Hammill Institute on Disabilities

Behavioral Disorders
2018, Vol. 43(2) 322–336
© Hammill Institute on Disabilities 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0198742917753157
journals.sagepub.com/home/bhd
SAGE

Republication of "Functional Analysis of Classroom Variables for Students with Emotional and Behavioral Disorders"

Glen Dunlap[1], Lee Kern[2], Maria dePerczel[1], Shelley Clarke[1], Diane Wilson[1], Karen E. Childs[1], Ronnie White[1], and George D. Falk[3]

Editors' Note

This is the third in a series of three classic article reprints celebrating the rich history of impactful scholarship published in Behavioral Disorders in the 1970s, 1980s, and 1990s. For each classic article in the series, original authors provide an accompanying commentary discussing the original manuscript, its implications, and developments in the area since publication of the article (see Algozzine, 2017*, and Elliott, 2017**, for commentaries on the classic articles from the 1970s and 1980s, respectively). In this issue, we reprint the classic article by Dunlap and colleagues (1993) on the then emerging practice of functional assessment for students with emotional and behavioral disorders. The paper describes a comprehensive functional assessment process that became a model for research, policy, and practice. We thank the two lead authors of this paper, Glen Dunlap and Lee Kern, for their insightful commentary, which provides a context and implications for the original study, and discusses future directions in this area.

*Algozzine, B. (2017). Toward an acceptable definition of emotional disturbance: Waiting for the change. *Behavioral Disorders, 42*, 136-144.
**Elliott, S. N. (2017). The social validity of "Acceptability of behavioral interventions used in classrooms": Inferences from longitudinal evidence. *Behavioral Disorders, 43*, 269-273.

Abstract

Functional assessment and functional analysis are processes that have been applied successfully in work with people who have developmental disabilities, but they have been used rarely with students who experience emotional or behavioral disorders. In the present study, five students in elementary school programs for severe emotional disturbance participated in a comprehensive functional assessment process designed to yield a useful understanding of their desirable and undesirable behaviors. Interviews, record reviews, and direct classroom observations led to the development of individualized hypotheses regarding relationships between classroom events and the occurrence of target behaviors. Subsequently, direct manipulations (i.e., functional analyses) were conducted to test each of the hypotheses in the context of regularly-occurring classroom activities. These analyses demonstrate important influences that certain classroom variables can exert over individual student's behavior. The process and results are discussed with regard to the need for improved methods for understanding student responding, and the benefits that functional assessment can offer programs of educational and behavioral support.

Over the past decade, procedures for managing desirable and undesirable behaviors have been refined through the development of individualized functional strategies of assessment and analysis. A growing literature has shown that careful pre-intervention assessments can lead to proactive positive programs that are effective in reducing problematic responding and that avoid reliance on punitive interactions (e.g., Carr, Robinson, & Palumbo, 1990). A number of authors have underscored the importance of the assessment process and have stated that the effectiveness of positive interventions can be dependent upon the validity of these pre-intervention assessment data (Carr et al., 1990; Horner et al., 1990). Indeed, the processes of functional assessment and functional analysis have emerged as significant

[1]Department of Child and Family Studies, Florida Mental Health Institute, University of South Florida, Tampa, FL, USA
[2]University of Pennsylvania School of Medicine, Children's Seashore House, Biobehavioral Unit, Philadelphia, PA, USA
[3]Department of Special Education, University of South Florida, Tampa, FL, USA

Received 29 January 1993; Initial acceptance 22 April 1993; Final acceptance 10 May 1993

topics of research and application, especially with regard to serious problem behaviors dis played by people with severe disabilities (Iwata, Dorsey, Slifer, Bauman, & Richman, 1982; Lennox & Miltenberger, 1989).

