40% of the children under the age of 16 exhibited "stereotyped or repetitive, apparently purposeless motor activity"; 13% displayed benign self-injurious behavior; and less than one-tenth of one percent showed severe self-injurious behavior involving intense head banging, eye-gouging, and biting of the extremities (Corbett & Cambell, 1981). Simply using the criterion of behavior which results in physical harm or tissue damage, Baumeister and Rollings (1976) reported a prevalence of self-injurious behavior of between 10 and 17 percent in mentally retarded persons in institutions (all ages). Using a questionnaire survey of the entire population of an institution (1300 retarded persons), Maist, Baumeister, and Maist (1977) carried out a factor analysis which indicated self-injurious behavior was associated with more profound retardation, specific indications of neurological impairment, and higher rates of stereotyped and aggressive behavior. Schroeder, Schroeder, Smith, and Dalldorf (1978) provided some interesting data on three yearly repetitions of an interview survey in an institution: Overall prevalence for self-injurious behavior was 10%; severe cases had longer histories of self-injurious behavior and more severe retardation; cases referred to a special "behavior modification program designed to control self-injurious behavior" were judged improved, so that some individuals identified in one year were not exhibiting self-injurious behavior in the second or third years. The fact that the number of persons identified each year was constant seems, therefore, rather unusual.

In what is probably the most comprehensive review available, Baumeister (1978) makes a number of important points regarding stereotyped movements. He points out that referring to such behavior as self-stimulation implies a theoretical assumption regarding the behavior's purpose or function. He reported data on persistence, such as the finding that there were no cases in his institutional survey in which a resident who exhibited stereotyped behavior on admission subsequently did not. He points out that variable baselines in such behaviors are the norm, rather than a reflection of the unreliability of observation, and describes a number of social and general environmental conditions that seem to relate to lowered or elevated levels of stereotypic behavior. Relying on data derived from the AAMD Adaptive Behavior Scale, Borthwick, Meyers, and Eyman (1981) reported that institutional residents display much higher levels of disruptive (e.g., "damages property"), self-injurious (e.g., "does physical violence to self"), and stereotyped (e.g.,"rocks back and forth") behaviors than those in any other type of facility. Whether this is a consequence or a cause of institutionalization cannot be determined retrospectively, of course.

Probably the most interesting insight into this rather old problem of cause and effect has been provided by Landesman-Dwyer in studies of the match between types of environments and types of behavior. She observed adult residents over a three month period who did not exhibit severe behavior problems and identified five different patterns or clusters of adaptive behavior. These clusters had considerable significance for the adaptation of clients to changed living environments and deinstitutionalization. One cluster was characterized by high levels of stereotypic behavior, non-verbal sounds, little mutual interaction and little visual exploration of the environment; this group was not responsive to environmental changes that were thought valuable for promoting social and adaptive behaviors. Nevertheless, these handicapped persons has spent the greater part of their preschool and school-age years in institutional settings. Our investigation of a statewide sample of severely handicapped children from very early identification through a period of time in which they received adequate, community based and educational services might provide some information on the natural history of behavioral repertoires displayed by persons who live in the kinds of environments enjoyed by nonhandicapped persons.

Data Analysis and Results

Throughout these studies described above there were various themes: There has been the attempt to categorize the very diverse behaviors exhibited by some severely retarded individuals; the attempt to see how the presence of certain behaviors interfered with community adjustment or were related to other behavioral characteristics; concern with the prevalence and incidence of excess behaviors in this population; and interest in the longitudinal or developmental changes in such behaviors over time. It should be made clear that our normative study was not a study of prevalence. We know from our surveys, our clinical and educational contacts and the reputation of the project in the community that it was unlikely that any child between the ages of 3 and 9 exhibiting high rates of excess behavior would not have been brought to our attention during the project (at least on the island of Oahu), but a prevalence study must survey the entire relevant population or a random sample thereof and cannot be based upon referrals. We do refer to this as the normative study, however, because it does allow some index of the percentage of occurrence of the types of behaviors we trained our observers to identify in a group defined by the presence of some kind of excess behavior. To the best of our knowledge, the total sample of experimental and comparison children reported throughout our studies represent all severely handicapped children between the ages of birth and nine who additionally exhibited multiple behavior problems, whether "serious" or more minor in nature. Thus, to some extent, their repertoires and behavioral histories across the data collection period of the project does represent normative data on a representative sample of severely handicapped children being served in community settings.

Sample Analysis of Largest Number of Children

We divided the analysis of the data in two ways. One way was to look at the children, in the comparison group for each school semester starting in the spring of 1980. In each of the five semesters the composition of the children in the comparison group varied: Some children left the islands, some were added to the group, some were experimental children who could no longer be observed in detail. This method maximizes the number of children making up each semester's group, but does not allow one to consider changes over time as the subjects change over time as well; in a way, this can be seen as a five-fold replication of a sample of children exhibiting excess behavior in a given semester in a typical school year. As the data across the five replications are voluminous, we have selected for the purposes of this report the third period of observation in which a total of 32 children were included.

Some behaviors on our excess list that were not exhibited by any of these 32 children consist of:

04 Spinning self
17 Sky or table writing
22 Pulling eyelash
38 Choking self
39 Digging in rectal area
42 Pulling other's hair
44 Pulling or pushing other
45 Pinching other
46 Choking other
47 Smelling other
48 Mouthing other
49 Biting other
50 Shadow play
74 Rumination
76 Chewing/sucking hair
77 Pica
83 Smearing feces
85 Breaking objects
87 Tearing or ripping material

All the other behaviors on our list were shown by at least one student on one occasion. Some of the very low prevalence responses were seen very briefly in one or two of the children only. The number of behaviors that occurred in one percent or more of the time in the sample (i.e., the total duration of the behavior of all children showing it divided by the total number of children in the sample) are shown in Table 5-1. Certain generalizations can be made regarding these data. Situational effects are very striking. As we have seen over and over again, excess behavior is more frequent in free play situations than under instructional conditions. This is not an artifact of teachers imposing physical restraints on the child or interfering with the occurrence of an excess directly as part of an intervention program (e.g., by briefly holding the child's hands at his/her sides prior to presenting individual instructional trials) since observers would enter a special code when such physical or topographical prevention occurred; the percentage duration figure would then be automatically adjusted to delete the "restraint" time period from all calculations by our computer programs. On the few occasions in which an excess behavior was more apparent in one-to-one instruction (Situation 2), the behaviors could be considered more appropriate; for instance, although clapping, signing and vocalizing can be excess behaviors, it is likely

96

Table 5-1: Most frequently occurring behavior in the comparison sample (32 children) derived from one observation period (as third in a series of five).

Behavior	Situation	n*	Mean Percent Duration	Overall Mean Percent Duration**
93 Finger/hand posture	1	11	21.2	7.3
	2	13	4.3	1.7
	3	10	6.6	2.1
25 Mouthing body parts	1	17	8.8	4.7
	2	15	5.7	2.7
	3	17	10.4	5.5
26 Mouthing objects	1	11	14.7	5.1
	2	8	4.8	1.2
	3	8	8.4	2.1
14 Finger rubbing	1	13	13.6	5.5
	2	3	11.7	2.9
	3	9	8.7	2.4
13 Finger flicking	1	18	18.7	10.5
	2	19	11.7	6.9
	3	17	16.1	8.5
01 Body rocking (sitting)	1	13	12.4	5.0
	2	7	5.4	1.2
	3	11	10.6	3.6
15 Vocalization	1	27	15.3	12.9
	2	29	12.2	11.1
	3	20	12.7	7.9

* Number of children in that situation showing the behavior

** Total percent duration divided by the total number of children (N=32)

Table 5-1: (Continued)

Behavior	Situation	n*	Mean Percent Duration	Overall Mean Percent Duration**
86 Tongue Movements	1	12	6.3	2.4
	2	11	12.1	4.2
	3	11	9.8	3.4
11 Hand/arm flapping	1	13	11.7	4.7
	2	10	3.2	1.0
	3	9	5.8	1.6
60 Cry, no tears	1	8	13.1	3.3
	2	9	7.8	2.2
	3	4	14.9	1.9
78 Head weaving	1	8	10.4	2.6
	2	8	7.6	1.9
	3	7	13.9	3.1
15 Finger tapping	1	7	12.6	2.8
	2	8	2.2	0.6
	3	8	4.9	1.2
10 Hand clapping	1	3	9.4	0.9
	2	8	10.9	2.7
	3	9	10.5	3.0
07 Leg swinging	1	5	9.4	1.5
	2	9	7.9	2.2
	3	8	4.6	1.2
23 Rubbing face	1	12	6.7	2.5
	2	14	4.9	2.2
	3	14	6.2	2.7

Table 5-1: (Continued)

Behavior	Situation	n*	Mean Percent Duration	Overall Mean Percent Duration**
69 Spontaneous Verbalization	1	4	6.9	.9
	2	7	13.0	2.8
	3	4	16.7	2.1
66 Vocalization supraglottal	1	7	3.7	.8
	2	7	10.1	2.2
	3	4	7.0	.9
19 Pulling clothes	1	5	18.1	2.8
	2	3	6.9	.6
	3	4	8.2	1.0
12 Hands pressed on ears	1	2	5.2	.3
	2	2	4.6	.3
	3	2	34.0	2.1
75 Teeth grinding	1	2	39.3	2.5
	2	1	48.1	1.5
	3	1	3.2	0.1
89 Pounding on objects	1	9	7.0	2.0
	2	9	4.9	1.4
	3	8	4.7	1.2
28 Scratching skin	1	6	9.4	1.8
	2	7	1.6	.3
	3	7	5.3	1.1
21 Rubbing eyes	1	5	2.5	.4
	2	8	2.0	.5
	3	6	7.6	1.4

Table 5-1: (Continued)

Behavior	Situation	n*	Mean Percent Duration	Overall Mean Percent Duration**
64 Blowing	1	5	9.7	1.5
	2	6	6.9	1.3
	3	3	6.8	.6
80 Gesture/sign	1	2	4.4	.3
	2	3	14.6	1.4
	3	4	8.6	1.1
70 Facial grimace	1	2	6.1	.4
	2	6	4.7	.9
	3	8	5.3	1.3
72 Staring/gazing	1	4	9.4	1.2
	2	3	1.7	.2
	3	4	2.7	.3

that appropriate behaviors were actually responsible for much of the percentages recorded under these categories. In any further use of this behavior observation system, therefore, these categories require revisions to remove this ambiguity; we often regretted including some of these categories in the excess group. Similarly a few behaviors such as facial grimace, etc., were more frequent during group instruction; it is possible that these excess behaviors were, in fact, rudimentary social responses.

Obviously some behaviors have relatively high overall percent durations because a few children engaged in them much of the time and many children engaged in them some of the time. Very few excess behaviors were revealed by more than half of the children, indicating the enormous variety and range. The most frequent behaviors were ones which were developmentally typical of this group of younger children, such as mouthing objects, finger mannerisms, rocking and so forth. More elaborated responses such as aggressive behaviors were rare (although one incident of "43: hitting other" was noted in one child).

To give some indication of the average occurrence of our "environmental" codes, the mutually exclusive codes in various categories of child, teacher, and other events (Observer 2 System), Table 5.2 gives the information on number of occasions the event was noted, the average percent duration of those occasions, and the overall average of all possible times. Some of the codes and their situational differences represent environmental differences and provide a simple validation of the Observer 2 System. For example, teachers are rarely present in free play, peers are never "not present" in group , food was never present in free play, etc. On the affect category, the children are coded as predominantly neutral (87% of the time, and either happy or sad approximately 7% and 4% of the time respectively). Visual attention during one-to-one instructions seemed rather good: 49% of the time on nearby objects and over 18% of the time on the teacher. Instruction took place with children sitting most of the time, as we had requested. The teachers were able to keep the students on-task 43% of the time in one-to-one instruction, only 27% of the time in group instruction. In the latter case, excess behaviors were high (44%) although still not as high as in free play (66%). As children in formal instruction were able to spend a quarter of the time on-task despite simultaneous excess behaviors, we must recognize that many of these stereotyped, inappropriate behaviors do not seem to interfere with teaching -- although whether they interfere with learning cannot be determined from our da' . As we have noted before, we were impressed with how little of the teacher's time could be described as approving, although if codes 5 and 6 are combined, 18% of the one-to-one and 10% of the group instruction situations involved teacher approval.

Table 5-2: Environmental event codes (Observer2 system) for comparison sample (N=32) for third observation period

		Situation 1			Situation 2			Situation 3		
		n	Mean PD	Overall	n	Mean PD	Overall	n	Mean PD	Overall
AFFECT	0 Neutral	31	90.0	87.3	32	86.0	87.5	29	87.6	79.4
	1 Happy	16	9.5	4.8	24	13.5	10.2	19	12.0	7.1
	2 Sad	11	14.1	4.8	10	13.1	4.1	8	16.4	4.1
	3 Angry	←				NO OCCURRENCE				→
VISUAL REGARD	0 Space	14	34.1	14.9	17	15.8	8.4	16	17.6	8.8
	1 Distant	29	38.2	34.6	28	28.8	25.2	28	24.5	21.4
	2 Proximal	27	47.0	39.7	32	48.7	48.7	29	40.9	37.0
	3 Self	13	7.3	2.9	9	6.9	1.9	12	11.5	4.3
	4 Teacher	5	10.8	1.7	23	18.5	15.9	26	11.9	9.7
	5 Adult	6	10.7	2.0	7	4.0	.9	8	4.6	1.1
	6 Peer	1	1.9	.3	1	.6	.1	21	10.0	10.9
	7 Eyes closed	5	6.1	3.0	9	5.5	4.2	9	5.7	4.4
POSITION IN SPACE	0 Lying	12	42.7	32.3	6	5.0	2.5	1	3.0	.5
	1 Sitting	25	61.2	47.8	32	97.3	97.3	29	98.3	89.1
	2 Crouching	13	18.6	17.4	3	1.7	.5	1	.8	.1
	3 Standing	15	30.1	14.6	7	6.6	1.4	5	8.2	1.3
	4 Crawling	13	5.6	2.3	1	2.0	.6	1	.8	.3
	5 Walking	14	17.9	7.8	1	3.9	.1	3	1.5	.1
	6 Running	5	8.2	1.3	0	0	0	0	0	0
CHILD	0 Neutral	23	19.3	13.9	25	9.9	7.8	21	16.9	11.1
	1 Excess	30	65.8	61.7	29	26.6	24.1	27	44.1	37.2
	2 Excess & On Task	17	15.1	8.0	32	25.6	25.6	27	24.4	20.6
	3 On task	20	16.6	10.4	31	43.2	41.9	25	27.4	21.4
	4 Off task-appropriate	8	11.4	2.8	4	5.0	.6	3	3.0	.3
TEACHER	0 Not present	31	97.9	94.9	4	3.1	3.8	0	0	0
	1 Neutral	2	7.1	1.7	29	28.9	18.8	27	17.7	15.1
	2 Neutral & Contact	6	4.6	.9	29	21.1	19.1	23	12.5	9.0
	3 Neutral Mand	2	6.4	.4	27	25.8	21.7	25	18.1	14.1
	4 Neutral Mand & Contact	2	4.6	.3	28	16.3	14.2	20	9.0	5.6

102

Table 5-2: (Continued)

		Situation 1			Situation 2			Situation 3		
		n	Mean PD	Overall	n	Mean PD	Overall	n	Mean PD	Overall
TEACHER (contd.)	5 Approval	0	0	0	26	10.0	7.4	24	6.9	5.8
	6 Approval & Contact	0	0	0	29	8.2	7.4	23	3.7	2.7
	7 Disapproval	0	0	0	13	2.5	1.8	7	2.3	1.2
	8 Disapproval & Contact	0	0	0	11	3.2	1.1	7	2.4	.5
	9 Ignoring/ Attending to Peer	0	0	0	2	11.2	.7	29	42.0	38.0
PEER	0 Not Present	31	99.9	96.8	32	99.6	99.6	0	0	0
	1 Neutral	1	1.1	0	3	3.7	.3	29	95.8	86.8
	2 Approach	0	0	0	0	0	0	16	7.4	3.7
	3 Avoid	0	0	0	1	.6	.1	2	1.2	.4
	4 Aggression	←				NO OCCURRENCE				→
	5 Protest Minus Aggression	0	0	0	0	0	0	1	.8	.1
OBJECTS	0 None	1	5.4	.2	2	43.5	2.7	1	3.3	.1
	1 Program Materials	1	1.8	.0	28	90.7	79.4	27	94.3	79.6
	2 Toys	31	99.8	96.6	5	12.6	1.9	3	35.1	3.3
	3 Food/drink	0	0	0	1	81.5	2.5	1	[illegible]8	.0
	4 Program Materials & Food	0	0	0	7	55.1	12.5	5	48.6	7.6
	5 Toys & Food	←				NO OCCURRENCE				→
	6 Materials & Toys	0	0	0	1	42.6	1.3	0	0	0
	7 Materials & Toys & Food	←				NO OCCURRENCE				→

Longitudinal Analysis Over 2 1/2 Years (3 Observation Periods)

If we wish to examine changes in excess behaviors over time, we need to consider the same group of children looked at longitudinally. For reasons already explained, only 16 children spanned the full three years of the project. The three points of observation presented here correspond roughly to the first, second, and third years of the investigations. In some cases variability within children might reflect different teachers and programs over the same school year. However, quite a number of the children had the same teacher all three years and if any taped sessions was considered particularly atypical it was usually re-taped later (such as a child sleeping through the entire session or having a sustained but unusual tantrum).

Table 5.3 displays the data for Situation 1. Once again, the group of 16 children produced many (64) excess behaviors at one time or another across the three years of data collection. Most of these behaviors were exhibited by a small number of children and for relatively short overall durations. All three indices of excess behaviors showed reductions over time: The number of children showing a behavior (i.e., the number of occurrences of all behaviors in the group of 16) decreased from 173 to 124 to 93, or decreased from 2.7 children per excess to 1.4 children per excess. The mean percent durations of these children showing the behavior at the given observation time remained very much the same (5.8, 5.8, 5.5) and the overall percentages (thinking of those children not showing the behavior or showing it 0 percent) decreased from 1.63 to 1.28 to 1.13. It is difficult to know whether these changes are clinically meaningful. However, it is encouraging that excess behavior in general appeared to be less frequent in the comparison group over time.

Table 5.4 presents a number of additional variables of potential interest to considerations of change over time. These variables partly answered the question just posed: There is no strong evidence that this group of 16 children were by the third year showing more positive affect or more on-task behavior. Teacher approval rates remained very much the same (but of course the teacher could have changed) and the peers were showing less, not more, approach behavior).

Generally, these results indicate to us the need to supplement conventional child outcome data with various samples of this kind. However, if these data were to be used for evaluating classrooms, teachers, or children's outcomes, more frequent samples would certainly be required.

Clustering of Excess Behaviors Within Children

By gathering observational data on the comparison children it is also possible to attempt to see how behaviors tend to co-occur in children, that is, whether they fit into meaningful clusters or "syndromes" (see

TABLE 5-3: LONGITUDINAL CHANGES IN EXCESS BEHAVIORS IN A SAMPLE OF THE COMPARISON GROUP FOR SITUATION 1-FREE PLAY

BEHAVIOR	1st YEAR			2nd YEAR			3rd YEAR		
	n	Mean Percent Duration	Overall Percent Duration	n	Mean Percent Duration	Overall Percent Duration	n	Mean Percent Duration	Overall Percent Duration
1	6	9.9	3.7	6	12.6	4.7	5	10.5	3.3
2	1	11.9	7.4	2	2.5	.3	0	0	0
3	0	0	0	2	1.7	.2	0	0	0
5	6	2.5	.9	1	1.4	.1	0	0	0
6	0	0	0	0	0	0	1	.4	0
7	3	1.8	.3	1	11.6	.7	0	0	0
9	1	.9	.0	0	0	0	0	0	0
10	4	1.8	.5	1	3.1	.2	2	19.5	2.4
11	7	10.1	4.5	6	4.2	1.6	3	18.3	3.4
12	1	3.6	.2	0	0	0	1	4.2	.3
13	10	23.6	14.7	8	15.6	7.8	5	11.5	3.6
14	4	13.1	3.3	6	26.0	9.8	7	28.7	12.6
15	5	11.9	3.7	5	16.2	5.1	3	11.9	2.2
16	1	4.9	.3	1	1.7	.1	0	0	0
17	1	.7	.0	0	0	0	0	0	0
18	1	2.7	.2	0	0	0	0	0	0
19	2	1.3	.2	1	19.9	1.2	0	0	0
21	1	1.3	.1	4	2.3	.6	1	7.1	.4
23	4	3.5	.9	5	12.0	3.7	5	3.1	1.0
24	0	0	0	1	.6	.0	1	10.4	.6
25	9	13.9	7.8	9	8.2	4.6	4	26.9	6.7
26	4	5.2	1.3	4	4.0	1.0	2	1.8	.2
27	1	2.3	.1	0	0	0	0	0	0
28	4	1.6	.4	2	15.7	2.0	1	1.0	0
29	1	12.0	.7	0	0	0	1	1.8	.1
30	0	0	0	0	0	0	1	1.2	0
31	1	1.4	0	1	14.2	.9	1	1.3	0
32	0	0	0	1	1.3	0	0	0	0
35	0	0	0	1	7.4	.5	0	0	0
36	1	4.25	.3	0	0	0	1	.1	0
37	2	67.5	8.4	1	2.1	.1	0	0	0
41	0	0	0	1	1.8	.1	1	.1	0

(TABLE 5-3 Continued)

BEHAVIOR	1st YEAR			2nd YEAR			3rd YEAR		
	n	Mean Percent Duration	Overall Percent Duration	n	Mean Percent Duration	Overall Percent Duration	n	Mean Percent Duration	Overall Percent Duration
43	0	0	0	0	0	0	0	0	0
51	3	3.0	.6	1	1.4	0	0	0	0
52	2	2.9	.4	1	7.2	.4	2	1.8	.2
53	1	5.4	.3	1	1.8	.1	0	0	0
54	4	3.1	.8	1	1.4	.1	3	5.1	.9
55	4	2.9	.7	1	.8	0	1	6.7	.4
56	5	1.6	.5	2	1.2	.2	1	1.2	0
57	4	4.6	1.1	0	0	0	1	1.7	.1
58	0	0	0	0	0	0	1	.8	0
59	0	0	0	0	0	0	1	1.0	0
60	3	33.1	6.2	5	16.8	5.27	1	24.7	1.5
62	2	1.9	.2	3	9.1	1.7	1	1.6	.1
64	1	.8	0	0	0	0	0	0	0
65	13	14.0	11.4	11	14.4	9.9	12	15.7	11.7
66	4	1.5	.4	3	2.1	.4	1	.9	.1
67	0	0	0	1	2.2	.1	0	0	0
68	0	0	0	0	0	0	0	0	0
69	0	0	0	1	5.3	.3	0	0	0
70	1	.4	0	1	1.0	.1	1	1.0	.1
72	3	23.1	4.3	2	21.1	2.6	1	16.14	1.0
73	1	2.8	.2	1	5.6	.3	0	0	0
75	2	15.6	1.9	2	39.3	4.9	0	0	0
78	4	2.7	.7	3	3 9	.7	4	4.1	1.0
79	1	1.5	0	1	3.2	.2	1	1.9	.1
80	0	0	0	1	5.4	.3	1	.8	0
86	7	6.3	2.8	1	.4	9	4	6.4	9.11
88	2	3.2	.4	0	0	0	2	10.6	1.3
89	10	4.8	3.0	3	18.5	3.5	2	4.6	.6
90	0	0	0	0	0	0	1	1.3	.1
91	4	6.9	1.7	2	2.7	.3	3	3.9	.7
92	2	3.8	.5	2	6.3	.8	0	0	0
93	9	11.0	6.2	4	17.9	4.5	2	51.6	6.5
TOTALS	173	369.75	104.2	124	375.1	81.9	93	353.3	72.3
N=64 (Behavior)	2.7	5.8	1.63	1.9	5.86	1.28	14	5.5	1.13

Quay [1972] for a general discussion on early efforts to empirically define such patterns based upon checklists and other data sources obtained for large numbers of children). The first successfully gathered session of observation was identified for each child (regardless of when in the research project it was obtained) and the percent duration of each variable coded represented the child's score on each variable. These data were then factor analyzed and also subjected to a cluster analysis; separate analyses were conducted by situation. Perusal of these analyses did not reveal results which are readily interpretable. First, we did not find higher order clusters of such behaviors as self-stimulatory or disruptive behavior. This may be because, as already demonstrated, the incidence of aggressive, disruptive, or self-injurious behavior in this group of children was quite low. Secondly, although there were a few high loading items on each factor as can be seen from Table 5.5, the factors did not account for a very large proportion of the variance (6.4% for Factor 1, 4.1% for Factor 7). Thus, the patterns may well arise from the idiosyncratic patterns of perhaps only one or two children who were exhibiting unique behavior to a large degree. Unless all children exhibit all behaviors to some extent, the factor analysis merely recaptures the cluster of behaviors exhibited by certain children in the sample. We are carrying our different clustering methods to try to resolve this issue.

Discussion and Conclusions

Much still needs to be learned of the behavior patterns of severely handicapped young children in natural community and educational environments -- by which we mean classrooms and homes and other non-institutional settings in which they are expected to function. Children in such environments, as we have repeated emphasized, are not in neutral settings from the point of view of intervention with excess behaviors. Teachers, explicitly and implicitly, initiated procedures in which learning could occur. Many of the more serious (in terms of physical harm) excess behaviors are simply not seen in these students, nor do they develop over a three-year time period. On the other hand there is not much indication in these findings that substantial, clinically or educationally significant change takes place over the same time period, although the behaviors which did occur did not deteriorate further. Further longitudinal studies are urgently needed in which alterations in excess behaviors and positive response development are simultaneously monitored. This aspect of our project provides, hopefully, an impetus to further studies and suggests that observational samples are a rich source of new data, particularly if that individual variability can be adequately taken into account. Differences in setting, which appear quite minor (such as one-to-one versus group instruction) have much greater impact on frequencies of these behaviors than does the passage of time. In the next chapter we will examine the impact of specific experimental manipulations.

TABLE 5-4: LONGITUDINAL CHANGES IN SLECTED CODE CATEGORIES OVER 3 YEARS ON OBSERVATIONAL SAMPLES:

BEHAVIOR/CODE CATEGORY	SITUATION	FIRST YEAR N.	FIRST YEAR MEAN PD	FIRST YEAR OVERALL	SECOND YEAR N.	SECOND YEAR MEAN PD	SECOND YEAR OVERALL	THIRD YEAR N	THIRD YEAR MEAN PD	THIRD YEAR OVERALL
0 AFFECT NEUTRAL	2	16	77.8	77.8	16	88.3	88.3	15	83.8	78.5
1 AFFECT HAPPY	2	9	29.2	16.4	13	12.7	10.3	7	20.2	8.8
4 VISUAL REGARD-TEACHER	2	12	12.8	9.6	10	19.3	12.4	7	6.2	2.7
2 CHILD EXCESS + ON TASK	2	15	35.9	33.6	16	29.3	29.3	15	28.2	26.5
2 CHILD EXCESS + ON TASK	3	13	14.6	11.8	13	24.6	20.0	14	17.7	15.4
3 CHILD ON TASK	2	14	26.3	23.0	15	39.3	36.9	15	25.1	23.6
5 TEACHER APPROVAL	2	11	7.6	5.2	12	6.3	4.8	13	5.7	4.6
5 TEACHER APPROVAL	3	8	4.4	2.2	11	4.4	3.0	13	4.3	3.5
6 TEACHER APPROVAL + CONTACT	2	10	6.2	3.8	15	7.7	7.2	13	9.2	7.5
6 TEACHER APPROVAL + CONTACT	3	10	3.7	2.3	13	3.5	2.9	12	4.5	3.4
2 PEER APPROACH	3	5	13.9	4.3	7	12.5	5.5	8	5.8	2.9

Table 5-5: Factor analysis of 49 children's excess behaviors as exhibited on one occasion of observation (Promax solution)

SITUATION 1

FACTOR I		II		III		IV		V		VI		VIII	
Hand clapping	.98	Stripping	.98	Body impress	.95	Object banging	.90	Ear poking	.93	Throwing objects	.90	Hand flapping	.81
Pulling clothing	.97	Lint picking	.98	Grunt	.95	Object flicking	.90	Biting other	.91	Spinning self	.89	Finger posture	.76
Biting self	.97	Temper tantrum	.98	Rumination	.95	Climbing	.90	Hands on ears	.79	Pica	.88	Tongue movements	.74
Clicking voc.	.91	Toe walking	.97	Jerky movement	.70	Bear walking	.90	Facial grimace	.79	Staring, gazing	.71	Spinning self	.61
Head weaving	.91	Jumping	.93			Rubbing eyes	.61	Head slapping	.73	Mouthing body parts	.68		
Rubbing face	.85												

Chapter 6

Intervention Studies

A smaller number of "experimental" children were identified as subjects[1] for the series of intervention studies. Procedures used to identify these subjects were described in Chapter 2. The series of individual interventions was conducted for periods of time ranging from several weeks to nearly two years. Each investigation involved a specific hypothesis regarding response-response relationships in the child's repertoire, including tests of "keystone" as well as reciprocally related behaviors. In most cases, a change in instructional strategies was involved within the context of existing IEP objectives, hence a new IEP meeting was not required (though parent permission was obtained for each of these specific intervention plans). In some cases (e.g., Child 05 described in this chapter), a major program change was made and we thus participated in planning a new IEP with the child's parents and the school's instructional team.

The intent of each of these intervention studies was to investigate strategies which would allow teachers and other clinicians to more efficiently establish positive outcomes through optimal multiple positive effects which they could predict and thus monitor. One major emphasis of these studies was to identify least intrusive behavioral interventions, involving primarily changes in teacher behavior and instructional arrangements. The other major emphasis of this component was to investigate decreases in particular excess behaviors as a consequence of learning an incompatiole and/or functionally alternative skill. Each of the studies was planned following the exhaustive review of intervention research with the particular excess behavior/s and in close consultation with the child's teacher and other relevant caregivers (e.g., the parents). This chapter presents detailed descriptions of three of these intervention studies; similar detail for all the studies and further information for these can be found in Evans and Voeltz (in preparation).

The experimental design of each study was single-subject (e.g., multiple baseline, reversal, interrupted time series) but included:

(1) multivariate dependent variables such as factor scores as well as traditional percentage occurrence changes in the frequency/duration of single target responses (e.g., Child 06 described in this chapter); and

(2) tests of statistical significance for differences in behavior across experimental (treatment) phases (e.g., Child 05 and Child 08 described in this chapter).

---

[1]Each of the three children described in the case studies has been given a "pseudonym" first name, i.e., the names used throughout this chapter are not their real names.

CHILD 08 - KATHY

This case study will be reported in a descriptive style in order to reveal some of the problems of logic, design and data analysis that we were continually working to overcome. In some ways it serves as a case study in scientific methods relevant to the study of a clinical case.

Introduction

Kathy was 5 years 11 months of age when first videotaped and observed by the project in December 1979. During the first project year she was in a self-contained class for SMH children on a public elementary school campus located near Pearl City on the island of Oahu. During the school years 1980-81 and 1981-82 when the data reported here was gathered, she was placed in a similar program in a public school on the leeward side of the island because she was now looked after by her father and paternal grandmother, her mother having separated from Kathy's father and left Hawaii. Kathy had been known to the principal investigators since she was three years old and in a State of Hawaii Health Department program for the early identification and assessment of children who are presented as "diagnostic enigmas". She had been the subject of an investigation by a Master's degree Special Education student while at that setting; this project had involved an attempt to decrease Kathy's hand mannerisms which will be described later. Although partially successful, an increase had been noticed in teeth grinding and blowing. Thus Kathy had already been identified as a potential interesting child for further investigation of response relationships.

At the diagnostic center she had been described as severely retarded with some autistic-like behavior. The latter referred to her apparent non-compliance in certain training situations, her limited social reactions, and her rather numerous excess behaviors. In 1980 at the age of six, we obtained a total raw score of 76 on the TARC. Kathy is toilet regulated, and is ambulatory although she has a cumbersome gait which suggests mild cerebral palsy. She has no language or communication skills, and minimal self-help skills. Perhaps her most major limitation with respect to this latter domain was her failure to make deliberate manual responses: she did not pick up objects placed in front of her (including food items when hungry) and would not grasp objects placed in her hand, or reach for objects. The only functional manual response ever observed was that she would extend her arms and use her hands to support herself and would do this rapidly on a balance ball or similar apparatus to maintain her upright position. Much of the time she exhibited a hand mannerism in which her left hand, with fingers extended or slightly bunched, would be brought up against the open palm of her right hand in a repetitive, clapping movement; both thumbs were fixed. Sometimes this action involved grasping of her fingers, so that the way this response was coded varies slightly across time and between observers.

Baseline Information

An initial assessment of Kathy's excess behavior was carried out by having her teacher and parent (in this case, her mother) complete a checklist of excess behaviors that was used in our coding system. As these behaviors were not defined in detail, some discrepancies between teacher and parent were noted; however, after resolving these discrepancies, the following behaviors were reported:

Body rocking (both sitting and standing)
Jerky movement
Hand clapping ("all the time" reported mother)
Finger flicking ("occurs if one hand is restrained and she can't clap" - teacher)
Cry - tears
Shriek/scream
Blowing
Vocalization (supra-glottal and babbling)
Facial grimace
Staring/gazing
Teeth grinding ("a lot of times" - mother)
Temper tantrum
Rubbing eyes or face ("when tired or crying or nose is running" - mother)

The first phase of formal data collection with Kathy took place in spring of 1980. During this period we were refining our data-gathering methods, our analytical procedures, or improving observer agreement and response definitions. We had expected to obtain relatively steady response baselines and to examine the relationships among behaviors over time. However, when percentage durations of her more frequent behaviors were plotted over observation sessions, there was marked variability and few or no discernable patterns of responding. One plausible explanation of this is that the intensive, real-time analysis of a short segment of time sampled from one or two days per week reveals a variability in behavior that is masked by time samplings (interval recording) methods currently in vogue. We tend to forget that published baselines of some inappropriate behaviors may appear regular (or "steady") because a sampling unit has been selected which is insensitive to the true variability in behavior.

Another source of richness of detail we were better able to discern than to deal with was the large number of excess behaviors that the observers were able to note in Kathy (and other children investigated). As we pointed out in our original proposal, excess behaviors have an origin or natural history that is little understood and has never been systematically investigated, so it is important to capture fleeting or momentary behavior that might increase in intensity later. However from the analysis point of view, it is difficult to summarize a behavior that an observer may have noticed for a few seconds; in Kathy's case, these were responses like tongue movements, scratching skin, mouthing body parts,

object flicking. To represent the dominant behaviors of the first observation period, any response occurring for 2% of the total duration of any observation session was listed, and in the following table are shown the mean percent durations of these behaviors over the first three observations and over the last three observations of the period, by situation.

TABLE 6.1 KATHY: FIRST OBSERVATION PERIOD

Mean percent durations of most prominent behaviors for first three observations and for the last three observations.

Behavior	Situation 1		Situation 2		Situation 3	
	First	Last	First	Last	First	Last
01 Body Rocking	35.6	34.0	0.4	2.9	4.2	1.8
10 Hand Clapping	53.1	51.1	29.3	1.9	36.3	36.9
13 Finger Flicking	11.4	52.1	20.3	38.1	40.1	47.9
23 Rubbing Face	1.5	0.0	0.5	1.8	1.4	1.3
25 Mouthing Body Parts	1.6	1.7	0.3	1.2	1.6	1.7
60 Cry, No Tears	0.0	0.0	6.6	1.2	1.4	0.4
62 Shriek/Scream	0.6	0.2	0.0	0.2	0.9	0.0
64 Blowing	15.9	0.0	3.2	0.0	8.7	0.3
65 Vocalizing	31.4	27.1	11.1	8.8	4.8	4.3

Although no formal statistical comparisons have been made, it can be seen that the only behavior which seemed to substantially decrease over the school term was Blowing. Situation 1 produced very much greater body rocking than the two instructional situations, and somewhat more of her hand/finger mannerisms; because of definitional difficulties, it is best to consider behaviors 10 and 13 together as a unit, not concurrent responses.

Second Baseline Phase

It must be remembered that our intervention research took place within the context of children placed in public school settings. Our role as experimenters had to be tempered by our more veridical role as consultants. It is in accordance with this reality that our interests in target selection, teacher decision-making and judgement, and educational validity form a cohesive whole. Our second year of observation of Kathy finds her in a new school with a new teacher who was cooperative, but not controlled in any way by our project staff. At the beginning of

the school year this teacher introduced a program for reducing the finger/hand mannerism. Her reasons for choosing this target were, according to an interview with her, that the behavior disruoted all other learning, interferred with other activities, and was considered negative by her parents. The intervention chosen was a non-contingent physical restraint in which each half-an-hour Kathy's hands were to be held down for five minutes. The data to be recorded (which it transpiredwere not systematically kept) was to time how long after that restraint she would keep her hands down.

Table 6.2 provides the summary results of observation periods in the middle of the Fall term (when we resumed observation) and three observation periods at the end, just before our planned intervention.

TABLE 6.2 KATHY: SECOND OBSERVATION PERIOD

Mean percent durations of previously identified behavior for first three observation sessions and for the last three observation sessions prior to intervention.

Behavior	Situation 1		Situation 2		Situation 3	
	First	Last	First	Last	First	Last
01 Body Rocking	41.4	78.5	1.1	0.0	18.7	0.0
10 Hand Clapping	0.0	0.0	0.0	0.0	1.5	0.0
13 Finger Flicking	72.8	82.4	50.7	64.5	43.6	64.0
23 Rubbing Face	2.2	5.3	9.5	0.0	7.1	5.6
25 Mouthing Body Parts	0.0	0.0	0.0	0 0	4.5	6.4
60 Cry, No Tears	0.0	0.0	2.5	0.0	5.8	5.5
62 Shriek/Scream	0.0	0.0	0.0	0.0	0.0	0.0
64 Blowing	19.7	53.1	7.3	25.0	9.9	4.3
65 Vocalizing	14.3	11.6	3.1	7.1	2.8	0.0

First Intervention

After observing Kathy's program and considering effects on behavior, it was very apparent that much of her educational curriculum needed to be revised. It was noted that Kathy enjoyed being hugged and treated affectionately, and would often walk over to an adult and stand close to him/her, rubbing her hands. One instructional objective was to teach her

to cross her arms in front of her as a gesture indicating she wanted a hug. A second program suggested was to introduce training on the electronic flute during group instruction (Wuerch & Voeltz, 1982, pp. 43-48). The major intervention program, however, was to try to obtain a simple, functional manual operant from Kathy. There were two interesting facets to such a program. One was that if such a response could be taught to Kathy, would its acquisition result in a reduction in the excess behavior of hand movements? This is a complex question, in terms of experimental design, as the attempt to teach the incompatible response requires some kind of direct intervention as well, when the excess behavior is of such high frequency. The second facet is that we believe it is important to make the actions of any intervention potentially functional. By this we mean that the response requirement or the intervention be a normal action. For example, brief physical restraint of Kathy's hands would result in her holding her hand flat on the table top; this is not a normative response for a six-year-old, whereas grasping the side of a puzzle, bowl, or the end of an electronic flute requires stabilizing actions for many skill sequences. Thus our suggested intervention required the physical prompt to interrupt Kathy's finger rubbing and return her hands to the play activity required, which was simply moving her hands in a bowl of beans or rice. In order to give this very simple manipulation an additional function, music via headphones was made contingent upon maintaining her play with the rice -- removing her hand resulted in the music being switched off.

To monitor progress on this manipulation task, the teacher was requested to record the duration of time that Kathy maintained her hands in the bowl of rice and the number of times she repeatedly placed her hands in the bowl without physical assistance. The former measure could not be obtained as the teacher had great difficulty in switching on the music when Kathy's hands were in the rice and switching if off when she removed them. In fact, the inexact relationship between the response and the contingency may have contributed to the program's lack of success, because no instances of spontaneous placement of hands in the bowl was recorded, even after three months of daily programming with this task. Because of the recording difficulty, weekly probes were also initiated. During the probes the bowl of rice (and sometimes beans or macaroni) were placed in front of Kathy with the instruction "play with the ____ Kathy". If there was no response within one minute the stimulus was withdrawn. This was done with a shiny mobile, a cookie, the electronic flute, and a push-button toy called the Big Mouth Singer, each time with appropriate instructions to play, touch or take the object. The data recorded during these five probe trials each week indicated that Kathy made no spontaneous response with her hands and had not done so by the end of the school year in June.

Excess behavior recorded during this same three month period is shown in Table 6.3. Eighteen observation sessions were obtained. Because of continued variability in response durations, those sessions have been

TABLE 6.3 KATHY: FIRST INTERVENTION PERIOD

Mean percent durations of previously identified behaviors, averaged over three observation sessions thereby yielding six observation periods for the eighteen sessions of observations gathered.

	Situation 1						Situation 2						Situation 3					
Period:	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
01 Body Rocking	4.1	41.0	49.5	35.9	18.8	50.6	5.9	17.4	9.8	3.2	6.9	6.8	8.5	7.8	5.9	35.2	21.3	27.7
10 Hand Clapping+																		
13 Finger Flicking	50.7	40.5	36.6	62.9	37.2	32.1	71.8	49.4	70.8	44.4	50.9	39.0	78.2	66.5	87.1	58.3	62.7	41.1
23 Rub Face	3.9	4.9	11.7	14.4	8.2	13.4	10.6	3.4	10.1	19.7	1.0	1.5	10.5	2.9	18.8	0.3	0.0	7.7
25 Mouth Body Parts	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.5	0.8	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0
60 Cry, No Tears	5.6	0.0	14.4	0.0	0.0	0.7	0.8	0.0	0.7	3.3	0.0	4.1	0.0	0.0	0.5	0.0	0.3	1.5
62 Shriek/ Scream	3.1	0.0	2.0	0.0	0.0	0.9	0.0	0.0	5.4	2.5	0.0	9.1	0.0	0.0	0.0	0.6	0.0	6.4
64 Blowing	3.9	5.9	8.7	12.2	0.0	19.3	3.8	9.7	5.3	16.2	0.0	28.3	8.8	5.4	10.7	3.4	0.0	11.4
65 Vocali-zation	18.4	4.1	5.7	19.0	11.6	12.8	12.8	2.4	0.9	4.4	3.6	6.3	9.1	0.0	0.0	0.9	0.8	2.9
86 Tongue Movement*	5.6	8.6	9.9	4.5	11.2	29.9	3.2	1.0	11.3	2.1	7.4	10.9	6.1	4.7	1.7	14.4	35.7	30.4

+ This behavior deleted to assist observers; behavior is subsumed under 13 - Finger Flicking.

* This behavior added to list as increase in frequency noted.

divided into six blocks and the durations averaged for the three sessions within each block or period. Examination of this table shows a number of interesting features. Body rocking continues to be a significant response, with much more rocking being present in free play. Finger flicking, the largest response, continued, although slighly below baseline levels. Other responses originally identified continue to be present but at low levels. A new response, tongue movements, appeared and seemed to get worse as the intervention period continued.

Second Intervention

The second intervention took place during the next school year. Kathy now had a new teacher and our response clustering techniques were in place. Six sessions of additional observational data were obtained as baseline and subjected to factor analysis: One analysis of the first three sessions combined and one of the second three observation sessions combined. The results of these analyses are not listed here in detail. As we have noted elsewhere these analyses are not always easy to interpret. In the case of Kathy, the factor solution for Situation 1 was a single factor in which the highest loading items were the finger mannerism pattern (Excess 10, 13, 14), sitting and rocking (Excess 1) with a distant stare. In Situation 2 a similar pattern of inappropriate behavior was seen but in addition a further factor in which "no excess" and a number of more positive child characteristics were related to positive teacher affect.

As no particularly consistent pattern or response cluster appeared in the factor analyses of the two baseline periods, it was felt that in order to try to see how responses were related, a temporary intervention designed primarily to have a rapid effect on a particular response seemed appropriate. Two drawbacks were immediately apparent: (1) it could not be justified, following our model, to use some kind of physical restraint to reduce behavior artificially; and (2) there would be no justification for introducing an intervention with a child such as her, that is designed primarily for experimental purposes (choice of target issue).

Fortunately, however, there seemed to be some value in changing, if possible, Kathy's affective responses. The factor analyses suggested that when Kathy's affect was positive, excess behaviors were less. Secondly, there was a persistent belief among those who had worked with her that she is "negativistic" or noncompliant and that she is capable of higher-level responses than she reveals. These two considerations suggested that a suitable short-term influence on Kathy's behavior would be to introduce low-demand, high reinforcement, interaction, with the teacher responding to her and playing with her interactively exactly as one would socially interact with a young infant.

The longer-term intervention plan for Kathy was based on a continuation of the logic of the intervention attempted the previous year. Kathy's predominant excess behavior was her hand manipulation. It was

also striking that a major deficit in her repertoire was a lack of any reaching, grasping, or manual manipulation, even of objects she appeared interested in (moves towards and stares at). Thus it was proposed that if Kathy could acquire a simple manual operant in which she could reliably produce some manual response, this would be the first step towards increasing play with objects and decreasing the hand/finger excess behaviors.

Interventions:

(1) Short-term influence: non-demanding social interaction. This "intervention" consisted solely of infant-type social interaction. The goal of the procedure was to make Kathy happy, that is, make her smile or laugh and appear to enjoy the interaction. Teacher behaviors were limited to (i) talking positively to Kathy with an exaggerated positive affect, (ii) vocalizing, making repetitive playful sounds, (iii) smiling, (iv) touching, such as stroking her arms or head, placing a finger on her nose or lips and naming these body parts simultaneously, (v) tickling, rubbing and higher intensity tactile stimulation, especially of her hands, (vi) imitating sounds that Kathy made, (vii) playing baby games like pat-a-cake, clapping, keeping time to the singing of a song. Only during these latter activities would there be any attempt to physically restrain or prevent Kathy from producing her hand mannerisms. Kathy's affective/social responding was then used to control the rate and intensity of these interactions, so that if she seemed to be getting upset, resisting, pushing the teacher away or herself withdrawing, the activity would be changed and the intensity reduced.

(2) Long-term influence: acquisition of a simple manual operant. This intervention consisted of teaching Kathy to produce the response of pressing a simple switch in order to turn on music for a short period of time. The manipulandum was a double micro-switch which could be pressed either by a single finger, or by letting her hand drop down on the switch or a number of other simple responses. The device was fully automated, unlike in the previous year; any response on the manipulandum provided 10 seconds of pre-recorded music. Generalization of the response to an appropriate toy was tested by means of the "Big Mouth Singer" game and the electronic flute.