Although the term has been used in various ways, *functional* assessment can be defined as a process of identifying functional relationships between environmental events and the occurrence and nonoccurrence of a target behavior. Thus, a functional assessment consists of the methods and procedures that are used to identify associations between the behavior and variables in the environment. These procedures include any techniques that help to identify these relationships such as rating scales (Durand & Crimmins, 1988), interviews (Bailey & Pyles, 1989; O'Neill, Horner, Albin, Storey, & Sprague, 1990), and direct observations (Mace & Lalli, 1991; Mace, Yankanich, & West, 1989; O'Neill et al., 1990; Touchette, MacDonald, & Langer, 1985). A principal objective of functional assessment is to derive clear hypotheses about the relationship between the environment and the behavior of interest (Repp, Felce, & Barton, 1988). Typically, the relationships that are identified can be categorized as involving either: (a) the operations of a reinforcement contingency including positive reinforcement (e.g., the undesirable behavior is reinforced by the contingent attention provided by specific classroom personnel) and negative reinforcement (e.g., the undesirable behavior is reinforced by the removal of, and thus the escape from, an unpleasant task requirement); or (b) a controlling antecedent stimulus (e.g., the undesirable behavior occurs during a specific instructional activity or when an unpleasant lesson is presented for an excessive period of time).

Functional analysis is considered to be an important component of the functional assessment process in which the identified variables are directly manipulated in order to verify or clarify the hypothesized relationships. The direct manipulations are conducted in the context of reversal or alternating treatments designs and have been referred to as experimental analyses (Iwata, Vollmer, & Zarcone, 1990) or hypothesis testing (Dunlap & Kern, in press; Repp et al., 1988). Although many of the functional analyses that have been reported in the literature have been conducted in clinical or other analog circumstances, there are examples that have been conducted in classrooms (e.g., Cooper et al., 1992) including some that have been conducted by classroom teachers (e.g., Lalli, Browder, Mace, & Brown, in press). To date, the vast majority of research on functional assessment and functional analysis has been conducted with individuals who have developmental disabilities. In contrast, very little data attest to the efficacy of functional assessment with individuals whose primary dis abilities are described as emotional or behavioral rather than intellectual. However, a few relevant illustrations recently have been reported. Dunlap, Kern-Dunlap, Clarke, and Robbins (1991) described a detailed case report that involved a 12-year-old girl whose multiple disabilities included emotional challenges, and who was enrolled in a classroom for students with severe emotional disturbance. These authors conducted an extensive process of functional assessment that yielded four hypotheses relating features of the girl's educational curriculum to the occurrence of her serious problem behaviors. In 4 days of hypothesis testing (functional analysis) in the girl's classroom environment, the hypotheses were validated and refined. The subsequent intervention involved a substantial revision of the girl's curriculum and produced considerable increases in desirable behavior and a virtual elimination of behavior problems.

In a more recent example (Kern, Childs, Dunlap, Clarke, & Falk, 1993), an 11-year-old boy, who was described as emotionally disturbed, participated in a process of functional assessment designed to identify classroom variables related to his excessive crying and self-injury. In this process, multiple methods of information gathering were used including an interview with the student himself. Five hypotheses were developed and tested during the student's ongoing classroom lessons in three academic subjects. The functional analyses served to confirm the hypotheses and led to interventions that were used successfully in each of his academic classes.

Although these examples are encouraging, there is still very little evidence that the processes of functional assessment and functional analysis are feasible and informative in applied settings with students identified as having emotional and behavioral disorders. This is a key deficiency in the literature because students with emotional and behavioral disorders are distinguished from students with other (e.g., developmental) disabilities in ways that could have a distinct impact on the assessment process. For example, students with emotional and behavioral disorders generally function within the normal range of intelligence and their challenges may be largely in areas such as interpersonal relations, emotionality, and control over inappropriate behaviors.

The purpose of this investigation was to evaluate the applicability of functional assessment and functional analysis with students described as having emotional and behavioral disorders. To do this, the authors developed individualized hypotheses and conducted systematic functional analyses with several students. Furthermore, we were interested in deter mining the feasibility and validity of these assessment strategies in the complex environments of special education classrooms. Therefore, all of the assessment procedures were developed and conducted within the context of the students' ongoing educational programs.

General Method

Participants and Setting

Serving as participants in this study were 4 boys and 1 girl. These students were referred by their classroom teachers because they exhibited ongoing behavior challenges that

Table 1. Student Characteristics.