Design: Both intervention programs were conducted each day. In order to investigate their immediate effects on behavior, each constituted the one-to-one instruction during an observation session on an alternating schedule. In order to observe specific effects of operant response acquisition in excess behavior, independent measurement of response acquisition and actual production of appropriate motor responding, including manipulation of toys, was recorded by the teacher. After nearly 2 months (6 observation periods) various aspects of the program were altered in order to meet the needs of the classroom and improve the quality of the instruction. Six more observational data points were then obtained -- this phase is called Treatment 2.

Results

The results of this design were analyzed by means of a Multivariate Analysis of Variance. The main effects were Treatment (Play or Program), Situation (1 or 2) or Phase (Baseline, Treatment Phase 1, Treatment Phase 2). The dependent variables were the observational code categories of interest. Three periods of observation made up the data for each cell of the design.

To summarize the results of this complex analysis, each excess behavior variable was examined to see if there were any significant main effects or interesting interactions. The main effect of Situation was highly significant for 01-rocking ($S_1 > S_2$; $p. < .001$) and 13-finger flicking ($S_1 < S_2$; $p. < .001$). The main effect of Treatment was significant for 13 (Play < Program, $p. < .05$) and for 86-tongue movements (Play > Program, $p. < .01$). The crucial main effect, of course, was for Phase. Here excess behaviors 64-blowing, or 65-vocalizing, showed significant differences; in both cases the percent duration of the responses was higher during intervention:

Behavior	Baseline	Treatment Phase 1	Treatment Phase 2	
64	0.0	5.4	10.1	$p. < .01$
65	3.4	18.1	14.1	$p. < .01$
13	38.0	16.8	22.3	$p. = .09$

Excess behavior 13 has been added as it approached significance and as the Situation by Treatment by Phase interaction was significant ($p. < .05$): The nature of this interaction can be seen in the following table of percent durations:

Table 6.4
Excess Behavior 13: Percent Duration

		Baseline	Treatment Phase 1	Treatment Phase 2
Play	S_1	21.3	11.8	13.2
	S_2	52.3	10.6	24.1
Program	S_1	26.1	4.1	14.2
	S_2	52.0	40.9	37.7

It can be seen from this table (which also summarizes the design of the study) that finger flicking was reduced significantly during the treatment phases with the greatest effect being the reduction during the play-oriented treatment. Note that the reductions in finger flicking in Situation 1 (which always produced less finger flicking that Situation 2) is a demonstration of the generalization of the treatment effect, however the effect of the two treatments on free play (Situation 1) are actually confounded. Similarily it should be noted that the Baseline phase is common to both types of treatment (Play or Program) which explains why the percent durations were quite similar during baseline, except for the situation effect which had not been noted previously to be markedly less during free play.

By looking at difference in teacher behavior the design provides something of a check on teacher "integrity". The interactions for the play treatment condition clearly indicated high levels of smiling, laughing, positive interaction and gentle physical contact. We observed that this particular teacher had some difficulty implementing this style of interaction. None of the teacher behavior categories showed main effects due to the type of treatment although Teacher-neutral did decrease from baseline to the two treatment phases ($p < .05$). This suggests that the effect of the style of interaction on Kathy's behavior was possibly due to the lowered demands during the Play condition rather than to the teacher's affect per se. As to the overall reduction in finger-flicking, it seems doubtful that this was significantly related to the manual operant training in the sense of increased manual responsiveness. Kathy spontaneously depressed the microswitch on three occasions through the four months of daily programming. She also depressed a button on the "Big Mouth Singers" on two occasions and patted at the mobile. She did not pick up the cookie or reach out for it.

Furthermore, although statistically significant reductions in finger flicking were observed, significant increases in blowing or vocalizations were obtained. As there were no increases in cry-no tears or shriek/scream it could be argued that the increase in vocalizations is positive. However it had been frequently noted that blowing seemed be be reciprocally related to Kathy's finger flicking and this was once again seen in these results. Overall, her duration of blowing was less than at the end of the treatment phase of the previous year.

Discussion and Conclusions

As was mentioned at the beginning of this particular case study, the major purpose of this description is to provide an example of the enormously complex design and interpretation problems that we encountered in obtaining multiple response measurements in real time. Clinically the results are somewhat disquieting. After three years of quality school programming with considerable consultation input on our part, Kathy continued to exhibit many excess behaviors and had not acquired the very fundamental skill of voluntary (or at least functional) use of her hands

to manipulate her environment. A simple manual operant was only just beginning to emerge at the end of the second treatment phase. While we cannot help but feel that the technologies to teach such responses are available, Kathy's teacher did not seem to be able to utilize either the technology we provided not did she consistently apply principles of systematic instruction familiar to her. To a large extent, Kathy's limited progress could be attributed to the limited amount and quality of "engaged time" in her program, i.e., time during which she was actively involved in actual instruction. The educational validity evaluation model proposed in Chapter 7 of this report would enable programs to determine more precisely the reasons for intervention outcomes. It may also be that the time at which a child's first operant responses are being established is a crucial one, and that public school programs may need to provide additional temporary instructional personnel during this learning phase just as personnel may be temporarily assigned to deal with behavioral crises.

Without having complete control over the environment, our three observational situations did nevertheless help to provide much information regarding influences on target responses. Striking differences were observed across situations, even in behaviors which seemed highly stereotypic. In Kathy's case, excess behaviors tended to be higher in free play except that some responses, notably blowing and finger flicking, seemed to increase when demands were being placed on her. Highly structured trial by trial programming for a child as severely handicapped as Kathy might contribute less to skill acquisition than would the sort of interactions described under the Play condition. Or, perhaps more cautiously expressed, interventions which themselves seem to increase excess behavior could be modified and adjusted so that excess behaviors are minimized.

Finally this study revealed just how much more work needs to be done on monitoring many excess behaviors. Some behaviors increased or decreased over the three year period with little lawful regularity; new behavior, not originally seen, did emerge, and others which had concerned Kathy's teachers seemed to gradually decrease without formal intervention, such as crying or screaming. Long "baselines" as we were able to gather show only too clearly the fluctuations in behavior which can be masked by the less sensitive recording techniques which are the accepted standards of the field. On the other hand, our efforts to reverse systematic response relationships were less successful than we had hoped for at the begininning of this project.

Child 06: Danny

Introduction

Danny had been included in our comparison sample during the previous (1980-1981) school year, and became an experimental subject during the 1981-1982 year. For both periods, he was enrolled in a self-contained class for SMH children on a public elementary school campus in Waipahu on the island of Oahu. He had two different teachers for the two years he was included in the project. His teacher for the "comparison" year was only currently completing his master's degree in special education but was considered to be an excellent teacher with several years previous experience in a program for severely handicapped children with severe cerebral palsy. He left the Waipahu teaching position to accept a similar position in Honolulu district where his would be the first classroom for SMH teenagers located on a regular secondary education (as opposed to elementary) campus. Danny's teacher during the year in which this study was conducted had just received her M.Ed. in special education under Voeltz's supervision; she had in fact completed her graduate practicum semester in Danny's class during the previous semester, under the supervision of Danny's teacher that year. This teacher was less experienced but was also considered to be excellent, and in our judgment the continuity in Danny's program was maintained.

Danny was seven years old when this study began. He was diagnosed as profoundly retarded, was nonambulatory with suspected motoric (cerebral palsy) involvement, was not toilet trained, and was cortically blind. With regard to his visual impairment, it was our judgment and that of his teachers that he could perceive at least light-dark differences and may have been able to discriminate the presence of persons and large objects. Although his legs and feet were extremely thin and weak (he is in the lowest quartile for size and height for his age), he was able to stand with support for brief periods of time, and he was regularly engaged in physically assisted walking with hip support through strategies developed by and under the supervision of the physical therapist. He appeared to be making good progress in developing these various motor skills, and we felt he would probably be walking within the next few years.

Danny's educational programs emphasized motor development and adaptive behavior. He had no independent feeding, toileting, or dressing skills, for example, and seemed to be unable to engage in a task such as scooping cereal from an adapted bowl into his mouth with his (adapted) spoon. Thus, one of his IEP goals was to increase self-feeding skills; to do this, the program described changes in levels of assistance through fading of full and partial physical prompts. Additional programming included an auditory tracking objective, in which Danny was required to localize a sound source by turning his head in the correct direction, and a toilet regulation program.

Danny was referred to our project almost immediately upon entering the school program because of the extent and severity of his excess behavior. His most serious behavior was self-biting. Unless his hands were physically restrained, he placed one or both (sometimes in a "fist" covered by a portion of his t-shirt) into his mouth where he sometimes merely sucked on them but more typically bit both hands, especially at the base of his thumbs.

Danny's hands were heavily calloused from previous self-biting injuries to the skin tissue, the skin on his hands was broken in various places most of the time, and he nearly always had a severe skin rash on his face, hands, arms and upper chest area; the rash was presumed to be a consequence of the constant presence of saliva and the concomitant irritation to tissue in those areas. When Danny's hands were not in his mouth, he "clapped" them together--in full or half-fists--nearly continuously. In addition, he engaged in almost constant head weaving, eye rubbing, face slapping, teeth-grinding, bouncing up and down on his bottom in a sitting posture, and (as implied above) drooling. He periodically (i.e., every few minutes or even seconds at certain times) emitted clicking and shrieking vocalizations, ruminated, and if objects were present, he would pound his fists or kick his feet against them. Danny's teacher during the 1980-1981 year had purchased cloth "mittens" and arm splints for use whenever Danny was not specifically being programmed so that the open sores on his hands would heal and the rash--which the teacher felt might by now be exacerbating Danny's behavior--would clear.

Danny had been observed by us on many occasions (i.e., by Voeltz) prior to the study in conjunction with practicum student supervision in the school's SMH program, and the only functional use of his hands and/or arms which we had directly observed were use of hands/arms to change body position (e.g., roll from side to prone on the floor) and what appeared to be several instances of "normal" scratching in response to an itch.

A narrative observation recorded by Voeltz of Danny in October 1981, at the end of the baseline phase and immediately prior to the implementation of the first intervention phase, notes the following:

Body control, trunk control improved: Functional scratching of itch on arm; rolls over to prone; supports self on elbows in prone; rubs eyes & flicks hands while maintaining balance on side (lying); slaps face with left hand; waves bell and bracelet.

Clicking episodes: Only one clicking episode during 20 minutes of observation, occurred when teacher picked him up (communicative function?).

Self-biting: Lots of hand biting, face looks better (rash clearing), but hands are very wet (saliva), skin is shriveled, red and appears raw.

Teacher's ecological modifications: D. has his shirt off (as was true throughout baseline and intervention phases) since teacher felt that it was a source of irritation when soaked with saliva, and with it off, teacher can keep his skin dry; teacher frequently cleanses skin with washcloth soaked in warm water and baking soda; foot restraints, tight body control/hand control evident during eating, Danny eating with his right hand; during play program, Danny is seated cross-legged (so he cannot kick his legs).

Behavior during play program: D. exhibits lots of head weaving. He "plays" organ by first pounding the keyboard with his fist/s several times, then holds his hand in place for continued tone.

Programs Related Directly to Study during 1981-1982

Danny's teacher during the 1981-1982 school year continued the previous efforts to improve motor and adaptive behavior, and began a systematic effort to teach Danny more functional use of his hands. Two major programs in this area were a self-feeding and a play program. The teacher felt that much of Danny's excess motor movements were much like the generalized reflexive responses displayed by an infant when actually attempting to reach for, grasp, etc., an object. She felt that these "competing" and perhaps non-voluntary movements may be interfering with Danny's efforts to engage in an operant behavior such as self-feeding with a spoon. Therefore, for the self-feeding program, Danny was seated in his adapted chair with tray-table attached, his legs and feet were secured to the chair with velcro straps, and the teacher prompted him to maintain a grip on an adaptive "post" fastened to his tray-table in position for his left hand. He was learning to scoop food, initially with full physical assistance, from his bowl (secured to the table) with an adapted spoon in his right hand.

A second program implemented to increase functional hand use was a play program with a toy (table top style) organ. Danny was seated on a carpeted area of the floor for this program, and was supposed to "play" keys on the organ placed in front of him in response to the verbal cue "Danny, play toy." If he responded correctly by manipulating the keyboard with fingers or fist, he was rewarded with lavish social reinforcement; this included verbal ("good playing" etc.) and physical touching such as rubbing his arm/s and/or stroking his back. If Danny began to play with the toy when no verbal cue nad been given or began to engage in a negative behavior (such as pounding on the keyboard), the teacher followed such behavior with a brief contingent restraint, holding his hands down to his sides for a few seconds, with her head turned away in an "ignoring" posture.

Design of the Study

In addition to the various behaviors and cabilities noted above for Danny, there was some concern that he was tactually defensive. He became agitated when touched and most of the staff interpreted this as an effort to avoid touch or, at least, as a negative emotional reaction to touch. Thus, the social reinforcement which included physical touch might not actually be reinforcing to Danny. Secondly, the "baseline" play program appeared to discourage voluntary and spontaneous play behavior at the same time that it espoused to teach him independent play skills: If the play was in any way designed to replace self-stimulatory responses or the self-biting, and/or if Danny was going to learn to initiate play in the way a nonhandicapped child might, it seemed important to change the program. In particular, Danny was being trained to touch the organ only upon the teacher's verbal cue to play and, in fact, was being punished if he did so spontaneously. A more appropriate and functional play program would undoubtedly be one in which he was reinforced for spontaneous, appropriate play and not only for prompted play. Hence, Danny's intervention study was designed to investigate whether particular play training and reinforcement strategies would result in changes in his excess behavior.

Phase A: Baseline

Three observation sessions[2] were coded and analyzed for each of two, two-week periods, which will be referred to as Period 1 and Period 2 in certain representations of the data. This enabled us to conduct two factor analyses of the six sequential observation sessions (based upon three consecutive sessions each) across the four weeks of "baseline." During this phase, Danny was observed during one-to-one instruction in playing with the toy organ according to the more traditional massed trial practice model described in the previous section. In the free play situation, he was provided with the usual array of toys while seated on the floor mat. The group situation varied, but generally included only one other child.

Phase B1: Physical Touch + Social (Verbal) Reinforcement

Three observation sessions across a two-week period were coded during this phase, comprising one Period (3). The play program with the toy organ was modified as follows: presentation of the organ was the (natural) cue to play, and the teacher provided physical and/or verbal (instructional) prompts only when needed, i.e., when Danny did not respond to the natural cue. He was contingently reinforced for his organ play as during Phase A, with both physical touch (rubbing his arms, hugging him, etc) and lavish social praise (verbal "good playing," etc.). Also as in Phase A, he was seated on the teacher's lap on the floor mat, with his back in front of the teacher's midline as she was seated directly behind him. Thus, the one-to-one instructional situation observed during Phase B1 differed from Phase A in terms of cue presentation and the absence of a correction procedure contingent upon play prior to a verbal cue by the teacher, but was otherwise virtually identical, particularly with regard to the reinforcement contingencies in effect. During free play, Danny was observed with the usual array of toys but including the organ. The group situation was a play session with the organ involving one additional child.

Phase C: Social (Verbal) Reinforcement Only

Three observation sessions were conducted across the next two-week period, referred to as Period 4. One-to-one instructional sessions were identical to those described for Phase B, with one exception: only contingent social (verbal) praise was delivered following efforts to play with the organ. No physical touch was to occur as reinforcement. Although some touch did occur, this was kept to a minimum by the weekly procedural reliability monitoring of teacher behavior by Voeltz, who observed sessions in each phase and counted these behaviors, providing feedback to the teacher regarding their occurrence after each session.

---

[2]Throughout the phases of the investigation, training sessions (essentially identical to and including those observed by us) were conducted at least once daily. However, only a sample of those sessions were observed during each phase, e.g., only three out of at least ten training sessions during Phase B were observed at regularly spaced intervals across the two week period.

Phase B_2: Physical Touch + Social (Verbal) Reinforcement

During the final phase, five observation sessions were conducted across the subsequent month time period, to allow for an additional factor analysis replication within this phase based upon three (Period 5) and two (Period 6) observation session data points respectively. This phase was identical in all respects to Phase B_1 as described above.

Data Analysis

Three procedures were utilized to analyze the results of this investigation. First, the observation code categories for the particular excess behaviors of concern were monitored across phases as dependent variables. This was done according to the traditional procedures in which changes in mean percent duration or frequency occurrence of excess behaviors would be analyzed as a function of the treatment phases in the within-subject design. We also conducted a factor analysis of Danny's excess behaviors within each time "period" for the six time periods noted above (Periods 1 and 2 within Phase A, Period 3 within Phase B_1, Period 4 within Phase C, and Periods 5 and 6 within Phase B_2). These factor analyses were conducted in order to determine whether the intervention phases were related to changes in the structure of Danny's excess behavior interrelationships. The factor pattern derived for Period 2 of the Baseline Phase A was used as the frame of reference for all further comparisons of factors and factor scores across phases. Finally, since the factor analysis revealed some consistencies in that three factors appeared similar in structure in Situation 2 (one-to-one instruction) across phases of the study, factor scores were computed for three sets of behaviors which co-occurred:

Factor 1 = Drool/Saliva Swish (73) + Tongue Movement (86)

Factor 2 = Self-biting (32) + Hand Clap (10) + Subglottal Vocalization (65) + Jump/Hop (91)

Factor 4 = Head Weave (78) + Head Drop (79) + Object Pound (89)

A unit weighting procedure was used in which the factor score derived for each interval of an observation session would be based upon the number of behaviors within that factor which occurred during that interval, thus yielding an average number occurring across each phase for that factor as the dependent variable (each factor score mean). For each factor, the factor score ranged from a minimum of zero (for all factors) to a maximum of 2 for Factor 1, 4 for Factor 2, and 3 for Factor 4, respectively.

Results and Discussion

Table 6.5 (see pages 118 and 119) provides a display of Danny's excess behavior in the three observation situations for the four phases of the investigation. In general, Danny's excess behaviors during the one-to-one instructional situation showed changes across time which were only partially related to the intervention phases of the study. Two change patterns appear most common: (1) a sharp decrease following baseline, with the excess behavior maintaining at this lower (or zero) level

TABLE 6.5

CHANGES IN DANNY'S BEHAVIORS ACROSS SITUATIONS AND INTERVENTION PHASES

		Percent Duration Behavior Across Phases			
Behavior	Situation	Condition A Baseline $\bar{X}$	Condition B_1 Social and Physical $\bar{X}$	Condition C Social $\bar{X}$	Condition B_2 Social and Physical $\bar{X}$
Self-Biting (32)	Free Play	4.2	5.1	0	0
	One-One	3.15	.5	0	0
	Group	2.1	2.4	0	0
Hand Clap (10)	Free Play	13.5	15.9	16.9	9.0
	One-One	5.9	11.3	15.8	16.1
	Group	5.4	11.9	14.1	9.2
Object/Surface Pound (89)	Free Play	4.3	11.2	9.1	17.0
	One-One	19.1	5.0	6.7	5.2
	Group	14.0	4.2	5.1	4.3
Head Weave (78)	Free Play	14.7	12.9	11.1	10.5
	One-One	18.3	20.1	14.6	10.7
	Group	14.2	14.4	8.4	10.1
Head Drop (79)	Free Play	1.5	4.9	6.0	7.5
	One-One	6.4	12.1	17.0	12.2
	Group	5.5	13.9	11.0	6.8
Mouth Body Parts (25)	Free Play	20.9	23.7	30.9	23.5
	One-One	5.8	2.4	.8	1.3
	Group	8.2	5.0	11.0	2.5
Jump/Hop (91)	Free Play	2.6	.4	3.7	0
	One-One	5.4	1.8	6.9	.3
	Group	3.3	.8	1.5	.7

TABLE 6.5 (cont.)

		Condition A Baseline $\bar{X}$	Condition B_1 Social and Physical $\bar{X}$	Condition C Social $\bar{X}$	Condition B_2 Social and Physical $\bar{X}$
Sub	Free Play	6.0	2.6	1.0	1.3
Glottal Voc	One-One	6.2	2.5	.1	.8
(65)	Group	2.5	3.2	0	.2
Saliva	Free Play	2.9	2.2	6.0	6.24
Swish	One-One	11.2	6.9	7.3	6.2
(73)	Group	9.0	9.0	4.9	2.9
Tongue	Free Play	20.1	7.5	11.2	15.0
Movements	One-One	30.2	27.8	21.9	23.3
(86)	Group	25.3	24.4	16.6	25.4
Shriek/	Free Play	2.3	8.0	11.5	6.4
Scream	One-One	2.1	2.9	2.3	.2
(62)	Group	2.2	2.9	7.7	.2
Click	Free Play	1.7	.1	0	0
	One-One	4.2	0	0	0
(67)	Group	2.4	.1	0	.1
Teeth	Free Play	0	20.1	12.5	7.1
Grind	One-One	.9	34.8	31.6	21.7
(75)	Group	0	41.8	17.0	13.1

across intervention phases B_1, C, and B_2. This occurred for Self-Biting (which virtually ended after baseline), Object/Surface Pounding, Mouthing Body Parts, Drooling/Saliva Swishing, and Clicking; and (2) a progressive decrease across all phases, with the behavior highest during Phase A and/or Phase B_1 and lowest during the final intervention phase/s. This occurred for Head Weaving, Subglottal Vocalization, Tongue Movement, and Shriek/Scream (although this last behavior was steady across Phases A, B_1, and C, declining to near-zero during the last phase). Two behaviors increased rather sharply following baseline, and remained at that (higher) level across intervention phases: Hand Clapping and Teeth Grinding. Teeth Grinding had been at near-zero level during baseline (.9%), rose to 34.8% during B_1, and then declined across Phases C (31.6%) and B_2 (21.7%). These increases in both Teeth Grinding and Hand Clapping appear to be side effects associated with same-time decreases in other behaviors which are topographically similar (e.g., Self-Biting and Object/Surface Pounding). Only two behaviors showed changes in percentage occurrence which appear to be related to phases of the study: Jumping/Hopping decreased from 5.4% during baseline to 1.8% and .3% during the two Social (Verbal) + Physical reinforcement conditions, with a temporary return to a higher than baseline 6.9% duration during the intervening Phase C (Social/Verbal reinforcement only). Head Dropping was lowest during baseline (6.4%), rose during B_1 (12.1%), increased further during Phase C (to 17%) and returned to 12.2% during B_2. It would appear, then, that Danny's behavior showed an overall improvement corresponding to the implementation of the intervention phases in contrast to baseline, with some evidence that the Social/Verbal + Physical (Touch) reinforcement condition was associated with more improvement in comparison to the Social/Verbal reinforcement only condition. Unfortunately, without a return to the baseline conditions (the traditional toy-play program) we cannot attribute the improvements in behaviors such as Self-Biting to this intervention alone. As discussed earlier in Danny's data, his teacher had also implemented other procedures (i.e., a training program to improve self-feeding) which might be related to this improvement. We were, however, unwilling to risk a return to baseline with this severe behavior problem given the clear and apparently lasting improvement present throughout the intervention phases.

Self-Biting showed a decrease to zero by Phase C for the free play and group situations as well. Mouthing Body Parts did not show a similar improvement during free play and group; in fact, this behavior showed a significant increase during Phase C (to 30.9%) in comparison to Phases B_1 (23.7%) and B_2 (23.5%) and particularly Phase A (20.9%) in free play. This behavior was also higher than baseline during Phase C in group, though it occurred considerably less during Phase B_1 and showed a significant decline during B_2. Since the teacher was not present during free play, the percentage duration changes here could not be attributed simply to topographical interference. It does seem likely, however, that teacher physical (touch) reinforcement may have topographically interfered with Mouthing Body Parts during the group situation.