Name	Age (years)	Referral Behaviors	Intellectual Test Scores	Grade
Desi	10	Inappropriate verbal behavior toward staff (arguing, whining, threatening statements)	WISC-R Full Scale-93 Verbal-90 Perform-100	4th
Garth	11	Inappropriate verbal behavior toward staff and peers. Difficulty accepting consequences or instructions. Non-compliant behavior and property destruction	WISC-R Full Scale-94 Verbal-95 Perform-95	4th
Shaun	11	Off-task behavior. Easily frustrated. Crying, whining, and self-abusive behavior	Stanford-Binet Composite-98	5th
Sammy	10	Off-task behavior. Making noises, inappropriate gestures, and "silly" behavior in the classroom	WISC-R Full Scale-101 Verbal-90 Perform-114	4th
Ann	6	Off-task behavior. Leaving assigned area, talking out, and inappropriate staff interactions (ordering, yelling)	Stanford-Binet Composite-83	K

had been resistant to classroom-wide and individualized programs of behavior management. Desi, Garth, and Sammy were all 10 or 11 years of age and enrolled in the 4th grade; Shaun was 11 years old and enrolled in the 5th grade; Ann was 6 years old and enrolled in kindergarten.

All of the participants were identified as having behavioral and emotional challenges by the public school system and were enrolled in self-contained classrooms serving students with severe emotional disturbance (SEO). None of the students were receiving medication at the time the investigation was conducted. Additional information about the participants and their major presenting problems are listed in Table 1.

The research took place at two elementary schools that include programs for students with severe emotional disturbance. The special education programs at each school consist of four classrooms, all of which are staffed by a teacher and an aide. The schools are among those in a large school district that are collaborating with the authors in an ongoing program of applied research. Thus, the classrooms were also occupied frequently by one or two staff from the research program who served as data collectors and consultants. Of the participating students, 3 (Desi, Garth, and Sammy) were enrolled in one school and the other 2 students attended the second school. The assessments in this study focused on academic sessions held each day in the students' special education classrooms.

Dependent Variables

After referral, the students' teachers and the consultants identified desirable and undesirable behaviors that served as the focus of the assessments. The undesirable behaviors were selected because they occurred with sufficient frequency to be ongoing concerns in the classrooms and because they had not been managed successfully with previous interventions. Desirable behaviors were those that the teachers identified as particularly important to increase. The specific desirable and undesirable behaviors that were identified for each participant are listed in Table 2.

Functional Assessment

The functional assessment process was conducted in two phases. In the first phase, hypothesis development, information was gathered from a variety of sources to arrive at specific hypothesis statements. These statements were intended to describe a relationship between a participant's behavior and a variable that could be manipulated within the classroom environment. In the second phase of the assessment process, the hypothesis statements were tested empirically through procedures of functional analysis. Throughout the assessment process, the teachers and consultants worked collaboratively within the ongoing context of the classroom activities.

Phase I: Hypothesis Development

The first phase in the functional assessment process was the development of individualized hypotheses relating classroom events to desirable, and in particular, undesirable responding. Data for each participant were collected from a variety of sources including standardized instruments and archival records; discussions and detailed interviews with teachers, students, and other classroom staff; and direct observations.

Table 2. Dependent Variables and Recording Procedures for the Five Participating Children.

Name	Undesirable Behavior	Desirable Behavior	Measurement
Desi	No response or negative verbal or nonverbal response to adult initiation; noise making or off-task statements in academic context	Verbal or nonverbal response or initiation that is positive or neutral and appropriate to the task or social context	Frequency recording, divided into 1 minute intervals
Garth	No response or negative verbal or nonverbal response to adult initiation; noise-making or off-task statements in academic context; off task behavior during academic activities	Verbal or nonverbal response or initiation that is positive or neutral and appropriate to the task or social context; on-task responding during academic activities	15-second continuous interval recording
Shaun	Whining, crying, grunting, and other noise making and uninterpretable, garbled vocalizations	Appropriate verbalizations (requests, statements, etc.) in a normal tone of voice	Frequency recording divided into 1 minute intervals
Sammy	Off-task behavior during academic activities; no response or negative response or initiation to peers; noise making or inappropriate talking during academic activities	On-task behavior during academic activities	15-second partial interval recording, with 10 seconds of observation and 5 seconds of data recording
Ann	Running out of class room without permission; hitting, attempting to hit or throwing objects at another individual in classroom; property destruction; leaving contact with designated chair for at least 3 sec onds; failure to comply with instruction within 5 seconds	On-task behavior during academic activities	15 seconds partial interval recording, with 10 seconds of observation and 5 seconds of data recording