Interrelationships among Danny's excess behaviors were identified by the factor analyses conducted for each time period; Table 6.6 displays the results of the factor analyses of excess behavior during one-to-one instruction for Period 2 of Baseline and Period 3, the B_1 Intervention

TABLE 6.6

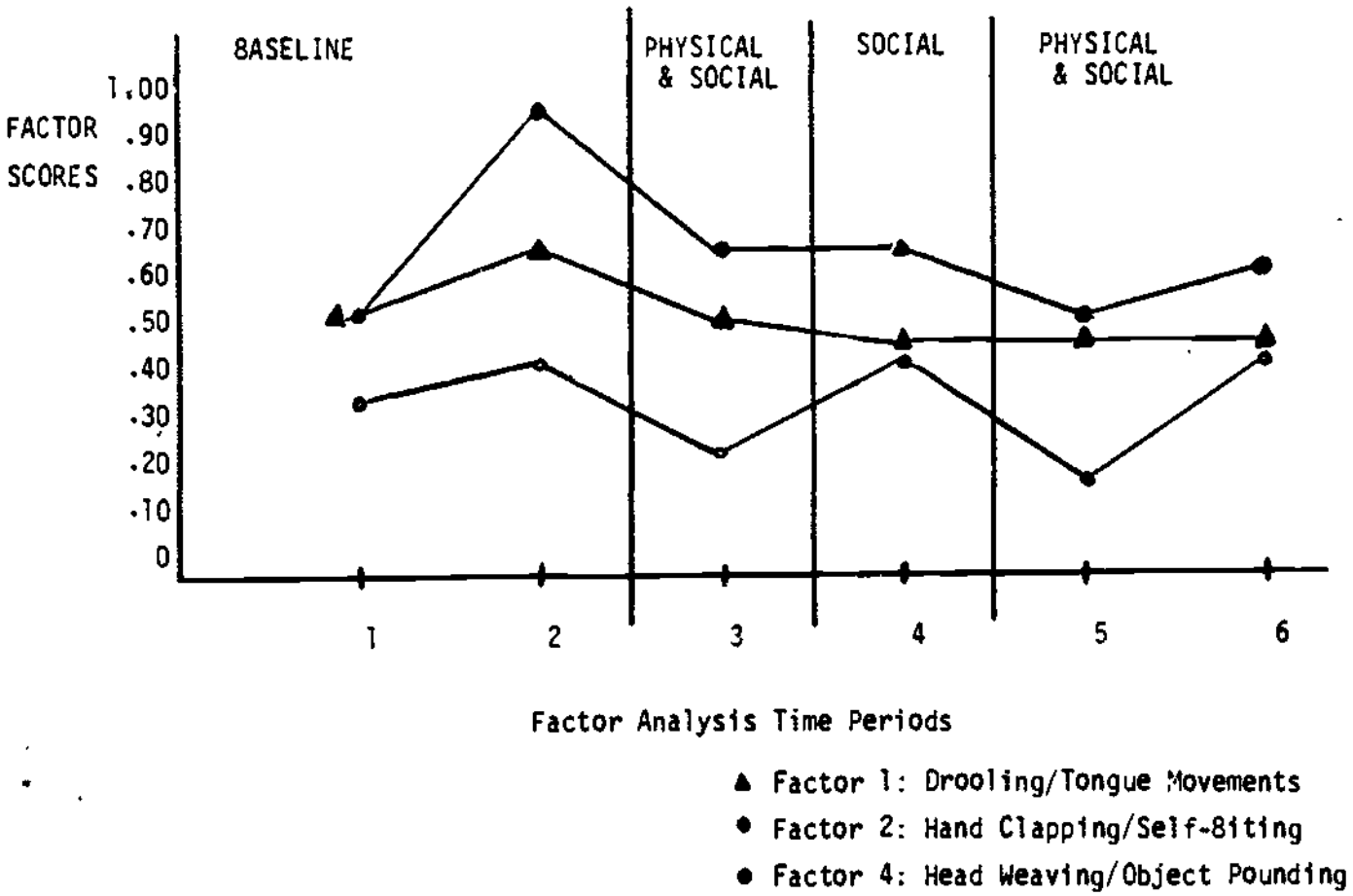
FACTOR ANALYSES OF DANNY'S EXCESS BEHAVIORS ACROSS INTERVENTION

Phases in Situation 2 (One-to-One Instruction)

Period 2: Baseline				Period 3: Social & Physical S^R			
Factor	Behavior/s	Loading	% Duration	Factor	Behavior/s	Loading	% Duration
1	Saliva Swishing	.89	12.5	2	Saliva Swishing	.58	6.9
	Tongue Movement	.52	35.5		Tongue Movement	.76	27.8
2	Hand Clapping	.37	7.0	3	Hand Clapping	.36	11.3
	Self-Biting	.47	6.3	[-	Self-Biting	n/a	.5]
	Sub. Glottal Voc.	.38	4.9	[-	Sub. Glottal Voc.	n/a	2.5]
	Jumping/Hopping	.58	6.4		Jumping/Hopping	.70	1.8
3	Mouth Body Parts	.90	4.5	-	Mouth Body Part..	n/a	2.4
4	Head Weave	.67	17.9	1	Head Weave	n/a	2.4
	Head Drop	.70	10.2		Head Drop	.50	12.1
	Object/Surface Pound	.55	24.9		Object/Surface Pound	.27	5.0
[	Teeth Grinding	n/a	.9]		Teeth Grinding	.68	34.8
5	Shriek/Scream	.43	3.0		Shriek/Scream	n/a	2.9
	Click Vocalization	.36	6.4		Click Vocalization	n/a	0
	Object/Surface Pounding	.38	24.9	[			]
	Jumping/Hopping	.37	6.4	[			]

(Social + Physical/Touch Reinforcement) Phase. Even though certain individual excess behaviors no longer occurred following baseline, the set of remaining behaviors which had appeared during the baseline factor analysis continued to show an interrelationship across subsequent phases. In particular, Factors 1, 2 and 4 could be identified throughout the intervention (see Table 6.6 and Figure 6.1). However, within these factors, certain behaviors had ceased to occur (Self-Biting in Factor 2) or were occurring at extremely low rates and failing to load significantly on the factor (Subglottal Vocalization in Factor 2 and Head Weaving in Factor 4) following baseline. Figure 6.1 displays the changes in these three factor scores across the phases of the study; as is explained on page 117, each graphed data point reflects the average unit-weighted occurrence of that factor within treatment phases and is affected by whether or not any of the behaviors occurred within a MORE "interval" as well as how many of those behaviors within that cluster co-occurred. Thus, even though Self-Biting scarcely occurred by Period 3, the increased occurrence of Hand Clapping results in an overall factor score of the same approximate level by the last treatment phase

FIGURE 6.1
Factor Scores for Factors 1, 2 and 4 During One-to-One Instruction with Danny across Phases.



in comparison to baseline. Similarly, while Object/Surface Pounding within Factor 4 has declined significantly during the intervention phases following aseline, Teeth Grinding increased significantly, loads on this factor, and thus results in a mean factor score remaining at the same level across phases.

Thus, while the intervention phase seems to be associated with the deletion of cerrain behaviors from a particular factor, the factor itself can still be identified and the remaining behavior/s within that factor are occurring more often such that the mean factor score does not decline significantly. These results support the hypothesis that when one behavior within a cluster of interrelated behaviors is decreased or extinguished, the remaining behaviors within that cluster increase in frequency or a new behavior appears to "take its place." In Danny's case, the decrease in his self-injurious self-biting would undoubtedly justify an increase in the other negative behaviors which occurred. Yet, these other excess behaviors (i.e., teeth grinding and, to a lesser extent, hand clapping) could potentially increase in frequency and intensity so that they would be judged to be equally serious. What must occur instead is a concomitant increase in positive behaviors (such as appropriate play during free time). Danny's behavior appeared to be improving in this direction so that over time, we would expect this alternative pattern to become established in his repertoire. But one should not expect an overall decrease in excess behavior following successful modification of a single target excess behavior in a child whose repertoire consists of primarily excess behaviors. Clearly, alternative behaviors will occur within any given "time-space" and, in the absence of any skills to display positive behavior, existing negative behaviors will perform this function.

Child 05: Paul

Introduction

Paul was enrolled in the same self-contained class for severely mentally retarded (SMR) children throughout the two years of his intervention study, 1980-1981 and 1981-1982. The classroom was located on a public elementary school campus in Waipahu on the island of Oahu, and was part of a "pod" of four special education classrooms arranged around an open air court. Paul's teacher during these two years was a student enrolled in the University of Hawaii's master's degree program in special education, which he attended part time while teaching full time. This teacher was the lead teacher of the special education classes for the second year of the study, and was oriented toward structure, behavioral control, was data-based, and generally organized all instruction around one-to-one, discrete trial sessions in the traditional behavioral format. However, he was enthusiastically interested in the skill cluster methods reflected in this intervention study. Once the intervention study with Paul was begun, this teacher proceeded to modify programs for all his students in the classroom based upon the skill cluster concept.

Same-age regular education peers also attended this particular school, with their classes being similarly grouped (generally by grade level) into physical "pods" of several rooms. In addition to Paul's class, the special education pod included two classes for severely multiply handicapped (SMH) children (child 06 was enrolled in one of these with another teacher) and one class for moderately mentally retarded children. Throughout the study, Paul participated in recess interactions with fourth grade regular education "Special Friends"; this particular school had conducted this program since 1978.

Prior to 1980, Paul attended a public school on the windward side of the island near his parent's home. He was one of the students enrolled in Hawaii's first class for severely multiply handicapped (SMH) children on a regular education elementary school campus. This class began in 1977, was taught (including Paul) by a master teacher who has continued to play a leadership role in Hawaii in innovative program design for severely handicapped children in Hawaii, and thus was regularly visited by teachers, parents, and administrators from throughout the state as each district moved to establish such integrated public school services for these children from 1977 on. Paul changed schools to the Waipahu site in 1980 when his parents placed him in the state institution for mentally retarded persons, since the Waipahu school was close to this facility in another school district. Generally speaking, then, Paul had been exposed to a high quality public education for severely handicapped children since school age.

At the completion of the study (the study ended, in fact, earlier than anticipated due to this move), Paul's school placement changed a third time. Due to placement out of the institution into a foster home, Paul was transferred to another school on another part of the island; this was also a public school classroom for severely handicapped children on a regular elementary school campus. We thought it unfortunate that this subsequent placement was seemingly unable to implement the intervention described in the next few pages. Paul's teacher at this new school was trained in the Teaching Research model and the classroom staff implemented physical restraint and contingency management strategies to manage his excess behaviors. In general, we saw a decline in Paul's behavioral performance in this new setting, supported by two systematic observations sessions conducted by us in the new school after the move.

Since this decline in behavior could be due to either the move or to the change in intervention procedures, we would need a number of observation sessions at least equal to our intervention phase B in order to investigate this issue. Unfortunately, it was not possible for us to schedule our observers so late in year three to travel to an additional school in another region of the island.

Diagnosis and Program

Paul was nine years old when the baseline phase of this study began (in November 1980) and eleven years old when the intervention phase data collection ended (in February 1982). He was diagnosed as profoundly mentally retarded according to a Cattell score of 9 attained in testing at the age of eight years, though his records referred to him as severely mentally retarded with an IQ of 34. At the age of eight years, eleven months, Paul attained a score of 88 (out of 194) on the TARC (Sailor & Mix, 1975). The etiology of his retardation was unknown, with the term "phychomotor retardation enigma" appearing throughout his records dating to infance. Paul had a severe myoclonic seizure disorder which was only partially controlled through medication throughout his life despite sophisticated medical care provided by a pediatric neurologist with an outstanding reputation in the area of developmental disabilities. According to his records, the first three years of Paul's life were characterized by almost continuous severe seizure activity. What had been surprising to htis neurologist and other specialized personnel was that the extensive seizure activity was associated with little motor involvement. While extensive damage ordinarily occurs following such a seizure history, Paul's gait and motor coordination were quite good for his age. In fact, most of Paul's excess behavior involved coordinated gross/fine motor activity such as running, climbing, and various hand/finger movements (e.g., spinning objects). Paul's appearance is normal, and he could be described as an extremely attractive child. His height and weight are within the normal range for his chronological age.

Paul experienced a grand mal seizure at schocl on the average of once a week, with perhaps one more such seizure occurring in the living environment. Following these seizures, he was physically exhausted and required sleep for anywhere from one to two hours. He seemed generally unable to function adaptively after a seizure, so that seizure activity at home in the morning affected his school performance as well. At other times, Paul's appearance was quite alert and he exhibited bursts of hyperactivity. Paul was toilet regulated but had occasional accidents, and has a longstanding history of problems in this area including extensive smearing of feces. The smearing was under control throughout this study, but had been a major factor in the parents' decision to place Paul in the state institution when he was eight years old. Paul also had a history of stripping (removing his clothing), another problem which was under control during the study--perhaps through interruption of any first effort to remove an article of clothing. Paul generally began the behavior by removing a shoe, and if interrupted at that point, would cease the behavior.

Paul's excess behavior repertoire might best be described as autistic-like. The following behaviors were observed by us and reported in the teacher and caregiver excess behavior checklist as most typical:

05 Jerky movement
06 Body limpness (apparently as a "task avoidance" strategy)
25 Mouthing body parts
26 Mouthing objects
28 Head slapping
41 Grabbing objects
55 Object spinning
56 Object dropping
60 Cry, no tears
62 Shriek/scream
70 Facial grimace
72 Staring/gazing
88 Bolting (apparently as a "task avoidance" strategy, for teacher attention, and/or to obtain preferred reinforcer)
94 Climbing (on furniture, etc.)

Table 6.7 summarizes Paul's IEP objectives and his current performance level during baseline, immediately prior to the intervention phase (October 1981).

Design of the Study

Most of Paul's excess behaviors appeared to serve several functions; task avoidance during instruction (e.g., running away, going "limp", screaming) was believed to be a major function of many of his behaviors. Also during instruction and at other times as well, he frequently attempted to obtain his favority toy (a small plastic ring) which he would manipulate repetitively in various positions for as long as he was allowed to play with this object. On occasion, he also appeared to display certain behaviors (e.g., "bolting" from a free play setting and at transition times) for teacher attention. Paul seemed unable to remain seated for his instructional sessions, virtually all of which occurred in the traditional discrete trial one-to-one format (as noted above) in which he was expected to remain seated and respond to a series of teacher mands and reinforcements. He reaction to this instruction and to confinement to any area during free play time was to be extremely disruptive. The "baseline" period reported here included a traditional contingency amangement program designed to increase appropriate instructional and free play behavior by allowing him contingent access to his preferred "self-stim" toy. Our data showed little change in Paul's behavior associated with this intervention.

The study reported here is an experimental investigation of the effects of a major change in instructional delivery pattern upon behavior. Guess and his colleagues and students at the University of Kansas have developed the Individualized Curriculum Sequencing (ICS) model, in which target responses are "reorganized" for instruction into more natural skill clusters, as they would actually occur in the natural environment and situation (cf. Holvoet, Guess, Mulligan, & Brown, 1980). By employing a distributive trial format (rather than massed trial) in which each individual skill is taught as part of behavioral chains, the ICS model is intended to facilitate acquisition, maintenance and generalization of the individual skill components. However, there is to-date no published, experimentally controlled evaluation of the effectiveness of this instructional strategy upon child outcome, e.g., as compared to a more traditional discrete trial format in which each isolated

TABLE 6.7 Paul's IEP
and Program during Baseline
(Review Date: 11/81)

	Instructional Goal/Objective	Paul's Performance
1.	Wash hands	Won't turn water on, though he can turn water off.
2.	Follow commands which are signed (ASL) and vocalized (receptive language)	Can follow: stand up, sit down, come here. Does not follow: Pick it up, put it down, hands down, go to the toilet, go to the table, go to the door, get your (school) bag.
3.	Travel from bus to classroom independently	Can do, but fails to bring bag into classroom from bus; at end of school day, does not go from classroom to bus (runs to bathroom instead).
4.	When presented with the object and/or picture, spontaneously sign "drinking fountain"	In progress.
5.	Correct use of napkin at snack and lunch times	OK, program ended.
6.	Imitations (gross motor)	"Some progress" but program dropped.
7.	Recognize own name printed on card	No progress, including program modification adding extra-stimulus prompt (a color code).

skill is instructed through massed trials (Mulligan, personal communication). We were particularly interested in whether the ICS skill cluster model might be more motivating to a student like Paul, who appeared to be so noncompliant in the traditional behavioral model. Specifically, would Paul's excess behavior and/or instructional performance change (i.e., improve) as a function of adopting this currently recommended "educational best practice" in program design?

In order to investigate this issue, an interrupted time series design was employed in which behavior during an extended "baseline" (Phase A)--his current educational program--is compared statistically to behavior during a subsequent intervention phase (Phase B)--a similar IEP but with instruction occurring in the skill cluster format. Paul's IEP review data of October 30, 1981, marked the end of Phase A and the beginning of Phase B. The experimental phases are briefly summarized below:

Phase A: "Baseline"

Thirty-one (31) sessions were observed across a period of three semesters of educational programming in a traditional, behavioral-developmental approach. Goals and objectives had been selected by teacher, parents and other members of the instructional team based upon the identification of "next-appearing" items appearing in sequence after those he displayed on various developmental measures. These goals and objectives include targeted skills in the areas of language, gross and fine motor, and self-help, and Paul's specific performance at the end of Phase A is displayed in Table

Phase B: Skill Cluster Intervention

Sixteen (16) sessions were observed during the period from November 1981 through February 1982 when the skill cluster intervention program was in effect. Each of Paul's IEP goals were translated into instructional objectives intended to be highly functional for Paul, i.e., they were designed to result in natural reinforcers for Paul in his environment. An instructional session for Paul under this new program design condition did not consist of a series of trials in which Paul was exposed to massed trials for a single objective from among those on his IEP. Instead, each instructional session represented an opportunity to practice a skill cluster, consisting of from two to six individual target behaviors appearing in a natural behavioral chain, i.e., as they might occur in the natural environment. Each of these individual targets representing a behavioral objective on his IEP was, in turn, part of at least two such skill clusters; again, as one might expect in the natural environment, a single language behavior would be practiced in more than one chain of behavior during the day.

Figure 6.2 presents the actual program summation sheet which was utilized for the one-to-one instructional observation coded by the observers throughout Phase B. This program summation sheet was generated by our project as a convenient format to assist the teacher during an instructional session. As can be seen from the sheet, certain discrete target behaviors might be "pulled out" for massed trial practice, as would be appropriate whenever additional practice might be needed. Hence, for this particular program, several individual targets were practiced separately from the observation session in massed trials, e.g., turning on the faucet. However, the observers coded

FIGURE 6.2 Paul's Skill Cluster for Phase B Observations.

Domains: Domestic & vocational (self-help, cognitive, language)

PROGRAM SUMMATION

Child P ____ Skill Cluster A

Teacher/s P , H , EA

Setting snack corner of room, near sink, at table

Child's position at beginning of session: seated at snack table

KEY: Cue/Correction Phases
(1) Initial Intrusive
(2) + Fading process

Response Objective
mass trial practice

Materials cup with name, dish drainer, snack (cookie), small toy reinforcer, cloth to wipe table.

page 1 of 3 pages

	NATURAL CUE	INSTRUCTIONAL CUE	PROMPTS/ CORRECTIONS	RESPONSE OBJECTIVE	REINFORCEMENT NATURAL	REINFORCEMENT INSTRUCTIONAL
GOAL: LANGUAGE (signing)	Cookies on table, snack time plus verbal cue: "[name]", [wait for attention] what do you want?"	(1) Teacher stands by drainer, holding cup in view 6 inches above surface, while asking P what he wants (2) As above, but only pointing to cup (3) As above, but T increases distance from cup (toward table) on successive days (4) T sits at table next to P, pointing toward drainer	(1) T prompt for sign "drink" (2) T models sign "drink"	P signs "drink"	Delayed primary drink (& cookies)	Social: "Good, you want a drink!"
GOAL: COGNITIVE (2-step command)	"P , get your cup and get a drink of water for snack."	(1) Repeat cues one at a time, (while prompting response)	(1a) T physical prompt to drainer, back to seat (1b) T models taking cup from drainer, gets water from faucet	P goes to drainer and gets cup, turns on faucet and gets water in cup. Returns to seat.	Primary: drink of water (after returning to chair)	

Domains: Domestic & vocational (self-help, cognitive, language)

PROGRAM SUMMATION

Child P

Skill Cluster A

Teacher/s P , H , EA

Setting snack corner of room, near sink, at table

Child's position at beginning of session: seated at snack table

KEY: Cue/Correction Phases
(1) Initial Intrusive
(2) + Fading process

Response Objective
mass trial practice

Materials cup with name, dish drainer, snack (cookie), small toy reinforcer, cloth to wipe table.

page 2 of 3 pages

GOAL: LANGUAGE (stanina)

GOAL: SELF-HELP

GOAL: DOMESTIC/SELF-HELP

NATURAL CUE	INSTRUCTIONAL CUE	PROMPTS/ CORRECTIONS	RESPONSE OBJECTIVE	REINFORCEMENT NATURAL	REINFORCEMENT INSTRUCTIONAL
Teacher holding P 's cup with drink.	Verbal cue: "P , what do you want?"	(1) T prompt for sign "drink" (2) T models sign "drink"	P signs "drink"	Primary: drink (& cookies)	Social: smiles, miscellaneous talk
P finishes drink & cup is empty--snack time continuing NOTE: REPEAT VARIATION OF LANGUAGE/COGNITIVE PROGRAM AS NEEDED DURING SNACK TIME	"P , do you want more drink? Get more water."	As above	P gets more water from faucet	As above	As above
P finishes snack, snack time over, verbal cue: "All pau; put your cup away & clean up."	Repeat detailed cues & provide detailed cues (e.g., take cup to sink"), one at a time (while prompting response)	T physical prompt	P takes cup to sink, rinses cup, puts in dish drainer (sink?)	Delayed preferred activity: ring toy	Social: "Good job, P !"

Domains: Domestic & vocational (self-help, cognitive, language)

PROGRAM SUMMATION

Child P

Skill Cluster A

Teacher/s P , H , EA

Setting snack corner of room, near sink, at table

Child's position at beginning of session: seated at snack table

KEY: Cue/Correction Phases
(1) Initial Intrusive
(2) + Fading process

Response Objective
mass trial practice

Materials cup with name, dish drainer, snack (cookie), small toy reinforcer, cloth to wipe table.

GOAL: DOMESTIC/SELF-HELP

NATURAL CUE	INSTRUCTIONAL CUE	PROMPTS/ CORRECTIONS	RESPONSE OBJECTIVE	REINFORCEMENT NATURAL	REINFORCEMENT INSTRUCTIONAL
Verbal cue: "P , clean up table."	(1) Extra verbal "Get cloth" & physical prompts (2) Extra verbal, plus gestural cues	T physical prompts	P gets cloth, turns on water & wets cloth, wrings out cloth, cleans table, returns cloth to sink, rinses, etc. Partial: cloth already ready, P gets & cleans table, returns cloth to sink (T rinses).	Preferred activity: ring toy for ___ minutes. Paired social: "Good job."	

GOAL:

GOAL:

the entire skill cluster instructional session, not massed trial practice, during the intervention Phase B. The program summation sheet also makes explicit those cues, correction procedures and consequences which are instructional vs. those which are considered "natural," i.e., instruction would be completed when the student can display the various target responses in the context of natural cues, correction procedures and consequences (not in the absence of any cues, etc.).

As has been discussed previously, observations occurred under three conditions (during both the baseline and intervention phases): Situation 1 (Free Play), Situation 2 (One-to-One Instruction), and Situation 3 (Group Instruction). During baseline, these observations had occurred during any of Paul's various instructional programs being conducted. During the intervention Phase B, Paul was always observed in the One-to-One instructional condition involving skill cluster A as depicted in the Figure. The group and free play conditions remained similar throughout baseline and intervention.