An interview protocol (available from the authors) was used for each student's assessment. This protocol was based extensively on the Functional Analysis Interview Form developed by O'Neill et al. (1990) and included a variety of questions designed to reveal student preferences and specific curricular and interactional variables that might consistently influence the student's behavior. Several of the questions asked respondents to identify specific classroom circumstances that were associated with the student's desirable and undesirable behavior.

Direct observations were collected for each student in the classroom context that was described by the teachers as the most problematic. In addition to documenting the rate of desirable and undesirable behaviors, the observations provided information regarding events that occurred as antecedents and consequences of undesirable behavior (Bijou, Peterson, & Ault, 1968). These observations also included records of the student's behavior in the context of teacher-student interactions. For example, after interviews suggested that Desi's undesirable behavior could be reduced with high rates of praise, direct observations were obtained on Desi's undesirable behavior as well as the praise statements that were delivered by Desi's teachers during naturally-occurring classroom interactions (cf. Lalli et al., in press). These observations suggested that undesirable behaviors were indeed reduced, but only when the praise statements described specific actions or accomplishments. Thus, the direct observations helped to confirm and refine the information from the interviews.

As data were collected, the consultants met with the classroom staff in order to formulate hypothesis statements describing functional relationships between classroom variables and the student's behavior. Criteria for hypotheses were that the statements had to (a) be based on information from the interviews and direct observations, and (b) identify specific variables that could be manipulated by the student's teachers within the classroom context. In addition, both the consultants and the teachers had to agree that the statements represented reasonable judgments from the accumulated information.

Two or three hypotheses were developed for each student. These hypotheses are presented in Table 3. As can be seen, the hypotheses differed across students and identified a range of variables that include antecedent and curricular events as well as consequences such as attention and praise.

Phase II: Functional Analysis (Hypothesis Testing)

General approach and design. In this phase of the experiment, the hypotheses that were generated in Phase I were tested empirically ·by manipulating the identified variables with in the ongoing context of the classroom routine. Each hypothesis for each participant was tested with a reversal design accompanied by direct observation of the dependent variables (Table 2). In general, the hypothesis testing was accomplished by conducting one session of one condition (e.g., the condition that typically occurred within the classroom and that was associated with relatively high levels of undesirable behavior) on one day and then changing the

Table 3. Hypothesis Statements for Each of the Five Participating Children.

Name	Hypothesis Statements
Desi	Desi's undesirable behavior will be reduced when: 1. the amount of specific praise is increased. 2. he evaluates the appropriateness of his verbalizations. 3. staff ignore all undesirable behavior.
Garth	Garth's undesirable behavior will be reduced when: 1. the amount of attention from staff is increased. 2. he evaluates the appropriateness of his verbalizations. 3. he is given a choice of academic tasks.
Shaun	Shaun's undesirable behavior will be reduced when: 1. long tasks are divided into portions. 2. school staff maintain close physical proximity.
Sammy	Sammy's desirable behavior will be increased when: 1. he evaluates his task engagement and ignores the activities of others. 2. he is not in close physical proximity to other students. 3. he is engaged in purposeful activities that result in a useful permanent product.
Ann	Ann's desirable behavior will increase when: 1. she is offered a choice of academic tasks. 2. she receives a high rate of adult attention and praise for desirable behavior. 3. she is engaged in a preferred academic activity.

condition (e.g., to the condition that was hypothesized to produce low levels of undesirable behavior) on the following day. This sequence was then repeated at least once for each hypothesis yielding, at a minimum, a comparison across three experimental phases. Although the majority of the analyses were conducted by alternating conditions on successive days, some exceptions occurred due to teacher preference or time constraints. The first hypothesis for Shaun, for example, was conducted by continuing each condition for a full week, while hypothesis 1 for Garth and hypotheses 1 and 3 for Ann were tested by implementing two or three conditions per day. In all cases, the sequence in which the conditions were presented was alternated to control for the possibility of order effects.