Analysis and Results

During Phase A, the data for 1 Observer 1 session and 4 Observer 2 sessions were lost due to observer error, MORE failures, etc., leaving a total of up to 26 sessions for which both Observer 1 and Observer 2 data were available. During Phase B, data loss reduced the total of available sessions to up to 14 sessions for which we had both Observer 1 and Observer 2 data.

The SAS 79.6 Interrupted Time Series sub-program was used to analyze baseline vs. intervention phase behavioral changes. Only data analyzed for situations 1 (free play) and 2 (one-to-one instruction) are reported here. For each situation, t-tests were conducted comparing mean percentage durations of both teacher and student behavior for selected codes in the two observation systems (Observer 1 and 2) for the Phase A (baseline) vs. Phase B (Skill Cluster Intervention) periods. For all but the teacher behavior codes and Paul's position codes, the SAS test version Auto Regressive Integrated Moving Averages (ARIMA) was also utilized to compute t-tests of baseline vs. intervention means which were adjusted for auto correlation. For these tests, a two-tailed confidence interval of .95 was used.

Tables 6.8 through 6.13present the mean percent durations, standard deviations, and the results of the non-independent t-test comparisons of these means during the baseline and intervention phases under the free play and one-to-one instructional situations. Beginning with Tables 6.8 and 6.9 displaying Paul's general Position, Affect and Task Relatedness, we can see that the free play condition was not associated with any significant differences in his Position, and differences in Affect and Task-relatedness were not significant according to the additional ARIMA tests conducted. In contrast to the one-to-one instructional condition (see below), however, three of the task-related behavior categories showed a significant auto-correlation during free play: "Neutral," "Excess" and "Excess + On-Task."

During the one-to-one instructional session observations (see Table6.9), the skill cluster intervention phase was clearly associated with higher percentages of standing and Walking, and less Sitting. Paul's affect was Neutral

TABLE 6.8 PAUL: Position, affect and task-relatedness in Situation 1 (Free Play).

Comparison of mean percent duration of behaviors across Baseline (Phase A) and Intervention (Phase B) while Paul is engaged in Free Play.

Behavior Category	Phase A (n = 26)[a] $\bar{x}$	SD	Phase B (n = 14)[b] $\bar{x}$	SD	t[c]	$p<$
Position						
0 Lying	16.9	.29	21.3	.28	-.46	ns
1 Sitting	52.4	.24	41.0	.20	1.54	ns
3 Standing	13.5	.13	12.9	.11	.17	ns
5 Walking	12.1	.08	14.3	.12	-.69	ns
6 Running	.6	.02	.3	.01	.78	ns
Affect						
0 Neutral	95.8	.07	99.8	.00	-3.08	.005
1 Happy	00.5	.01	00.2	.00	1.01	ns
2 Sad/Distress	03.7	.07	0	0	2.11	.05
3 Angry	0	0	0	0	—	—
Task-Related						
0 Neutral	12.5	.20	24.5	.24	-1.68	.10
1 Excess	68.4	.29	69.5	.25	-.12	ns
2 Excess & On-Task	14.0	.21	2.1	.02	2.84	.01
3 On-Task Appropriate	3.1	.05	2.6	.04	.25	ns
4 Off-Task Appropriate	2.0	.04	1.2	.05	.55	ns

[a]For Affect, n = 29

[b]For Affect, n = 15

[c]Non-independent

TABLE 6.9 PAUL: Position, affect and task-relatedness in Situation 2 (One-to-One Instruction).

Comparison of mean percent duration of behaviors across Baseline (Phase A) and Intervention (Phase B) while Paul is engaged in instruction.

Behavior Category	Phase A (n = 26)[a] $\bar{X}$	SD	Phase B (n = 14)[b] $\bar{X}$	SD	t[c]	p
Position						
0 Lying	0	0	7.1	.27	-.99	ns
1 Sitting	91.1	.18	67.8	.23	3.57	.001
3 Standing	6.5	.18	15.8	.10	-2.09	.05
5 Walking	2.0	.03	9.2	.07	-3.52	.01
6 Running	.2	.01	0	0	.98	ns
Affect						
0 Neutral	94.2	.10	1.00	.00	-3.13	.01
1 Happy	2.9	.05	0	0	1.5	ns
2 Sad/Distress	3.5	.09	0	0	2.14	.05
3 Angry	.4	.02	0	0	.85	ns
Task-Related						
0 Neutral	4.6	.06	10.8	.26	-.88	ns
1 Excess	16.7	12.5	2.8	.04	5.25	.0001
2 Excess & On-Task	45.9	.35	22.2	.14	2.99	.005
3 On-Task Appropriate	32.8	.30	64.3	.22	-3.45	.001
4 Off-Task Appropriate	0	0	0	0	—	—

[a] For Affect, n = 29

[b] For Affect, n = 15

[c] Non-Independent

significantly more of the time (to 100%) and Sad less (to 0%) during Phase B. There was significantly less Excess behavior (from 16.7% during baseline to 2.8% during the intervention phase); the ARIMA statistic was also significant, even though these behavior categories were not significantly auto-correlated. The Excess + On-Task category showed a significant decrease (ARIMA ns, though there was no significant auto-correlation), and the On-Task Appropriate category showed a significant decrease (ARIMA ns, though there was again no significant auto-correlation). None of Paul's behaviors in these categories were significantly auto-correlated in the one-to-one instructional situation.

Tables 6.10-6.11display Paul's excess behavior in the free play and one-to-one instructional situations. During free play, only the categories of No Excess and Staring/Gazing were significantly autocorrelated, while only Grabbing Objects and Bolting were significantly autocorrelated during one-to-one instruction. Thus, the non-independent t-test comparison would not be appropriately conservative for Bolting, a behavior which appeared to change significantly across phases according to this statistic (i.e., Bolting shows a reduction in mean percentage duration during skill cluster instruction which cannot be considered to be significantly different from that which occurred during the discrete trial sessions). On the other hand, Bolting does show a significant reduction during free play. Also during free play, Head Slapping decreased significantly. During one-to-one instruction, the category of No Excess shows a significant increase, while Object Spinning, Object Dropping, and Shriek/Scream decreased significantly.

As might be expected, there were no changes in teacher behavior across phases during Paul's free play observation situation (see Table6.12). There were, however, changes in teacher behavior during one-to-one instruction (see Table 6.13). The skill cluster intervention phase was associated with increases in the Neutral + Contact, Neutral Mand + Contact, and Attend Peer/Ignore categories, and with decreases in the Neutral Mand, Approval, and Disapproval + Contact categories.

Discussion and Summary

One of the results of this investigation which may be surprising to many investigators is the absence of autocorrelation across days for most of Paul's behaviors. This variability in Paul's behavior was not entirely unexpected based upon our visual inspection of graphs of his behaviors (and that of many of our other handicapped subjects as well), but it does represent a challenge to behavioral and educational interventions. If Paul's behavior cannot be reliably predicted from day to day, is it reasonable to assume that an intervention will be associated with a reliable improvement in behavior?

For three categories which do show changes, we were able to document improvements in Paul's excess behavior as a function of rearranging the presentation of performance trials into a natural chain of behaviors. The dramatic escalation of no excess, from 46% of the time during discrete trial instruction to 71% of the time during the skill cluster program approach, clearly supports this model for programming for Paul. Various individual excess behaviors also showed an increase during Phase B. For the more general behavioral categories monitored by Observer 2, excess codes declined dramatically while on-task codes increased just as dramatically. Interestingly enough, there were some slight

TABLE 6.10. PAUL: Excess Behavior in Situation 1 (Free Play).

Comparison of mean percent duration of excess behaviors across Baseline (Phase A) and Intervention (Phase B) while Paul is engaged in free play.

Excess Behavior	Phase A (n = 29) $\bar{x}$	SD	Phase B (n = 15) $\bar{x}$	SD	t[a]	p
00 No Excess	29.6	.24	40.8	.19	-1.60	ns
05 Jerky Movement	.1	.00	.1	.00	-.22	ns
06 Body Limpness	5.6	.20	0	0	1.08	ns
25 Mouth Body Parts	15.8	.25	24.9	.24	-1.15	ns
26 Mouthing Objects	9.1	.17	8.3	.09	.18	ns
28 Head Slapping	3.6	.07	0.2	.01	2.60	.01
41 Grabbing Objects	0	0	.4	.00	-1.85	.08
55 Object Spinning	5.2	.11	5.0	.13	.04	ns
56 Object Dropping	2.1	.03	1.2	.02	1.05	ns
60 Cry, No tears	0.1	.01	0	0	.72	ns
62 Shriek/Scream	2.7	.05	.8	.01	2.05	.05
65 Vocalization (Sub/Glottal)	12.2	.12	20.3	.10	-2.20	.05
70 Facial Grimace	.5	.02	0	0	1.30	ns
72 Staring/Gazing	3.9	.10	1.9	.07	.73	ns
80 Gesture/Sign	1.7	.06	0	0	1.12	ns
88 Bolting	10.1	.09	2.8	.03	3.87	.001
94 Climbing	.5	.02	0	0	1.10	ns

[a]Non-independent.

TABLE 6.11.PAUL: Excess Behavior in Situation 2 (One-to-One Instruction).

Comparison of mean percent duration of excess behaviors across Baseline (Phase A) and Intervention (Phase B) while Paul is engaged in one-to-one instruction.

Excess Behavior		Phase A (n = 29) $\bar{x}$	SD	Phase B (n = 15) $\bar{x}$	SD	t[a]	$p<$
00	No Excess	45.5	.23	70.5	.10	-4.99	.0001
05	Jerky Movement	.3	.02	0	0	.91	ns
06	Body Limpness	.4	.01	0	0	1.33	ns
25	Mouth Body Parts	.8	.02	0	0	2.04	.05
26	Mouth Objects	.2	.01	.7	.01	-1.41	ns
28	Head Slapping	.3	.01	0	0	1.51	ns
41	Grabbing Objects	.6	.01	.5	.01	.28	ns
55	Object Spinning	13.8	.10	1.7	.03	5.97	.0001
56	Object Dropping	3.3	.07	.4	.01	2.36	.03
60	Cry, No tears	.3	.02	0	0	.72	ns
62	Shriek/Scream	1.3	.03	0	0	2.40	.03
65	Vocalization (Sub/Glottal)	11.1	.11	14.8	.08	-1.15	ns
70	Facial Grimace	.8	.02	0	0	1.82	.08
72	Staring/Gazing	.5	.02	0	0	.95	ns
80	Gesture/Sigh	8.2	.10	9.7	.07	-.52	ns
88	Bolting	3.8	.05	.9	.01	3.00	.005
94	Climbing	0	0	0	0	--	--

[a]Non-independent.

TABLE 6.12. PAUL: Teacher Behavior in Situation 1 (Free Play).

Comparison of mean percent duration of behaviors across Baseline (Phase A) and Intervention (Phase B) while Paul is engaged in Free Play.

Teacher Behavior	Phase A (n = 26) $\bar{x}$	Phase A (n = 26) SD	Phase B (n = 14) $\bar{x}$	Phase B (n = 14) SD	t[a]	$p<$
0 Not Present	95.0	.07	96.9	.04	-.90	ns
1 Neutral	1.7	.06	0	0	1.07	ns
2 Neutral & Contact	.9	.02	.9	.01	-.13	ns
3 Neutral Mand	1.5	.02	.9	.02	.85	ns
4 Neutral Mand & Contact	.3	.01	1.0	.01	-1.60	ns
5 Approval	0	0	0	0	—	—
6 Approval & Contact	0	0	0	0	—	—
7 Disapproval	.3	.01	0	0	.99	ns
8 Disapproval & Contact	.2	.01	.2	.01	.04	ns
9 Attend Peer/Ignore	0	0	0	0	—	—

[a]Non-independent t-test.

TABLE 6.13. PAUL: Teacher Behavior in Situation 2 (One-to-One Instruction).

Comparison of mean percent duration of behaviors across Baseline (Phase A) and Intervention (Phase B) while Paul in engaged in one-to-one instruction.

Teacher Behavior		Phase A (n = 26) $\bar{x}$	SD	Phase B (N = 14) $\bar{x}$	SD	t [a]	$p<$
0	Not Present	.2	.01	7.5	.27	-1.03	ns
1	Neutral	13.6	.12	20.0	.13	-1.60	ns
2	Neutral & Contact	5.8	.07	22.5	.15	-3.86	.001
3	Neutral Mand	54.2	.18	14.5	.10	9.09	.0001
4	Neutral Mand & Contact	8.1	.08	22.0	.15	-3.20	.005
5	Approval	13.3	.09	4.4	.04	4.47	.0001
6	Approval & Contact	2.3	.05	4.9	.04	-1.59	ns
7	Disapproval	1.0	.02	.6	.02	.61	ns
8	Disapproval & Contact	.6	.01	0	0	2.07	.05
9	Attend Peer/Ignore	.7	.02	3.5	.05	-1.88	.08

[a]Non-independent t-test.

improvements in various behaviors while Paul was observed in free play also, although this observation condition was virtually identical across baseline and intervention phases.

The obvious confounding variable in these definite changes in Paul's behavior is, however, teacher behavior. The mean percentage duration of various instructional behaviors changed considerably during one-to-one instruction, including significant increases in two "neutral" contact categories (Neutral + Contact and Neutral Mand + contact) and significant decreases in the Approval, Disapproval + Contact, and the Neutral mand categories. Would Paul's behavior have changed similarly across baseline and intervention phases in which these teacher instruction behaviors only had been manipulated, but within the context of similar session structures? Our study design and results do not allow us to answer this question. We do know that the changes in teacher behavior during one-to-one instruction did not also occur in the free play condition during the skill cluster phase, although certain of Paul's behaivors had changed in both conditions. It would be interesting, then, to control for certain types of teacher behaviors across instructional phases. To some extent, however, at least some change in teacher instructional behaviour would be an inevitable function of the change to a skill cluster instructional design. In general, the positive changes in certain of Paul's behaviors support the change in instructional format. While these data were not included in our data collection, reports of improved skill acquisition during Phase B by Paul's teacher offered equally encouraging support for further utilization of this skill cluster approach to programming.

Chapter 7
Implications for Future Research and Practices

Evidence of response interrelationships obviously complicates the process of educational assessment and planning (Voeltz & Evans, 1982). Kara and Wahler (1977) noted that predictions of multiple behavior changes following the manipulation of a single target behavior could only be made from their factor analysis. Our research indicates that such predictions, again, can be made based upon multivariate analyses of children's repertoires, but that this process is enormously complicated. The kind of complex observations and data analyses undertaken here are likely to be viewed as prohibitively expensive and thus of limitted utility for application in children's programs. To date, proponents of behavioral assessment have emphasized its practicality and relatively low "cost" for use by teachers and other clinicians, but the picture changes once multiple effects become the focus of an intervention. It is true, of course, that professionals in related fields who provide input for such children have been far less conservative in recommending, for example, extensive diagnostic/neurological examinations which seldom provide information relevant to educational prescription (Bricker & Campbell, 1980). Perhaps, then, it would be appropriate to rethink the level of sophistication required for meaningful behavioral assessment. For children who are severely cognitively delayed and additionally present complex behavior problems, the cumulative effects of considering only a few educational/behavioral targets a year-with little to no regard for the effects of each behavior upon the child's total repertoire within a single year and across time-are unlikely to produce maximum results. Ultimately, we may have to question the wisdom of planning handicapped children's precious educational time on the basis of minimum data probes conducted for single, arbitrarily selected target behavors.

At the very least, our procedures and results suggest the relevance of multivariate behavoral assessment for serious investigation of severely handicapped children's educational needs. Where clinical-experimental settings have the resources and technology to implement these kinds of complex behavioral assessment procedures and investigate multiple outcomes, they should do so. Research in this area must move away from the less intensive, short term, single-target manipulations which now typify the field. Lifelong planning for severely handicapped children must include a consideration of multiple effects as part of the criteria for maximum efficiency and effectiveness in programming.

Practical Alternatives for Special Education

Since most educational programs are unlikely to have the technology to systematically monitor children's total repertoires, several alternative procedures can and should be utilized to allow the collection of information useful for documentation of meaningful behavior change. Ironically, both behavior therapists and special education training programs have consistently advocated the use of certain complex assessment and evaluation strategies which are not well-suited for use in educational programs and which yield little information useful for educational planning. We have explicitly critiqued strategies recommended for use in special education classrooms to determine child progress and program effectiveness, particularly single-subject experimental designs and developmentally-based assessments (Voeltz & Evans, in press). In their place, we proposed an assessment/

evaluation model which would be feasible in public school program, for use by teachers in classroom settings, and which does allow for decisions regarding multiple and meaningful outcomes and appropriate instructional planning on a day-to-day basis.

This instructional evaluation model poses three major questions regarding the effects of interventions which must be answered in order to evaluate programs for children. Each will be summarized below, and specific recommendations will be made regarding measurement processes which could be used to address each issue. More detail is available in Voeltz and Evans (in press), but briefly, the following criteria must be considered:

(1) Has the behavior change occurred and is the change a function of the program?

A simple time series analysis should be sufficient to establish behavior change. This should include periodic systematic data probes conducted before, during, and after instruction in multiple (relevant) situations and environments. For some programs, a daily class record or a "diary" rotating from school to home may be adequate to document pupil performance changes.

Where behavior shows a clear pattern of expected improvement over time--as is typical of successful skill acquisition--this change can reasonably be attributed to the effects of the intervention. The teacher is likely to be well aware of ususal threats to internal validity, such that single subject designs to control for such threats are unnecessary. Furthermore, we expect multiple effects (generalization and maintenance) in many cases; most single-subject designs cannot accomodate such changes. Finally, teachers must be able to continuously adjust individualized programs as needed, on a daily basis in some cases, and cannot utilize an "evaluation" model which does not allow such adjustments within a "treatment phase".

(2) Did the educational intervention occur as specified in the intervention plan (Educational integrity)?

Teachers must periodically monitor the instructional environment, including observations of such variables as teacher affect and faithfulness to the written program plan, physical arrangements, daily activities (e.g., did the program even occur the planned number of times during a given week?), etc., in addition to monitoring the individual performance of the student. An informal functional analysis of the instructional program should be regularly conducted as a source of valuable information on effective strategies for individual students.

In order to monitor the educational integrity of children's programs, teachers and ancillary staff must be prepared to self-evaluate and engage in peer evaluations. In addition, simple records of the actual instructional time available for each program can be kept, which would eventually yield valuable information regarding the amount of instructional time needed for learning particular skills across children.

(3) Is the resultant behavior change meaningful? Is it beneficial for this child's eventual outcome (empirical validity)? Is the change valued by those in the natural environment of that person (social validity)?

Empirical validity. Teachers must collect information on possible multiple effects of changes in children's behavior. By formulating informal hypotheses regarding predicted changes in several behaviors concurrent with successful modification of a particular target behavior or skill, these behaviors can be probed systematically during (preintervention) assessment and during the intervention phase. A structured interview could be conducted (over the phone or with a short questionnaire), asking parents and other caregivers to specify behavior patterns where they suspect that certain behaviors always occur together, one before another, etc. These "hypothesis" could then be used to decide which behaviors seem most relevant to monitor more systematically where this is possible. Alternatively, parents could be interviewed again following a change in the intervention target regarding possible multiple effects. Eventually, this "informal" or clinical data would accumulate to support particular formal hypothesis which could be tested experimentally.

Eventually, programs could establish the validity of particular goal selection choices as ones which do, in fact, result in optimal effects upon the child's total repertoire. Data on successful skill acquisition should also focus upon demonstrations that the child can use the new skill to perform an essential life-function, rather than requiring that children perform an isolated target skill in an artificial massed trial format according to arbitrary accuracy and reliability criteria which are idiosyncratic to that situation.

Social validity. Persons who are in a position to significantly affect outcomes for severely handicapped persons--including the success of a community placement, etc.--should participate in the specification of priority goals and instructional objectives. Thus, parent priorities for instruction should be incorporated into instructional plans, and employers should be asked what they consider the most important skills and behaviors needed for successful adjustment in their setting, etc. Once such behaviors have been acquired, the relevant persons in the community can also help specify how much of a behavior or what degree of accuracy is needed before the skill is actually meaningful in the natural environment.

Input from parents, etc., can easily be obtained through initial telephone consultations, interviews in the criterion environment where the teacher would also observe the level of skill performance needed for fluency, and evaluations of the student's actual performance in those criterion environments following instruction, in which the performance would be rated by parents, etc. The major concerns would be that the behaviors to be taught are in fact valued by the child's actual environments, and that the behaviors acquired are mastered sufficiently to reflect actual demands of those environments.

Selecting Priority Goals for Children with Multiple Needs

When we began our research effort, we were concerned that no clear guidelines exist as to when a goal or objective included on a handicapped child's Individualized Education Program (IEP) might target the modification of an excess behavior, rather than the acquisition of a new skill. First, existing criteria for goal selection in general are generally applied to a single skill or excess behavior in isolation, without regard to others in the child's repertoire or the environmental context. For example, Heads' (1980) 5-point checklist for selecting potential individual goals does not deal with the selection of one goal over another of equivalent concern. Second, recommendations which are made regarding whether to teach a skill or modify an excess first are not consistent in the behavioral and educational intervention literature. For example, Koegel, Egel, and Dunlap, (1980) emphasize the need to reduce problem behavior prior to instruction, based upon previous (though limited) evidence that certain behaviors may actually prevent learning (Koegel & Covert, 1972). Alternatively, Gaylord-Ross (1980) and Schroeder, Mulick and Schroeder (1978) recommend teaching an incompatible skill, which performs the same function as the excess behavior for the handicapped child, as a most effective and thus preferred strategy to reduce problem behavior; this position can be supported by evidence that contingency management studies dealing with excess behavior fail to produce lasting behavior change in many, if not most cases (Carr, 1980; Derer & Hanashiro, 1982). In the absence of any empirically-based concensus regarding the selection of priority goals and appropriate intervention strategies, educational planning and programming on behalf of behavior disordered severely handicapped children will be particularly problematic. Parents, teachers, and other professionals and caregivers could only rely on personal perceptions and biases regarding the advantages and disadvantages of targeting certain behaviors for change and selecting from among several possible intervention plans. While misjudgements in educational planning are always unfortunate, they are particularly so where the children involved have severe learning problems and multiple programmatic needs. Each decision may or may not result in successful behavior change or skill acquisition, but also represents an opportunity lost, since other behaviors were not modified and other skills were not taught. For some children - such as autistic children--there is evidence that certain excess behaviors may remain in the child's repertorie for years despite all intensive efforts to decrease or extinguish them through behavior management procedures. If skill instruction is actually postponed until such behaviors are under control, these children would lose already limited learning time. The ultimate consequence for the child and his/her caregivers is clear: sooner or later, the child reaches the age at which a free and appropriate special education ends, and unless certain skills necessary for adult functioning have been acquired, the handicapped person's opportunities to participate maximally in integrated community environments and opportunities will be severely limited. Planning and evaluation become crucial. Decisions made during each school year must reflect a consideration of the child's total repertoire, his/her needs with reference to the demands of current and future environments, and provision of sequential opportunities for learning across the child's school career which will ultimately produce optimal individual gains or maximum participation in society as an adult.