All of the manipulations were conducted in the context of the ongoing classroom routine. Consultants were present to collect data and to assist in conducting the manipulations. The testing was conducted during sessions that the teachers had identified as being problematic. The length of the sessions was not changed from the typical class room routine and usually ranged from 25 to 35 minutes. Data were collected during a consistent portion of the session for 15 or 20 minutes.

Procedures for testing hypotheses. The manner with which the hypothesis testing (see Table 3) was conducted for each participant is described below:

Desi. Throughout these manipulations Desi was assigned independent textbook activities that were similar in response requirements across each of the hypotheses tested. During the testing of hypothesis 1 (high specific praise vs. low specific praise), high specific praise conditions were those in which Desi received praise for specific desirable behaviors at least once every 2 minutes (based on observations during the hypothesis development phase). Low praise conditions were those in which staff gave the student specific praise less frequently than once every 5 minutes.

During the testing of hypothesis 2 (self-monitoring vs. no self-monitoring), self-monitoring conditions were those in which a tape recorder sounded a bell every minute, and at the time the bell sounded, Desi recorded whether he had positive interactions (e.g., not arguing, speaking at appropriate times) during the preceding interval by checking either Yes or No on a sheet of paper at his desk. The accuracy of the student's recording was evaluated by having a data recorder score the same behaviors, and feedback regarding accuracy was given to the student following each session. No self-monitoring conditions were those in which the student did not use the self-monitoring procedures.

During the testing of hypothesis 3 (ignoring vs. no ignoring), ignoring conditions were those in which staff ignored all undesirable behaviors exhibited by the student. No ignoring conditions were those in which staff responded to the majority of undesirable behaviors with some form of attention (e.g., redirection, mild reprimand).

Garth. During the testing of hypotheses 1 and 2, Garth was engaged in independent work involving typical textbook and worksheet tasks or group activities. The tasks were similar across conditions in terms of the response requirements. For the testing of hypothesis 1 (high attention vs. low attention), high attention conditions were those in which Garth received from the staff one or more positive or neutral statements per minute throughout the session. This rate was based on observations during the hypothesis development phase. Low attention conditions were those in which he received less than one positive or neutral interaction every 2 minutes. The testing of hypothesis 2 (self-monitoring vs. no self-monitoring) was conducted in the same manner as described for Desi.

Hypothesis 3 (choice vs. no choice) was tested in spelling. Choice conditions were those in which Garth was given a menu of 5 different spelling tasks (e.g., writing the word 3 times; writing definitions for 20 words) and was allowed to choose which task he would complete that day. In addition, he was provided the option of changing tasks at any time during the session (cf. Dyer, Dunlap, & Winterling, 1990). During the no choice conditions, the assignments were the same as in the choice sessions but the assignments were selected by the teacher and changing tasks was not permitted.

Shaun. Hypothesis 1 (short vs. long) was tested during reading while Shaun completed worksheets. During short conditions, Shaun received a series of 2-3 worksheets with each worksheet presented after the previous worksheet was completed. During long conditions, he was presented with worksheets that were similar in content; however, these worksheets were presented all together at the beginning of the session in a stapled packet.

The second hypothesis (near vs. far) was tested by manipulating staff proximity. In near conditions, staff remained within.3 m of the student when giving instructions or answering questions about academic tasks. In far conditions, staff maintained a distance of at least 1 m from the student when giving instructions or answering questions about academic tasks.

Sammy. Hypothesis 1 (self-monitoring vs. no self-monitoring) was tested in two settings, English and spelling. During all conditions, Sammy independently completed tasks from textbooks or worksheets. Self-monitoring conditions were those in which the student listened to a tape recorder which sounded a bell every minute. At the time of the bell, he recorded whether he had minded his own business (e.g., ignored others' conversations) for the preceding interval by checking Yes or No on a sheet at his desk. Accuracy was monitored with the same procedures that were described above for Desi. No self-monitoring conditions were those in which the student did not use the self-monitoring procedures.

The testing of hypothesis 2 (distanced vs. close physical proximity) was conducted in a music class in which students were engaged in teacher-directed group activities while seated in rows. During the _ close proximity condition, Sammy sat in the front row *very* close to other students so that there was unavoidable physical contact. During the distanced conditions, the students were disbursed so that Sammy sat a row behind and a few seats to the side of other students.

The testing of hypothesis 3 (purposeful vs. analog) was conducted in spelling class. During the purposeful condition, the student made a crossword puzzle using the words from his spelling list and was told in advance that the crossword puzzle would be completed by another student. During the analog condition, the student placed the words in alphabetical order, wrote them three times each, and made up sentences using the words.

Ann. During the testing of hypothesis 1 (choice vs. no choice), choice conditions were those in which Ann was presented with a menu of six comparable academic tasks which were typically assigned in the class. She chose a task from the menu with the option of changing tasks upon completion of at least half of the chosen task. During no choice conditions, the student was presented with similar tasks but in this condition the tasks were chosen by her teacher.

For all sessions of the testing of hypothesis 2 (high attention vs. low attention), Ann was presented with typical academic tasks which were similar in response requirements. High attention conditions were those in which a staff member, seated next to the student throughout the session, praised her for on-task behavior at least three times per minute. Low attention conditions were those in which the staff member, seated at least 6 feet away, *gave* praise statements for on-task behavior only *very* rarely (i.e., once *every* 5 minutes).

Hypothesis 3 (preferred vs. nonpreferred) was tested during math assignments in which the task was counting. The student's preferences were identified through a series of systematic observations (Dyer, 1987) prior to the hypothesis testing. For the preferred activity, the student counted Lego blocks by colors as she constructed objects. For the nonpreferred activity, Ann counted objects on a worksheet and then colored them.

Measurement and Interobserver Agreement

All observations were conducted within the regular classroom settings by observers who had been trained previously to record desirable and undesirable behaviors of students in SEO classrooms. Observer training involved attaining at least 85% agreement on video-taped classroom vignettes and/or during in *vivo* classroom practice observations. During the investigation, data for Desi and Shaun were collected by counting the frequency of each targeted behavior within continuous 1-minute intervals. Data for Garth were collected using an interval system in which the occurrence or nonoccurrence of targeted behaviors were recorded during continuous 15-second intervals. For Sammy and Ann, a partial interval method of recording was used. In this system, the 15-second intervals were divided into 10 seconds of observation followed by 5 seconds of recording.

Data were collected for a total of 15 or 20 minutes per session. Intervals were cued by an audiotape that observers listened to through inconspicuous earphones. During sessions in which interobserver agreement was assessed, two observers listened to the same audio tape with separate earphones. Independence of observations was assured by positioning observers at sufficient distances and angles so that recording sheets could not be overseen. Interobserver agreement for data collected using the frequency within interval method (Desi and Shaun) was calculated by using a Block-by-Block Agreement method (Bailey & Bostow, 1979) in which observer scores were compared interval by interval for a given reliability session. Agreements were counted when both observers recorded an occurrence with in the same interval. That is, if both observers recorded four occurrences within a given interval, the number of agreements for that interval would be four. Disagreements were counted when there were differences between the observers on the number of occurrences within an interval. For example,

one disagreement existed if Observer 1 had recorded two occurrences while Observer 2 recorded three. Percentage agreement was calculated by totaling the number of agreements and disagreements across all intervals for the session, then dividing agreements by agreements plus disagreements, and multiplying by 100. This calculation yielded a percentage agreement between the observers for each session (Kazdin, 1982).

For Garth, Sammy, and Ann, interobserver agreement was calculated for occurrence (OC), nonoccurrence (NO), and total agreement (TA). Agreements between observers for occurrence and nonoccurrence were defined as intervals scored in an identical manner by two observers. Disagreements were those intervals in which the second observer scored the interval in a manner differing from that of the first observer. Percentage agreement was calculated by dividing the number of agreements by the number of agreements plus dis agreements and multiplying by 100.