A major product of our literature review and research effort, therefore, was the development of a decision model to assist teachers and caregivers on making systematic choices from among multiple potential instructional and behavioral objectives and intervention procedures for individual children (Voeltz,

Evans, Derer,& Hanashiro, 1982). This model is based upon the following major assumptions regarding behavior modification in school settings and the responsibilities of clinicians toward the children they serve:

(1) The IEP is an Educational Plan

By legal statute, a handicapped child's experiences in any special education program must reflect the goal priorities specified in the Individualized Education Program (IEP). The purpose of this IEP is to provide a plan for education according to goals and priorities designed to establish new skills in children with severe deficits rather than focus upon reducing excess behavior: that is, the IEP is the framework for developing skills needed by the handicapped individual for optimal participation and independence in integrated community environments.

(2) Not all Excess Behaviors are Priority Targets

Since any instructional or behavior management program involves the consumption of valuable teaching (staff) and learning (child) time, use of such time to decrease rather than increase behavior in a child's repertoire should occur only when unavoidable. Behaviors that interfere with the performance of a needed skill might be justified as targets for interventions. Behaviors that prevent children from functioning capably in their social environment may also warrant consideration as priority targets for change. On the other hand, social pressure, expedience, or personal preferences of caregivers lack such clear ethical justification as reasons to target a behavior for change: such intentions are not directly related to the child's best interests, but tend to serve the needs of others. Behaviors identified for reasons that fail to consider the child's total repertoire and potential outcomes should be dealt with by means other than the child's educational program.

(3) To Decrease a Behavior, Increase a Skill

There is considerable evidence that excess behaviors are best remedied, in terms of producing lasting and generalized behavior change, by replacing them with the skills needed to deal more positively with the situations associated with such problem behavior. We recommend that professional resources and research efforts be concentrated on the development and dissemination of positive alternatives to modify negative behavior through educative approaches. Positive functional skills which provide children with strategies to interact appropriately with and gain control over their environments are readily maintainable and generalizable. On the other hand, a behavioral intervention designed to decrease an excess behavior will not produce lasting, generalized behavior change if the child has no alternative strategy to accomplish the function performed by the excess. Thus, the task is to identify and teach a positive behavior that can replace the negative one precisely because it addresses this function for the child.

A Clinical Decision-Model Reflecting Educational Best Practices

The clinical decision model developed by the Behavioral Systems Intervention Project was designed to (1) formalize procedures which professionals currently use in an intuitive, non-specific manner (see Chapter 4); (2) incorporate available

professional and legal-ethical guidelines; (3) synthesize existing empirical data on behavioral interrelationships (see Chapters 1, 5 and 6) and effective interventions (see Chapter 6 and Derer and Hanashiro, 1982); (4) allow for the incorporation of a teacher's personal knowledge regarding each child's behavior; and (5) requires simultaneous and systematic consideration of each child's skill acquisition and behavioral needs while the IEP is being formulated and revised throughout programming. The model is, thus, intended to reflect current recommendations regarding educational best practices, in particular: (1) the major emphasis is always upon teaching the child new skills, and (2) the major emphasis is always upon utilizing least intrusive (or "normalized") intervention strategies which are feasible in actual public school environments.

Figures 7.1 - 7.3 illustrate the series of questions which the user must answer in order to select priority goals and appropriate interventions on behalf of children with multiple needs. The three levels of the model reflect differences in the severity of the behavior, effects on potential child outcomes, and the various professional and legal-ethical criteria relevant to the selection process. Each decision point requires skilled judgements, for which data might not always be available, either with respect to that particular child or for a particular behavior problem across children. The model is not intended to provide an error-proof or rigid prescription for intervention; eventually, it may be possible to recommend that a certain behavior (e.g., finger flicking) should or should not be modified based upon information regarding the outcome of that decision for a representative number of handicapped children, and it might be possible to pinpoint specific strategies for use with specific behaviors, but available information is insufficient to support such prescriptions.

The Flow Chart. The flow chart is organized into three levels, but all decisions regarding excess behaviors begin at Circle A on Level I (see Figure 7.1). Behaviors are not determined in advance to be at any particular level, although those which fall into Level I will be clearest immediately to persons who know the child. Whether or not a behavior is addressed at Level II (see Figure 7.2) or Level III (see Figure 7.3) depends upon the answers to several questions, and a Level III decision is the result of having dealt with all of the issues from preceding levels. Thus the flow chart can be conceptualized as a systematic movement through a sequence of considerations reflecting the seriousness of consequences for either changing or choosing not to change the behaviors identified during assessment.

Level I decisions focus on excess behaviors that pose a threat to the life of the child or are likely to result in irreversible physical harm. Level II decisions focus on behaviors that have direct serious consequences for the child, in that they may be dangerous to others or have potential for becoming more serious in the future. Level III decisions focus upon behaviors whose negative effects tend to reside in the child's social environment. The consequences of these behaviors are indirect and include factors such as community acceptance and damage to the environment. More information on use of the model is provided in Voeltz, Evans, Derer, and Hanashiro (1983), and our manual details procedures and individual examples for use by teachers (Evans, Derer, Voeltz, & Hanashiro, (1982), but a brief summary of the behavior levels will be provided here:

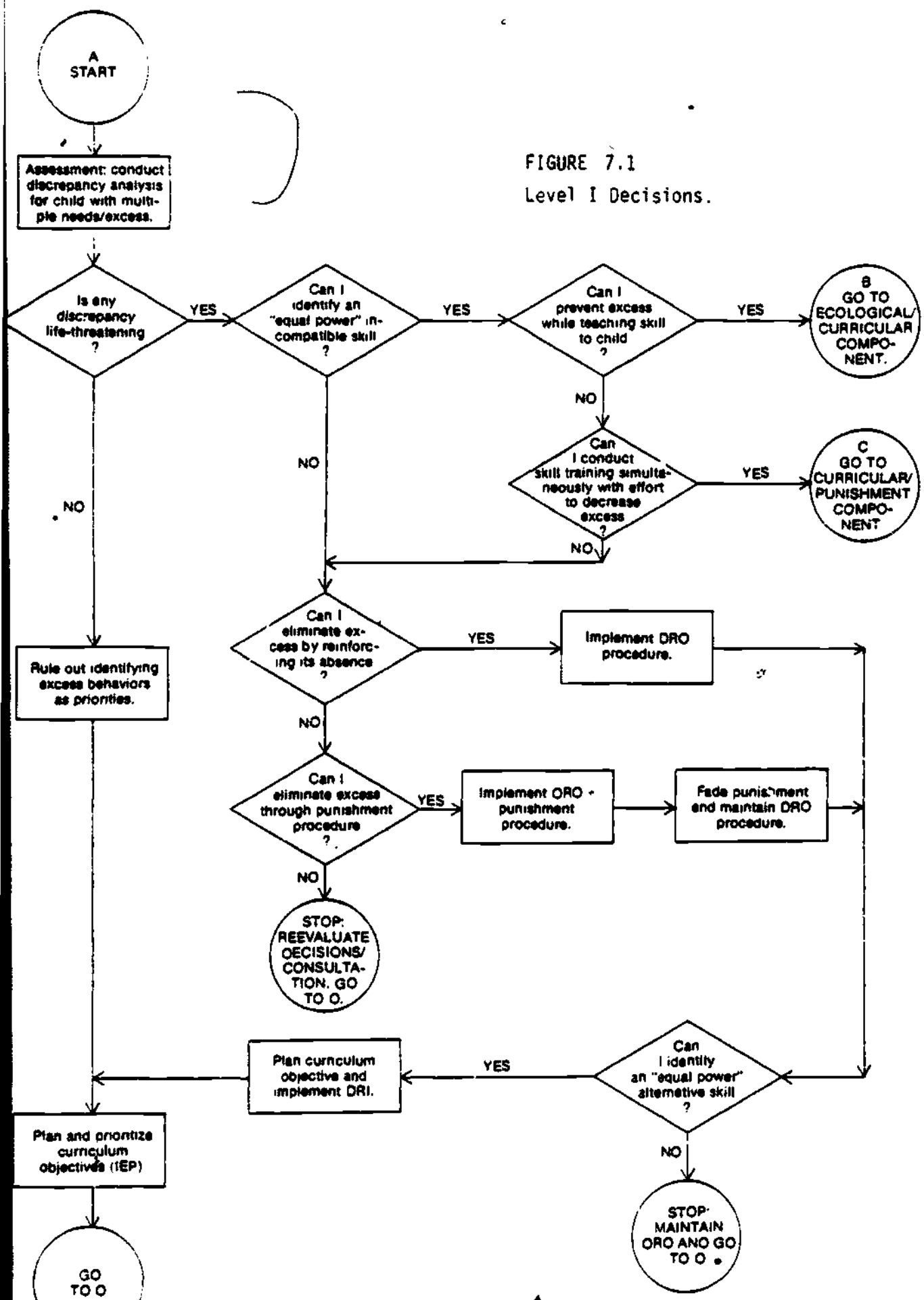
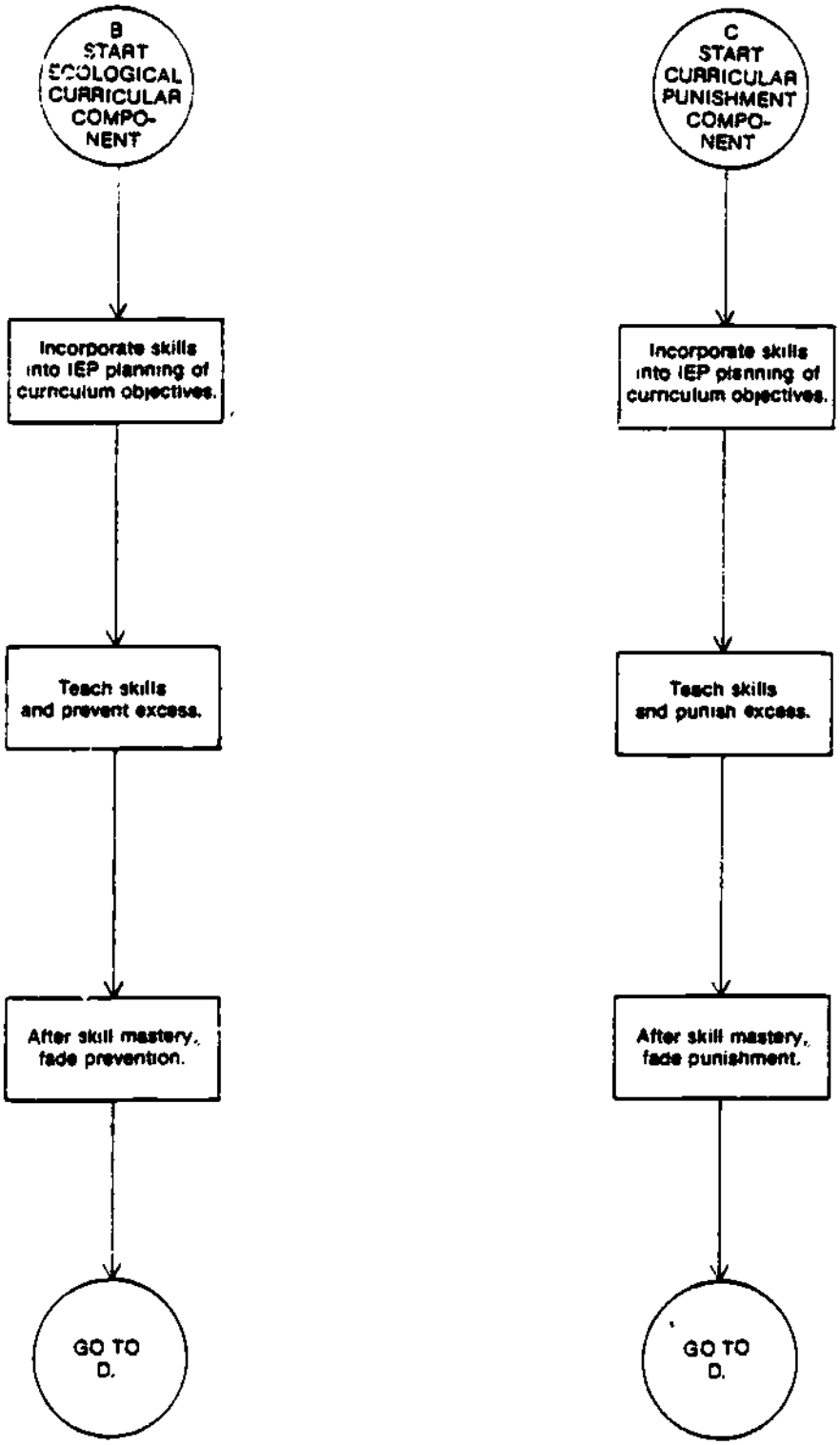


FIGURE 7.1
Level I Decisions.

FIGURE 7.1 (cont.)



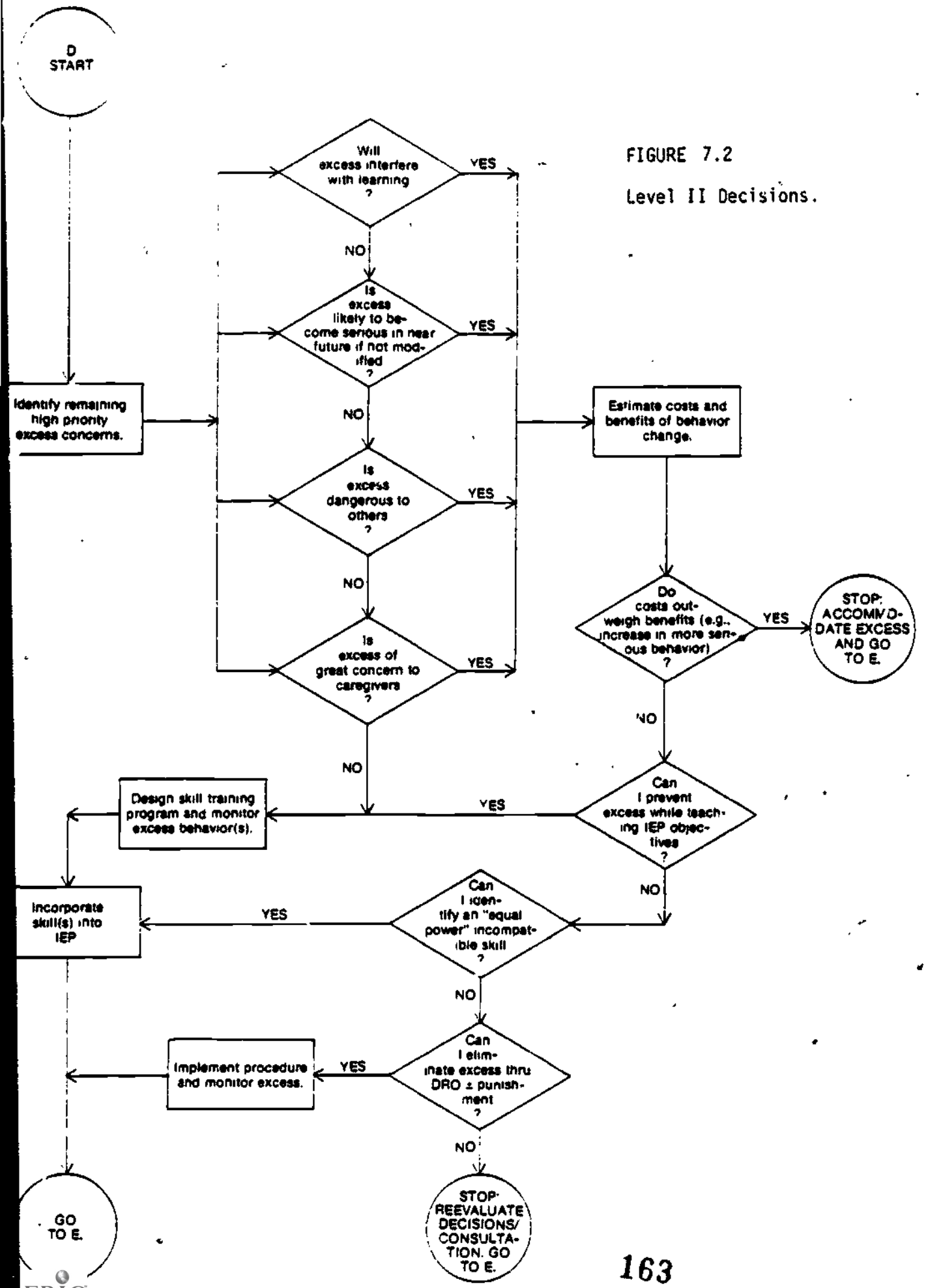
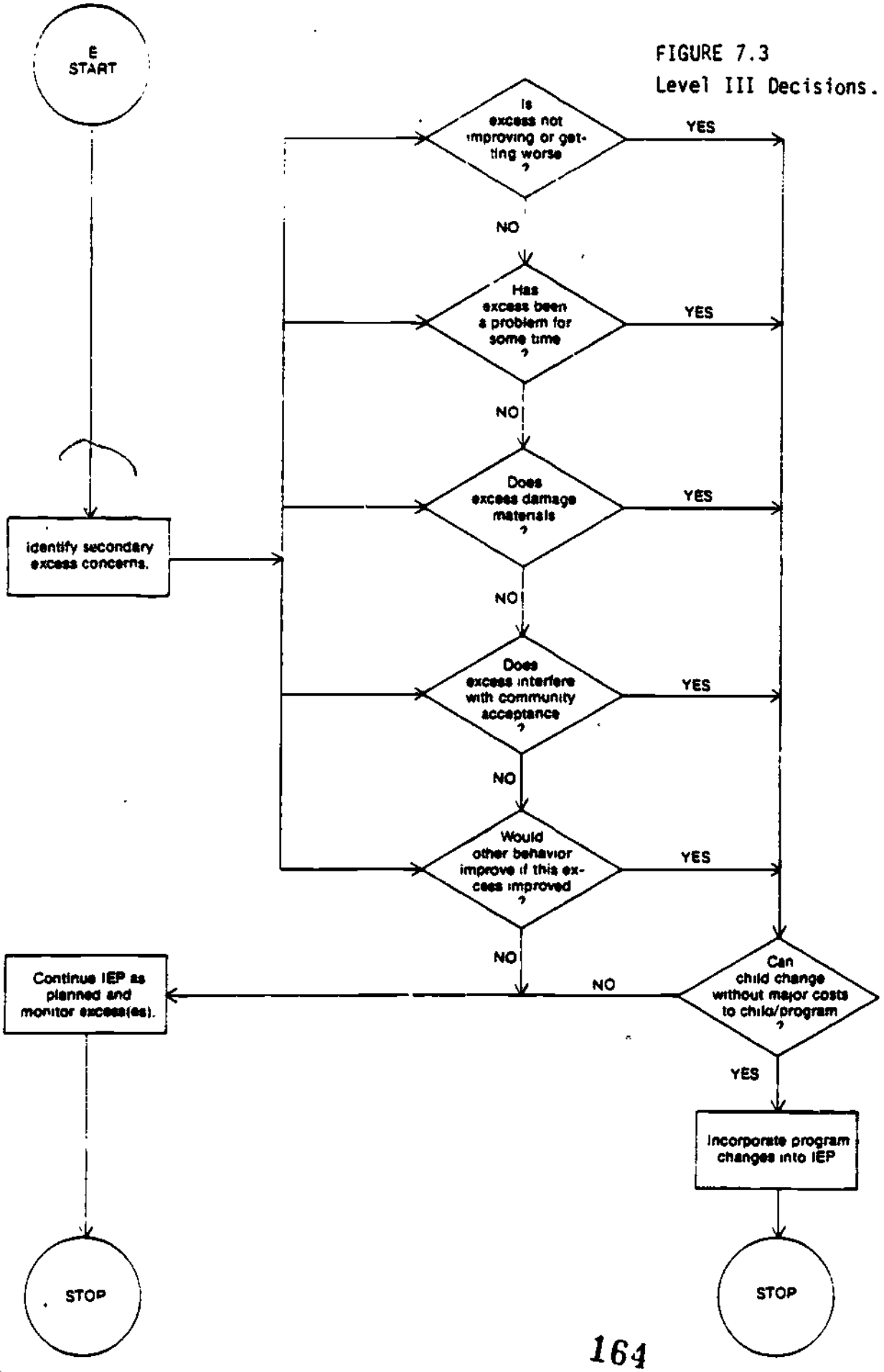


FIGURE 7.2

Level II Decisions.

FIGURE 7.3
Level III Decisions.



(1) Level I - Urgent Behaviors Requiring Immediate Attention

The first level of the flow chart (see Figure 7.1) considers the relatively rare instances of life or health threatening excess behaviors which cause irreversible physical harm. Behaviors such as eye poking and head banging could be considered life or health threatening if they occur with high frequency or intensity. Another example would be chronic vomiting with concomitant weight loss. The collateral effect of weight loss identifies this particular excess as life-threatening. Without the weight loss (or other physical effects such as dehydration or deterioration of the esophagus), the urgency of the behavior is reduced, and the excess would be more appropriately considered at another level of the flow chart. Behaviors that are neither life nor health threatening are ruled out as priorities on Level I. The clinician (teacher or consultant) and parents would then write IEP goals and plan and prioritize the curriculum objectives based upon discrepancies between the severely handicapped child's skills and those needed for maximum participation in targeted community settings (Brown, Branston, Hamre-Nietupski, Pumpian, Certo, & Gruenewald, 1979). With the IEP formulated, the user moves to Level II, Circle D.

If a life or health threatening excess behavior were identified, the next task is to select an equal power incompatible skill. An equal power incompatible skill has two major characteristics:

1. The skill is topographically incompatible with excess behavior in that the skill and the excess can not be performed simultaneously;

2. The skill generates sufficient opportunities for reinforcement, thereby allowing it to compete with the excess behavior.

In order to identify an equal power incompatible skill, the clinician needs to be aware of what reinforces the behavior and what skill could replace it while maintaining the same level of reinforcement. With a behavior such as chronic vomiting, social interaction in the form of negative attention or the opportunity to play with the vomit may function to maintain the behavior. If an equal power incompatible skill can be identified, the first choice for intervention would be an ecological/curriculum component, in which the teacher would train the skill while preventing the excess from occurring. This approach can be particularly effective when vomiting is being induced by hand mouthing or tongue and throat manipulations. A possible curricular intervention would be the training of a leisure skill requiring object manipulation with both hands. The physical redirection of the child's hands away from the face combined with social praise and the intrinsically reinforcing properties of the activity serve to identify the skill as incompatible and equal power. In the event that the response can not be prevented, a curricular/punishment component would be implemented in which skill training would be conducted simultaneously with efforts to decrease the excess. An example would be teaching appropriate feeding skills and momentarily removing the food paired with a verbal "No" when the child begins the response chain that leads to vomiting.

If skill training can not be conducted simultaneously with efforts to decrease the excess, the next decision point examines the possibility of eliminating the excess behavior through the reinforcement of its absence (differential reinforcement of other behaviors, or DRO). A major consideration in using DRO is whether the excess behavior occurs at a low enough frequency or duration to allow the reinforcement of other behaviors. Only if the excess can not be eliminated through [illegible] of DRO alone could a punishment procedure then be added to

the program. If the punishment procedure does not produce a reduction in the excess behavior, then a complete re-evaluation of the decisions and procedures must be conducted. This would be an appropriate time for practitioners to consider additional outside consultation. If the punishment procedure succeeds in establishing control over the behavior, then the punishment procedure is faded while continuing DRO. In the next step, an alternative skill is identified and a program is implemented to teach the skill--differential reinforcement of incompatible behavior (DRI). Finally, the curriculum objectives are planned and prioritized with DRI as part of the curriculum.