Overall, interobserver agreement was assessed for 51% of the sessions including sessions for each condition, each hypothesis, and each participant. Interobserver agreement was generally high across both dependent variables. Of the 70 reliability quotients that were calculated, the vast majority (over 87% of the entries) exceed 80% agreement, and more than 71% exceed a reliability criterion of 90%. The conditions that had lower percentages were associated with very few instances of the target behavior. Specific reliability quotients on desirable and undesirable behavior for each participant and each hypothesis can be obtained by writing to the authors.

Results

The results from the functional analyses for the five participants are presented in Figures 1 through 5. These figures summarize the data obtained during the testing of each hypothesis for the participating students. On each figure, levels of undesirable behavior are presented in the graphs on the left while data representing desirable behavior are shown on the graphs on the right. Specific definitions of undesirable and desirable behavior and the complete hypothesis statements have been presented previously (please refer to Tables 2 and 3).

Consider first the data for Desi which are shown in Figure 1. The graphs on the top of the page show the results for the testing of Desi's first hypothesis, analyzing levels of specific praise statements. Five sessions were conducted. In the first, third, and fifth sessions, praise was delivered infrequently while the second and fourth sessions contained relatively dense schedules of praise. These results served to confirm this hypothesis by showing that the high praise (HP) sessions were associated consistently with low rates of undesirable behavior and high rates of desirable behavior. In contrast, the low praise (LP) sessions contained elevated rates of undesirable behavior and decreased rates of desirable behavior. The center and bottom graphs on this figure show the results for the testing of Desi's second and third hypotheses. In both cases, rates of undesirable behavior reflect the expected pattern. That is, self-monitoring (hypothesis 2) and ignoring of undesirable behavior (hypothesis 3) both produced relatively low levels of undesirable responding. Rates of desirable behavior for these two hypotheses did not reveal a consistent pattern.

Figure 2 shows the results for Garth. Inspection of the data for undesirable behavior indicates that, in general, each of the hypotheses was supported. Levels of undesirable behavior were lower when he was (a) given a choice of academic assignments, (b) provided relatively high rates of adult attention, and (c) engaged in self-monitoring. Desirable behavior during these sessions occurred during the majority of intervals and thus there was little difference between the conditions. However, the few sessions with low levels (i.e., below 60%) of desirable behavior were all associated with the expected conditions.

Figures 3, 4, and 5 show the results for Shaun, Sammy, and Ann, respectively. All of these data support the hypotheses that were developed for these three participants. Specifically, in all cases, levels of undesirable and desirable behavior varied consistently and in the predicted direction with the condition that was implemented.

Table 4 summarizes the data from this investigation as means for each condition that was tested. In this table, the data listed under the A columns for undesirable and desirable behavior are associated with those conditions that represented typical classroom practice and that are depicted on each of the figures by the hatched bars. The B columns present the data from those conditions that were hypothesized to produce improvements in student behavior (represented by the solid bars in each of the figures). This table provides additional evidence that the hypotheses were supported by the functional analyses.

Discussion

This study supports and extends previous research on functional assessment. It offers a demonstration of functional assessment (hypothesis development and hypothesis testing) in a manner that is similar to previous studies with individuals with developmental disabilities (e.g., Carr & Durand, 1985; Lennox & Miltenberger, 1989), and it extends this literature to children with emotional and behavioral disorders. In addition, the study was conducted entirely within the ongoing contexts of the children's classroom activities. Although there have been case study illustrations of these latter two features (e.g., Dunlap et al., 1991; Kern et al., 1993), this may be the first demonstration of in-context functional assessment with multiple participants who are identified as emotionally and behaviorally disordered.