(2) Level II - Serious Behaviors Requiring Formal Consideration

On this level, guidelines are suggested for selecting priority goals from among the more serious but not life or health threatening behaviors (see Figure 7.2). These high, but not urgent, priority concerns are delineated by the immediate and serious consequences of the excess. At Level II, the user ask the following questions: Will the excess interfere with learning; is the excess likely to become serious in the near future if not modified; is the excess dangerous to others, and; is the excess of great concern to caregivers? If all of the answers to the questions are "no", a skill training program is designed and incorporated into the IEP, and the excess behavior is monitored. The user would then proceed to Level III, Circle E. If there is an affirmative answer to any of the previous questions, a cost-benefit analysis of behavior change must be conducted.

A cost-benefit analysis considers potential child outcomes resulting from a decision to decrease a behavior. For example, if the anticipated result of decreasing verbal aggression were a decrease in physical aggression, then a benefit would be realized. On the other hand, a potential cost of decreasing verbal aggression might be a decrease in positive verbal interaction. If the costs of behavior change outweigh the benefits, then the behavior must be accomodated. If the benefits outweigh the costs, then decisions similar to those for Level I must be made: Can the behavior be prevented while teaching current IEP objectives; can a new, incompatible skill be identified; can the behavior be eliminated through DRO combined with a punishment procedure? If the behavior can not be remedied through one of these three methods, a team re-evaluation should be conducted and perhaps outside consultation requested.

(3) Level III - Excess Behaviors Reflecting "Normal Deviance"

Level III considers secondary excess behaviors (see Figure 7.3). This level offers guidelines for dealing with excess behaviors whose negative effects tend to reside more in the child's social environment and whose presence does not directly threaten others or interfere with learning. Finger flicking or object banging might be considered on this level. After identifying the remaining excess concerns, the user asks the following questions: Is the excess not improving or getting worse; has the excess been a problem for some time; does the excess damage materials; does the excess interfere with community acceptance, and; would other behaviors improve if this excess improved? If all the answers to these questions are "no", then the IEP would continue as planned(i.e., the IEP would not be modified to include an objective to decrease the excess behavior), and the excess behavior would simply be monitored. If there is an affirmative answer to at least one of these questions, the user should then consider whether the child can change without major costs to the child or program.

The cost-benefit question on Level III examines the seriousness of the behavior, potential increases in other equally serious behaviors, and the proportionate program effort which would be needed in order to successfully modify the behavior. Level I and Level II behaviors involve consequences serious enough to warrant intervention regardless of whether or not the resultant use of staff time and resources represents a major inconvenience for the educational program. Level III behaviors have already been determined to be relatively minor in terms of consequences for the child, since answers to the more serious questions at Levels I and II were all negative for these behaviors. Thus, it is often not appropriate to use valuable program time and resources-- which are needed to teach the child new skills--to modify a Level III behavior. The user should consider the possible collateral effects resulting from efforts to change the behavior, what behaviors would remain in the child's repertoire if these behaviors were eliminated, what type of environmental restructuring can take place that would prevent the excess yet maintain efficient use of classroom facilities, staff time required for change, etc. If the behavior can be changed without major costs to the child or the program, the program changes would be incorporated into the IEP. If the child can change only with major costs, the excess behavior would be monitored and the IEP continued as originally planned.

SUMMARY

The flow chart is offered as a means for teachers, psychologists, and other behavior modification professionals to organize the various factors involved in selecting priority goals and making intervention decisions. Rather than providing answers, it specifies important considerations which must be addressed with regard to each child and each situation. In the absence of conclusive empirical data to support particular goal selection decisions, clinicians will still need to exercise their best professional judgements. We are currently working with several teachers who are using the model, and their input will be utilized to develop a revised version. One of the studies planned for the University of Minnesota Consortium Institute for the Education of Severly Handicapped Learners involves an experimental investigation of feasibility (for public settings) and effectiveness (for child outcomes).

In any event, decisions on whether or not to intervene with a particular behavior should not be based on the success ratio of a particular intervention, nor should a behavior be evaluated in isolation without consideration of the child's total repertorie. Rather, the decision process can be approached from a perspective that integrates available empirical evidence, a comprehensive needs assessment, systematic analysis of the seriousness of the behavior a consideration of potential child outcomes, and concern for ethics and legal precedent.

REFERENCES

Bachman, J.A. Self-injurious behavior: a behavioral analysis. Journal of Abnormal Psychology, 1972, 80, 211-224.

Baer, D.M. A note on the absence of a Santa Claus in any known ecosystem: a rejoinder to Willems. Journal of Applied Behavior Analysis, 1984, 7, 167-170.

Baldwin, V., personal communication, 1979.

Balthazar, E.E. & Phillips, J.L. Autistic-like behavior in mentally retarded residents. Madison, WI: Central Wisconsin Center for the Developmentally Disabled, 1977.

Baumeister, A.A. Orgins and control of sterotyped movements. C.E.Meyer (Ed.), Quality of life in severely and profoundly mentally retarded people: Research foundations for improvement. Washington, D.C.: AAMD, 1978

Baumeister, A.A. & Rollings, J.P. Self-injurious behavior. N.R. Ellis (Ed.) International review of research in mental retardation, Vol 8, New York: Academic Press, 1976.

Becker, J.V., Turner, S.M. & Sajwaj, T.E. Multiple behavioral effects of the use of lemon juice with ruminating toddler-age child. Behavior Modification, 1978, 2:2, 267-278.

Bergin, A.E. Some implications of psychotherapy research for therapeutic practice. Journal of Abnormal Psychology, 1966, 71, 235-246.

Bernal, M.E. & Miller, W.H. Electrodermal and cardiac responses of schizophrenic children to sensory stimuli. Psychophysiology, 1971, 7, 155-168.

Bornstein, M., Bellack, A.S., & Hersen, M. Social skills training for highly aggressive children;treatment in an inpatient psychiatric setting. Behavior Modification, 1980, 4, 173-186.

Borthwick, S.A., Meyers, C.E., & Eyman, R.K. Comparative adaptive and maladaptic behavior of mentally retarded clients of five residential settings in Western states. R.H. Bruinenks, C.E. Meyers, B.B.Sigfor, & K.C. Laking (Eds Deinstitutinalization and community adjustment of mentally retarded people. Washington, D.C. AAMD, 1981.

Bower, E.M. Primary prevention in a school setting. In G. Caplan (Ed.), Prevention of mental disorders in children. New York: Basic Books, 1961.

Bradtke, L.M., Kirkpatrick, W.J.,& Rosenblatt, K.P. Intensive play: a technique for building affective behaviors in profoundly mentally retarded children. Education and Training of the Mentally Retarded. 1972, 7, 8-13.

Brannigan, C.R. & Humphries, D.A. Human non-verbal behavior, a means of communication. In N.G. Blurton Jones (Ed.), Ethological studies of child behavior. London: Cambridge University Press, 1972.

Brennan, J.. & Freedland, K. Behavioral Observation, Data Management and Analysis Package. Honolulu: University of Hawaii Behavioral Systems Intervention Project, 1981.

Bricker, W.A. & Campbell, P.H. Interdisciplinary assessment and programming for multihandicapped students. In W. Sailor, B. Wilcox, & L. Brown (Eds.) Methods of instruction for severely handicapped students. Baltimore: Paul H. Brookes, 1980.

Brown, L. Introduction to educational programs for severely handicapped students. Minneapolis: Upper Midwest Regional Resource Center and the University of Minnesota, 1982.

Brown, L., Branston, M.B., Hamre-Nietupski, S., Pumpian, I., Certo, N., & Gruenewald, L. A strategy for developing chronological-age-appropriate and functional curricular content for severely handicapped adolescents and young adults. The Journal of Special Education, 1979, 13, 81-90.

Buium, N, Rynders, J. & Turnure, J. Early maternal linguistic environment of normal and Down's Syndrome language-learning children. American Journal of Mental Deficiency, 1974, 79, 52-58.

Butler, J.A. & Stenmark, S. P.L. 94-142: Can we evaluate outcomes? Washington, D.C.: Institute for Health Policy Studies, July 1981.

Carr, E.G. The motivation of self-injurious behavior: a review of some hypotheses. Psychological Bulletin, 1977, 84, 800-816.

Carr, E.G. Generalization of treatment effects following educational intervention with autistic children and youth. In B. Wilcox & A. Thompson (Eds.), Critical issues in educating autistic children and youth. Washington, D.C.. U.S. Department of Education, Office of Special Education, 1980.

Carr, E.G., Newsom, C.D., & Binkoff, J.A. Stimulus control of self-destructive behavior in a psychotic child. Journal of Abnormal Child Psychology, 1976 4, 139-152.

Carr, E.G., Newson, C.D., & Binkoff, J.A. Escape as a factor in the aggressive behavior of two retarded children. Journal of Applied Behavior Analysis 1980, 13, 113-129.

Casey, L.O. Development of communicative behavior in autistic children: A parent program using manual signs. Journal of Autism and Childhood Schizophrenia, 1978, 8, 45-59.

Cone, J.D. Assessing the effectiveness of programmed generalization. Journal of Applied Behavior Analysis, 1973, 6, 713-718.

Corbett, J.A. & Campbell, H.J. Causes of self-injurious behavior. P. Mittler (Ed.), Frontiers of knowledge in mental retardation, Vol. II Biomedical aspects. Baltimore: University Park Press, 1981.

Derer, K.R., & Hanashiro, R. Behavior Inverventi.on Index. Honolulu: University of Hawaii Behavioral Systems Intervention Project (Technical Report #9), 1982.

Elstein, A.S., Schulman, L.S.,& Sprafka, S.A. Medical problem solving: An analysis of clinical reasoning. Cambridge, Mass:Harvard University Press, 1978.

Epstein, L.H., Doke, L.A., Sajwaj, T.E., Sorrell, S. & Rimmer, B. Generality and side effects of overcorrection. Journal of Applied Behavior Analysis, 1974, 7, 385-390.

Evans, I.M., Derer, K.R., Voeltz, L.M. & Hanashiro, R. A manual for teacher decision making in selecting educational and behavioral management objectives and intervention strategies for children with multiple needs. Honolulu: University of Hawaii, 1982.

Evans, I.M., Freedland, K., Lipton, D., & Voeltz, L.M. Behavioral Sytems Intervention Project Observer training manual. Honolulu: Behavioral Systems Intervention Project (Technical Report #6), 1980.

Evans, I.M., & Voeltz, L.M. Designing behavioral programs for severely handicapped students: A practical decision model for educators. Baltimore: Paul H. Brookes, in preparation.

Evans, I.M., Voeltz, L.M., Freedland, K., & Brennan, J.M. Behavioral interrelationships in the design and evaluation of applied intervention research. Honolulu: University of Hawaii Behavioral systems Intervention Project (Technical Report #1), 1981.

Federal Register, 1975, 4-:35, 7412.

Flavell, J.E. Reduction of sterotypes by reinforcement of toy play. Mental Retardation, 1973, 11, 21-23.

Forehand, R. & Baumeister, A.A. Deceleration of aberrant behavior among retarded individuals. In M. Hersen, R.M. Eisler, & P.M. Miller (Eds.) Progress in behavior modification, vol. 2. New York: Academic Press, 1976.

Foxx, R.M. & Arzin, N.H. The elimination of autistic self-stimulatory behavior by overcorrection. Journal of Applied Behavior Analysis, 1973, 80, 1-14.

Frankel, F. & Simmons, J.Q. Self-injurious behavior in schizophrenic and retarded children. American Journal of Mental Deficiency, 1976, 80, 512-522.

Freedland, K., & Brennan, J. FAST Computer Program Library, Honolulu: University of Hawaii Behavioral Systems Intervention Project (Technical Report #11), 1981.

Gaylord-Ross, R. A decision model for the treatment of aberrant behavior in applied settings. In W. Sailor, B. Wilcox & L. Brown (Eds.), Methods of instruction for severely handicapped students. Baltimore: Paul H. Brookes, 1980.

Gaylord-Ross, R.J., Weeks, M., & Lipner, C. An analysis of antecedent, response, and consequence events in the treatment of self-injurious behavior. Educa- and Training of the Mentally Retarded. 1980, 15. 35-42.

Goldfarb, W. Self-awareness in schizophrenic children. Archives of General Psychiatry, 1963, 8, 47-60,

Gottman, J.M. Analyzing for sequential connection and assessing interobserver reliability for the sequential analysis of observational data. Behavioral Assessment, 1980, 2, 361-368.

Griffiths, W. Behavioral difficulties of children as perceived and judged by parents, teachers, and children themselves. Minneapolis, University of Minnesota Press, 1952.

Harris, A. Response class: A Guttman scale analysis. Journal of Abnormal Child Psychology, 1980, 8, 213-220.

Hart, B., & Risley, T. In vivo language intervention: unanticipated general effects. Journal of Applied Behavior Analysis, 1980, 13, 407-432.

Hayes, S.C. Single case experimental design and empirical clinical practice. Journal of Consulting and Clinical Psychology, 1981, 49, 193-211.

Heads, T.B. Ethical and legal considerations in behavior therapy, In O. Marholm, II (Ed.) Child behavior therapy. New York: Gardner Press 1978.

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

Holvoet, J., Guess, O., Mulligan, M., & Brown, F. The Individualized Curriculum Sequencing model (II): A teaching strategy for severely handicapped students. Journal of the Association for the Severely Handicapped, 1980, 5, 337-351.

Horner, R.O., Holvoet, J.,& Rinne, T. Socialization Programs Module. Lawrence, KA: Personnel Training Program for Teachers of the Severely Handicapped, The Kansas Nuerological Institute, 1976.

Hutt, C., Forrest, S.J.,& Richer, J. Cardiac arrhythmia and behavior in autistic children. Acta Psychiatrica Scandinavica, 1975, 19, 361-372.

Hutt, S.J. & Hutt, C. Direct observation and measurement of behavior. Springfield, IL: Charles C. Thomas, 1970.

Hutt, S.J., Hutt, C., Lee, O.,& Ounsted, C. A behavioral and electroencephalographic study of autistic children. Journal of Psychiatric Research, 1965, 3 181-197.

Johnson, W.L. & Baumeister, A.A. Self-injurious behavior: a review and analysis of methodological details of published studies. Behavior Modification, 1978, 2 , 465-487.

Kara, A. & Wahler, R.G. Organizational features of a young child's behaviors. Journal of Experimental Child Psychology, 1977, 24, 24-39.

Kazdin, A.E. Sympton substitution, generalization, and response covariation: Implications for psychotherapy outcome. Psychological Bulletin, 1982, 91, 349-365.

Kissel, R.C. & Witman, T.L. An examination of the direct and generalized effects of a play-training and over-correction procedure upon the self-stimulatory behavior of a profoundly retarded boy. AAESPH Review, 1977, 2, 131-146.

Koegel, R.L. & Covert, L. The relationship of self-stimulation to learning in autistic children. Journal of Applied Behavior Analysis, 1972, 5, 381-387.

Koegel, R.L., Egel, A.L.,& Dunlap, G. Learning characteristics of autistic children. W. Sailor, B. Wilcox & L. Brown (Eds.) Methods of instruction for severely handicapped students. Baltimore: Paul H. Brookes, 1980.

Koegel, R.L., Firestone, P.B., Kramme, K.W. & Dunlap, G. Increasing spontaneous play by suppressing self-stimulation in autistic children. Journal of Applied Behavior Analysis, 1974, 7, 521-528.

Lathey, J.W. Assessing classroom environments and prioritizing goals for the severely retarded. Exceptional Children, 1978, 45, 190-195.

Lichstein, K.L, & Wahler, R.G. The ecological assessment of an autistic child. Journal of Abnormal Child Psychology, 1976, 4, 31-54.

Lovaas, O.I. The autistic child: language development through behavior modification. New York: Irvington Publishers, 1977.

Lovaas, O.I., Freitag, G., Gold, V.J., & Kassorla, I.C. Experimental studies in childhood schizophrenia: I. Analysis of self-destructive behavior. Journal of Experimental Child Psychology, 1965, 2, 67-84.

Lovaas, O.I., Koegel, R.L., Simmons, J.Q. & Long, J.S. Some generalization and follow-up measures on autistic children in behavior therapy. Journal of Applied Behavior Analysis, 1973, 6, 131-166.

Lovaas, O.I, Litrownik, A., & Mann, R. Response latencies to auditory stimuli in autistic children engaged in self-stimulatory behavior. Behaviour Research and Therapy, 1971, 9, 39-49.

MacCulloch, J.J., & Williams, C. On the nature of infantile autism. Acta Psychiatrica Scandinavica 1971, 47, 295-314.

Maher, C.A. Developing and implementing effective Individualized Education Programs for conduct-problem adolescents: The goal-oriented approach to learning. Child Behavior Therapy, 1981, 3, 1-11.

Maist, C.R., Baumeister, A.A., & Maist, A.A. An analysis of variables related to self-injurious behavior among institutionalized retarded persons. Journal of Mental Deficiency Research, 1977.

Marholin, D., Steinman, W.M., Luiselli, J.K., Schwartz, C, Townsend, N.M., Groden, J. & Baron, M.G. The effects of progressive muscle relaxation on the behavior of autistic adolescents: a preliminary analysis. Unpublished manuscript, 1978.

Meier, J.H. Screening, assessment and intervention for young children at developmental risk. In B. Z. Friedlander, G.M. Sterritt, & G. E. Kirk (Eds), Exceptional Infant, v. 3: Assessment and intervention. New York: Brunner/Mazel, 1975.

Nelson, R.O. & Hayes, S.C. Some current dimensions of behavioral assessment. Behavioral Assessment, 1979, 1, 1-16.

Ornitz, E.M. The modulation of sensory input and motor output in autistic children. In E. Schopler & R.J. Reichler (Eds.), Psychopathology and child development: research and treatment. New York: Plenum Press, 1976.

Page, E.B. Tests and decisions for the handicapped: A guide to evaluation under the new laws. Journal of Special Education, 1980, 14, 423-483.

Quay, H.C. Patterns of aggression, withdrawal, and inmaturity. In H.C. Quay & J. S. Werry (Eds.), Psychopathological disorders of childhood. New York: John Wiley, 1972.

Rapoport, A. & Wallsten, T.S. Individual decision behavior. Annual Review of Psychology, 1972, 23, 131-176.

Rincover, A., Cook, R., Peoples, A., Packard, D. Sensory extinction and sensory reinforcement principles for programming multiple adaptive behavior change. Journal of Applied Behavior Analysis, 1979, 12, 221-223.

Rincover, A., & Koegel, R.L. Setting generality and stimulus control in autistic children. Journal of Applied Behavior Analysis, 1975, 8, 235-246.

Risley, T.R. The effects and side-effects of punishing the autistic behaviors of a deviant child. Journal of Applied Behavior Analysis, 1968, 1, 21-34.

Rollings, J.P., Baumeister, A.A.,& Baumeister, A.A. The use of overcorrection procedures to eliminate the stereotyped behaviors of retarded individuals: an analysis of collateral behaviors and generalization of suppression effects. Behavior Modification, 1977, 1, 29-46.

Sailor, W. & Mix, B.J. The TARC assessment system. Lawrence, KA: H & H Enterprises, 1975.

Scheff, T. Being mentally ill. Chicago: Aldine, 1966.

Schopler, E. & Reichler, R.J. Psychoeducational Profile (PEP). Chapel Hill, NC: Child Development Products, 1976.

Schroeder, S.R., Mulick, J.A.,& Schroeder, C.S. Management of severe behavior problems of the retarded. In N.R. Ellis (Ed.), Handbook of Mental Deficiency, 2nd edition. New York: L. Erlbaum Associates, 1978.

Schroeder, S.R., Peterson, C.R., Solomon, L.J., & Artley, J.J. EMG feedback and the contingent restraint of self-injurious behavior among the severely retarded: two case illustrations. Behavior Therapy, 1977, 8, 738-741.

Schroeder, S.R., Schroeder, C.S., Smith, B., & Dalldorf, J. Prevelence of self-injurious behaviors in a large state facility for the retarded: A three-year follow-up. Journal of Autism and Childhood Schizophrenia 1978, 8, 261-270.

Schuler, A.L. Communicative intent and aberrant behavior. Paper presented at the Council for Exceptional Children's Topical Conference on the Seriously Emotionally Disturbed, Minneanolis, Minnesota, August, 1980.

Shavelson, R.J.,&Stern, P. Research on teachers' pedagogical thoughts, judgements, decisions, and behavior. Review of Educational Research, 1981, 51, 455-498.

Siegel, E. Task analysis and effective teaching. Journal of Learning Disabilities, 1972, 5, 11-15.

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

Smith, P.K. & Connolly, K. Patterns of play and social interaction in preschool children. N.G. Blurton Jones (Ed.), Ethological studies of child behavior. London: Cambridge University Press, 1972.

Sroufe, L.A., Stuecher, H.U. & Stutzer, W. The functional significance of autistic behaviors for the psychotic child. Journal of Abnormal Child Psychology, 1973, 1, 225-240.

Stein, Z. & Susser, M. Public health and mental retardation: new power and new problems. M.J. Begab & S.A. Richardson (Eds.), The mentally retarded and society: a social science perspective. Baltimore: University Park Press, 1975.

Stolz, S.B., and Associates. Ethical issues in behavior modification: Report of the American Psychological Association Commission. San Francisco, Jossey-Bass, 1978.

Strain, P.S., Cooke, T.P.,& Apolloni, T. Teaching exceptional children: assessing and modifying social behavior. New York: Academic Press, 1976.

Strain, P.S. & Ezzell, D. The sequence and distribution disordered adolescents' disruptive/inappropriate behaviors: an observational study in a residential setting. Behavior Modification, 1978, 2, 403-425.

University of Washington. Uniform Performance Assignment System (UPAS), Birth-6 year level. Seattle: University of Washington, College of Education, 1978.

Voeltz, L. M. & Evans, I.M. Covariation of behavior in a preschool handicapped child with multiple behavioral problems. Paper presented at the VI International Congress of the International Association for the Scientific Study of Mental Deficiency, Jerusalem, Israel, August 1979.

Voeltz, L.M., & Evans, I.M. Behavioral systems observation system. Honolulu University of Hawaii Behavioral Systems Intervention Project, (Technical Report #5) 1979 (b).

Voeltz, L.M., & Evans, I.M. The assessment of behavioral interrelationships in child behavior therapy. Behavioral Assessment, 1982, 4, 131-165.

Voeltz, L.M. & Evans, I.M. Educational Validity: Procedures to evaluate outcomes in programs for severely handicapped learners. Journal of the Association for the Severely Handicapped, in press.

Voeltz, L.M., Evans, I.M., Derer, K.R. & Hanshiro, R. Targeting excess behavior for change: A clinical decision model for selecting priority goals in educational contexts. Child and Family Behavior Therapy, in press.

Voeltz, L.M., Evans, I.M., Freedland, K. & Donellon, S. Teacher decision making in the selection of eduational priorities for severely handicapped children. Journal of Special Education, 1982, 16, 179-198.

Voeltz, L.M. & Wuerch, B.B. Monitoring multiple behavior effects of leisure activities training on severely handicapped adolescents. L.M. Voeltz, J.A. Apffel, & B.B. Wuerch (Eds.), Leisure activities training for severely handicapped students: Instructional and evaluation strategies. Honolulu: University of Hawaii Department of Special Education, 1981.

Wahler, R.G. Some structural aspects of deviant child behavior. Journal of Applied Behavior Analysis, 1975, 8, 27-42.

Wahler, R.G., & Fox, J.J. Solitary toy play and time out: A family treatment package for children withaggressive and opoositional behavior. Journal of Applied Behavior Analysis, 1980, 13, 23-29.

Walker, H.M. & Buckley, N.K. Programmed generalization and maintenance effects across time and across settings. Journal of Applied Behavior Analysis, 1972, 5, 209-224.

Wilcox, B., & Bellamy, G.T. Design of high school programs for severely handicapped students. Baltimore: Paul H. Brookes Publisher, 1982.

Willems, E.P. Behavioral technology and behavioral ecology. Journal of Applied Behavior Analysis, 1974, 7, 151-165.

Williams, C. Self-injury in children. Developmental Medicine and Child Neurology, 1974, 16, 88-90.

Wolery, M.R. Self-stimulatory behavior as a basis for devising reinforcers. AAESPH Review, 1978, 3:1, 23-29.

Yarrow, M.R. & Waxler, C.Z. Observing interaction: A confrontation with methodology. In R.B. Cairns (Ed.), The analysis of social interactions: Methods, issues, and illustrations. New York: Erlbaum, 1979.

Zlutnick, S., Mayville, W.J., & Moffat, S. Modification of seizure disorders: The interruption of behavior chains. Journal of Applied Behavior Analysis, 1975, 8, 1-12.

APPENDIX A

Behavioral Systems

Observation System

176

Observer 1 System:

Pupil Affect/Attention plus Excess Behaviors

BEHAVIORAL SYSTEMS INTERVENTION PROJECT

Behavioral Definitions
3/20/80 Revision

Observer #1: Pupil Affect/Attention plus Excess Behaviors

AFFECT (1st digit)

0 Neutral--no discernible expression on face.

1 Happy/Excitement--smiling, laughing, vocalization accompanied by broad smile.

2 Sad/Distress--pouting mouth drawn out and down, quivering of body, brows furrowed.

3 Anger--jaw set, facial features contorted, face may become flushed, body may tense, veins on neck and arm may protrude.

VISUAL REGARD/ATTENTION (2nd digit)

0 Space--not focusing on anything in particular (e.g., staring at blank wall) for two or more seconds.

1 Distant Object--eyes focused on object(s) located further than arm's reach for 2 seconds or more.

2 Proximal Object--eyes focused on object(s) within reach for 2 seconds or more.

3 Self (no object)--eyes focused on body or body part for 2 seconds or more.

4 Teacher--eyes focused on teacher for 2 seconds or more.

5 Non-Teacher Adult (includes observer)--eyes focused on adult(s) other than teacher for 2 seconds or more.

6 Peer/s--eyes focused on peer/s for 2 seconds or more.

7 Eyes closed (minimum of 1 second).

Observer #1 Definitions
3/20/80 Revision

EXCESS BEHAVIOR/VERBALIZATIONS (3 - 4th digits and following)

01 Body rocking (sitting, kneeling, crouching position)--moving in the trunk from hips or waist, rhythmically back and forth, or side to side.

02 Body rocking (standing)--moving entire trunk (may actually lift each foot slightly off the floor, alternately) or moving trunk from the hips or waist back and forth, or side to side, 2 or more cycles.

03 Spinning self (slowly)--slowly twirling self in a full (not rapidly enough to lose balance) circle, taking 2 seconds or more for one cycle.

04 Spinning self (rapidly)--rapidly twirling self in a full circle, taking less than 2 seconds for complete cycle.

05 Jerky movement--non-purposeful movement which interrupts previous "flow" of activity (includes loss of balance while sitting, walking, etc., may "catch" self or actually fall).

06 Body limpness and floor sprawling--child does not support body weight (limbs lose muscle tone, become "dead weight" when liften, i.e., "passive resistance"). Includes child sliding out of chair.

07 Leg/feet swinging (may include foot tapping, leg tapping against table leg, etc.)--swinging of leg at hip or knee and/or feet at ankle and/or tapping foot/feet on floor or against table leg.

08 Bear walking (on all fours)--feet and open palm on surface, bent at waist, propelling self using feet and hands, 2 or more cycles, or standing in position for 2 seconds or more.

09 Toe walking/running--walking/running 2 or more steps on balls of feet or toes, heels not touching floor.

10 Hand clapping--pounding or clapping of hands together.

11 Hand/arm flapping--moving hands/arms rapidly in "fluttering" motion with movement in wrist, elbow and/or shoulder.

12 Hands pressed on ears--hands or fingers pressed on ears, where impression of behavior is that child intends to block out sound as a functional purpose.

13 Finger (hand) flicking--repetitive finger (1 hand) movements, one or both hands, child may/may not watch hand/s.

14 Finger rubbing--rubbing of fingers and thumb (finger/s and thumb or fingers need to touch each other in back and forth motion for at least one incidence).

15 Finger tapping--touching surface or body part with finger/s or fingertip/s in a forcible fashion, 2 or more times.

16 Peering at object/person through finger/hand opening--peering at object/person through an opening formed by child's hand/fingers.

Observer #1 Definitions
3/20/80 Revision

17 Sky or table writing--movement of fingers as if writing or drawing in the air or on surface, generally with index finger.

18 Hair flicking/twisting (own)--fingers and/or hands flicking hair, more than one cycle.

19 Pulling/twisting clothing (own)--grasping own clothing, holding grasp for one second or longer; may twist clothing around fingers and/or hands.

20 Pulling/twisting ears--grasping and pulling at ears.

21 Rubbing eyes--repetitively rubbing eye/s with palm of hand/s or surface of fist of hand/s; includes eye rubbing with flexed fingers (not fingertips).

22 Pulling eyelash--grasping eyelash between thumb and fingers or between fingers and pulling away from eye (may either release eyelash with pull, or may actually pull eyelid away from eye slightly with pull).

23 Rubbing face, nose, mouth--rubbing fingers/hands across facial area (from ear to ear, and top of forehead to chin) in more than one cycle; or holding hand/finger/s in contact with face for more than 2 seconds.

24 Nosepicking--inserting finger or object into nostril, at least 1/8 inch deep; repeated insertion of finger or object; repeated scratching of nasal passage, or insertion of finger or object and repeated scratching of nasal passage.

25 Mouthing body parts--moving lips on and/or over a body part, inside or against mouth; body part must be visible or body part covered by clothing (e.g., foot in sock, shoulder in shirt, etc.).

26 Mouthing objects--moving lips on and/or over a (nonedible) object (even-though part of the hand may also be in contact with mouth/lips, code 26 if object is in mouth), inside or against mouth; object must be visible. Include mouthing of clothing which has been stretched from usual position and is being held in the mouth by the child (e.g., shirt wrapped around hand).

27 Hair pulling (own)--grasping and tugging at own hair.

28 Scratching/picking skin--scratching of body part (with fingers and/or deliberate intense rubbing of body part with hand).

29 Genital touch/masturbation--repeated touching/rubbing of genital organs or clothing in genital area; or holding hand/s in contact with genitals or clothing in genital area more than 3 or more incidents within a one minute time period.

30 Eye poking--poking around edges of eyes and/or pushing in on eyeball, using tips of finger/s or object.

31 Ear poking--inseting finger or object into ear, at least half inch deep, where intent does not appear to be blocking out sound.

Observer #1 Definitions
3/20/80 Revision

32 Biting self (attempting also)--biting or attempting to bite own body parts; may or may not break skin.

33 Pinching self--grasping fold of skin between fingers or thumb and finger/s, applying pressure to fold of skin.

34 Face slapping--slapping of face (ear to ear, forehead to chin area) with hand/s object

35 Head slapping--clapping of head with hands (fist or open hand/s) or object.

36 Head banging--banging of head against an objector person.

37 Body slapping (not head/hands)/body rubbing--slapping of hand/s, fist/s against body part other than head or face (e.g., slapping palms against thighs); rubbing of feed, fingers, hands or fists against body surface in back and forth or up and down motion; 2 or more cycles.

38 Choking self--grasping own neck with hand/s, and applying pressure to windpipe.

39 Digging--poking in rectal area; pulling at clothing in rectal area, two or more times or lasting longer than 2 seconds.

40 Grabbing at others (includes clothing)--grasping another person's body part/clothing in a forcible fashion, holding grasp for one second or longer.

41 Grabbing object (attempting to)--grabbing objects other than appropriate manipulation of instructional materials during a task or manipulation of a t during free play. Includes taking materials the teacher has indicated sho ld not be taken and taking a peer's toy away during free play.

42 Pulling other's hair--grasping, tugging at another's hair.

43 Hitting other/s--hitting another person/s with body part (arm, hand, foot) or an object being held.

44 Pulling/pushing other--pushing a person by laying hand/s on his body and forcibly applying pressure; person may/may not move; grasping another's body part/clothing and applying force toward the source of the force, person may/may not move.

45 Pinching other--grasping fold of skin of another person between fingers or thumb and finger/s, applying pressure to fold of skin.

46 Choking other--grasping another person's neck with hand/s, applying pressure to windpipe.

47 Smelling other--bringing body part/person within 2 inches of nose or moving nose within 2 inches of an object.

48 Mouthing other--moving lips on and/or over another/s body part inside or against mouth; body part must be visible.

49 Biting other (attempting to)--biting or attempting to bite body parts of another person; may or may not break skin.

50 Shadow play--movements of object while observing its shadow for more than one movement cycle or for longer than 3 seconds; may involve moving head while watching shadow effects (e.g., moving position of head and watching chair rung shadow movement).

51 Smelling object--bringing object within 2 inches of nose or moving nose to within 2 inches of an object.

52 Object banging--grasping and banging an object against a body part, surface or another object.

53 Object tapping--tapping of object/s against a surface.

54 Object flicking--manipulating an object rapidly in "fluttering" motion, movement in wrists/elbows/finger joints.

55 Object spinning--acting upon an object so that it moves in a circular motion for one or more cycles.

56 Object dropping--grasping and releasing of object from a distance above a surface.

57 Throwing objects--grasping and forcibly releasing an object causing it to rapidly move through space.

58 Sweeping object/s off surface--bringing arm/hand across surface swiftly pushing object/s to the side of or off surface onto floor.

59 Tipping over furniture--acting upon a piece of furniture causing it to fall over.

60 Cry, no tears--whining, moaning, sniffling for 5 seconds or more or 2 or more times during a 1 minute period.

61 Cry, tears--whining, moaning, sniffling for 5 seconds or more or 2 or more times during a one minute period. Must include welling of tears tearing.

62 Shriek/scream--sharp, shrill, loud cry (above conversational level); may be abrupt with short duration or prolonged, continuous duration.

63 Grunt--low volume subglottal/glottal sound produced when air forced out of mouth and/or nose.

64 Blowing--blowing air out of mouth with/without saliva and/or tongue protrusion.

65 Vocalization, subglottal and glottal (includes humming)--vocalization produced mainly in the throat with no visible tongue, lip, teeth movements involved; most will be vowel sounds ("aa", "oo", "hmm", etc.) and single-syllable; not identifiable words.

182

66 Vocalization, supraglottal--vocalization involving visible movement/contact with tongue and teeth or lips, lips and teeth, lips, etc.; most will be consonant plus vowel sounds ("ba-ba", ma-ma", etc,) and may be multiple syllables; not identifiable words.

67 Clicking vocalization (tongue/lips)--loud, popping sound when tongue quickly drawn from roof of mouth and/or clicking sould when tongue quickly drawn from teeth to back of mouth (tongue movement need not be visible); or click-like sound produced by expulsion of air with friction between side tongue surface and teeth.

68 Echoed verbalization (immediate)--repeats wording said by instructor, peers, etc,, during present observation period.

69 Spontaneous verbalization--identifiable word/s spontaneously uttered by the child (not imitations of peer/teacher verbalization).

70 Facial grimace (includes squinting)--contortion of facial features, including squinting, furrowing eyebrows, wrinkling nose, drawing corners of mouth out and down; look may appear suddenly "excited,"

71 Eye crossing/rolling--bringing pupils of eyes to inside corners of eyes and/or moving pupils of eyes on sockets in different directions (in more than one cycle).

72 Staring/gazing--holding a fixed, glassy-eyed look for more than 3 seconds.

73 Saliva swishing (drooling)--audibly swishing saliva in mouth, and/or visible saliva, visibly escaping from corner/s of mouth.

74 Rumination--regurgitating food into mouth; may either re-swallow material or actually expell material from mouth.

75 Teeth grinding/clicking--audibly grinding and/or rapidly closing teeth together (mouth may be shut tightly or set in a grimace).

76 Chewing/sucking hair--placing of own hair into or against mouth.

77 Pica--eating nonedibles, chewing (one or more chewing movements) and/or swallowing noneoible materials; object must have been visible (e.g., bits of paper, material from floor, etc.) must be small enough to swallow.

78 Head weaving/shaking--movement of head in a side to side or front and back motion for 2 or more cycles; or pronounced, full circle motion or figure eight pattern motion of head, one cycle or more.

79 Head dropping--head drops abruptly forward or backwards, chin toward chest or neck stretched, facing ceiling.

80 Gesture/sign--specific body movement representing a word; either familiar gesture (e.g., pointing to toilet) or trained symbolic gesture,

81 Sign and vocalization--all vocalizations which are simultaneously paired with a sign or gesture.

82 Lint picking--grasping piece/s of lint from a surface (e.g., table top, floor, clothing), one or more cycles.

83 Smearing--manipulation and spreading of feces.

84 Temper tantrums--thrashing arms and legs, loud screaming and/or crying, accompanied by thrashing body movements.

85 Breaking object--manipulating an object causing it to crack/break into pieces.

86 Tongue movements--moving tongue around mouth and/or moving or holding tongue over lips for 3 seconds or more or 2 or more times during a 1 minute time period, not connected with either functional vocalization or eating, drinking; chewing-like motion of mouth (tongue may or may not be visible) not connected with eating, vocalization or drinking.

87 Tearing--manual ripping of a material (e.g., paper).

88 Bolting--physically moving away from control, authority of another person (e.g., running away from teacher while being led from one task to another, leaving chair during instruction, etc.), or a designated area where student should remain.

89 Pounding on object/surface (include body throw against object/surface)--forcible hitting a large object or surface with body part other than fingertips (e.g., hitting a table with fist or open palm, kicking a table leg, etc.) Includes body throwing against an object/surface (e.g., wall, table, door, etc.)

90 Stripping--complete removal of clothing, or definite attempt to at least partially remove clothing (e.g., pulling shirt with neck opening up to armpits with chest showing, opening a front-closing shirt, pulling pants down to a point where buttocks are visible, unzipping pants, pulling off shoes or socks, etc.)

91 Jumping/hopping--rapid up and down, verticle motion; in standing positon, entire body moves. feet may or may not actually leave floor surface. In sitting position, involves verticle jiggling--type movement of head and torso.

92 Object rubbing--rubbing object/surface with fingers or hands for 2 or more cycles or twice in a 30 second time period, e.g., fingertips in back and forth motion across toy surface, palms of hands rubbed side to side, back and forth on rug or floor surface

93 Finger/hand posture/finger touch--posturing of hand/s and/or fingers in rigid position for one second or more; restricted to finger/hand posturing or may also involve rigid positioning of arms shoulders, back, etc.

94 Climbing--physically pulling self on or over furniture and/or equipment.

Observer #1 Definitions
3/20/80 Revision

Notes:

1. When any excess behavior occurs as described by the definitons, enter the appropriate code regardless of the following conditions:
   a) Teacher instructs, e.g., "Clap your hands."
   b) Teacher shapes, e.g., teacher physically guides the child to flap his/her hands.
   c) Teacher approves, e.g., "That's good talking" in response to screaming.
   d) Behavior seems appropriate, e.g., spinning a top, "mouthing" balloon in order to blow it up.

2. Code the following behaviors last when entering multiple behaviors.
   a) vocalizations and verbalizations
   b) gestures, signs
   c) saliva swishing/drooling
   d) leg swinging
   e) sign and vocalization

3. When behaviors rapidly follow one another, with less than one second between behaviors, they are to be coded as one sentence. When there is more than one second between behaviors the break button should be pushed.

Observer 2 System:

Pupil General plus Environment

186

BEHAVIORAL SYSTEMS INTERVENTION PROJECT

Behavioral Definitions

3/20/80 Revision

Observer #2: Pupil General Plus Environment

POSITION IN SPACE (1st digit)

0 Lying--reclining posture with both trunk and limbs making contact with substrate

1 Sitting--position in which body rests primarily on buttocks

2 Crouching/kneeling--
Crouching--position in which body rests primarily on the feet with the knees and hips flexed and legs drawn close to the body
Kneeling--position in which body rests primarily on the knees
(includes crawling position without movement)

3 Standing--position in which the body rests primarily on one or both feet, legs relatively straight and perpendicular to the substrate

4 Crawling--movement forward/backward on hands and knees propelled by the limbs (2 cycles or more) (includes scooting on buttocks)

5 Walking--moving the body forward/backward at a moderate pace, placing one foot or knee on the floor before lifting the other. Includes walking in a circle.

6 Running, jumping, hopping
Running--moving the body forward at a rapid pace with both feet off the ground simultaneously during each stride (2 cycles)
Loco-jump/hopping--moving the body suddenly upward or horizontally by leg and foot extension, landing on one or two feet

CHILD RESPONSE (2nd digit)

0 Undifferentiated/Neutral--no response; no engagement in any type of behavior

1 Excess behavior--engagement in any of the excess behaviors listed for that child (note: does not include vocalization behaviors)

2 Excess + On-Task Appropriate--engaging in any of the listed excess behaviors (excluding vocalizations) while simultaneously focusing eyes on task and/or materials and/or teacher for at least a 3 second period; attempting to perform or performance of task required (e.g., swinging leg under the table while giving the requested shape to the teacher during an instructional task)

3 On-Task Appropriate--eyes focused on task and/or materials and/or teacher for at least a 3 second period; attempt to perform or performance of task required (includes incorrect and approximate responses, i.e., matching red to green)

4 Off-Task Appropriate--appropriate play or manipulation of task materials but no attempt to perform specified task objective (e.g., child required to throw ball to instructor but instead bounces and catches ball repeatedly)

Observer #2 Definitions
3/20/80 Revision

OBJECTS IN VICINITY* (3rd digit)

0 None--no visible extraneous objects present (toys, equipment) other than necessary, usual furniture such as desk, chairs, etc.

1 Program material/s--materials, toys equipment, etc., which the teacher has presented in the situation for instruction during the present period with that child and/or peer/s

2 Toy/s--objects or equipment child can manipulate during free time; not used for instructional purposes during persent period, may be identified as "toy" as a child-selected item (e.g., piece of string being flicked)

3 Food/Drink--any edibles which are visible to or in contact with the child (e.g., in child's hand or mouth; in cup on table in front of child)

4 Program material/s + Food/Drink--presence of both materials (see code 1) and any edibles (see code 3)

5 Toy/s + Food/Drink--presence of both objects (see code 2) and any edibles (see code 3)

6 Program material/s + Toy/s--presence of teacher selected program materials, (see code 1 above) and object/toy (see code 2 above)

7 Program material/s + Toy/s + Food/Drink--presence of teacher selected program material/s (see code 1 above), toy (see code 2 above) and any edibles (see code 3 above)

TEACHER RESPONSE (4th digit)

0 N/A (not present)--teacher not in vicinity

1 Neutral--no visible, audible or discernible "affect"/expression by the teacher

2 Neutral plus Contact--no audible, discernible expression by teacher while physically touching child

3 Neutral Mand--instructional verbal command and/or gestures given by the teacher with no significantly audible or observable positive/negative expression, including an instructional 'latency" (waiting for the child to respond) of approximately 3 seconds if teacher behavior does not change

4 Neutral Mand + Contact--neutral mand (see code 3) while physically touching child

5 Approval--facial and body gestures which indicate approval/positive affect (e.g., smiling, clapping); may or may not include verbal positive social reinforcement (e.g., words like "Good!" "Nice work!" or vocalization, e.g., "mmm" as child eats primary reinforcer", but after such a verbalization a new code is entered only if the teacher's affect changes--not merely because teacher is no longer talking--or if code 6 occurs

---

*"Vicinity" should be defined as within child's immediate sphere of influence and/or within the child's reach without a major change of body position. In most cases, an object in "vicinity" would be in view within arm's reach of the child; exception: child is seated across from another child engaged in ball play--ball considered in vicinity as long as it remains in play between the two children

6 Approval + Contact--teacher approval (see code 5 above) while physically touching child (e.g., pats child's head, kisses, hugs child, etc.)

7 Disapproval--facial and body gestures which indicate disapproval/negative affect (e.g., frowning, shaking head, etc.). May or may not include negative verbalization (e.g., "No, stop it!", "No biting!") or vocalization (e.g., "ouch!"), but after such a verbalization/vocalization a new code is entered only if the teacher's affect changes--not merely because teacher is no longer talking--or if code 8 occurs

8 Disapproval + Contact--teacher disapproval (see code 7) while physically touching child (e.g., physical restraint, putting child's hands down, etc.)

9 Attending to peer/ignoring--all of attention is focused on peer; or may be deliberate non-responding, refusal to attend to, avoid attending to, avoid interacting (verbal and physical) with target child (e.g., teacher turns away from child for 10 seconds in response to excess behavior)

PEER RESPONSE (5th digit)

0 N/A (not present)--peer not in vicinity

1 Neutral--no visible or audible response to subject's presence; no attempt to interact with subject

2 Approach--looks at subject for at least 3 seconds; (attempts to) verbally and/or physically interact with subject. May vocalize, smile while looking at subject, reach out and (attempt to) touch subject, move to decrease the distance between peer and subject

3 Avoid--increases the distance between himself and subject; physically moves away from subject, refusal to interact verbally/physically with subject (e.g., pulling arm away in response to touch by subject).

4 Aggression--makes physical contact/verbal action which amy result in injury to subject or intent to hurt subject (hitting, shoving, pinching, scratching, verbal taunts, swearing, etc.)

5 Protest minus Avoid/Aggression--peer exhibits distress, upset by obvious crying or other vocalization, facial expression. etc., in response to target child's behavior, but does not actually physically move away or pull away from target child (which is code 3) or strike back at target child (which is code 4)

NOTES

1 The following definitions should be referred to when determining On Task Appropriate verses Off Task Appropriate play behavior.

	On-Task Appropriate	Off-Task Aopropriate
Doll	Cradles, rocks, hugs, strokes hair, pats back with doll on shoulder, pretends to feed, bathe or dress/ undress.	Holds in a fashion other than as a baby, e.g. upside down by feet, pulls at body parts, rolls doll on floor, etc.
Pull-toy	Pulls toy by the string causing wheels to turn on floor surface, holds/pushes train and rolls it along surface.	Bobs toy up and down by pulling string upwards, pushes toy with wheels not in contact with floor, turns toy over and over, etc.
Busy box	Operates or attempts to operate any of the various devices, e.g. turn the dial, press the lever, etc.	Makes physical contact with the material but doesn't attempt to operate the various devices.
Ball	Rolls, tosses, catches, bounces or kicks the ball.	Holds ball.
Bolster/ Pillow	Sits on, lies, over/on, bounces on, hops on, jumps over.	Physical contact without any conscious attempts to use, e.g. happens to be learning against it.
Stacking rings	Puts rings on or takes rings off column in any order.	Manipulates rings in some fashion e.g. transfers from hand to hand, picks up and holds, squeezes.

2 When Off-Task Appropriate behavior occurs wit any excess behavior, c.de only excess.

3 Determine child response only on the basis of the task command and the child's response, not on the basis of the teacher response. E.g. teacher says, "match the cards", child picks up card while looking at it and matches to a sample, teacher says "no, it's this one". This would be coded as 3, On-Task Appropriate since the child attempted to complete the desired response. If the child had taken a pair of scissors and cut the cards, it would be coded as 4, Off-Task Appropriate.

4 If a behavior changes rapidly (e.g. teacher approval, disapproval, approval) with less than one second between behaviors, the first behavior coded can remain. Do not begin a new sentence on the basis of fleeting changes in behavior.