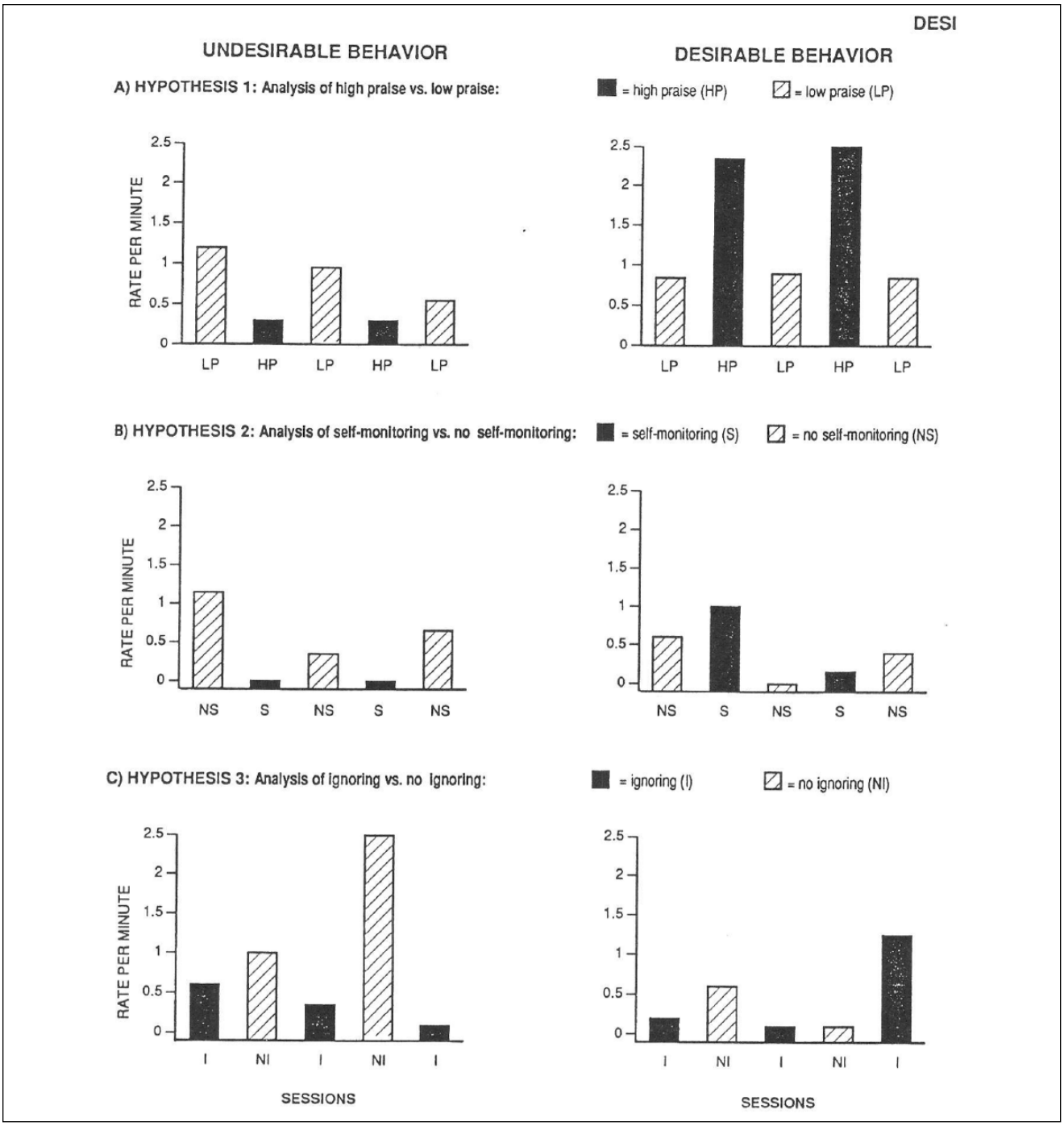

Figure 1. DESI. Results of the functional analyses for each of the hypotheses that were tested with Desi. Levels of undesirable (left) and desirable (right) behavior are shown for each of the three hypotheses.

The results of the functional assessment and functional analysis completed with each of the participants reflect a comprehensive and highly individualized approach to examining the variables maintaining challenging behaviors. In this investigation, the assessment process yielded several noteworthy observations. One finding is that the undesirable behaviors exhibited by each of the five participants were related empirically to observable environmental events in the children's classrooms. A related finding is that the assessment process identified different events for the different children and that there was considerable diversity in the identified variables. The hypotheses in this study implicated antecedent events (e.g., choice making), consequences (e.g., specific praise), and even strategies involving

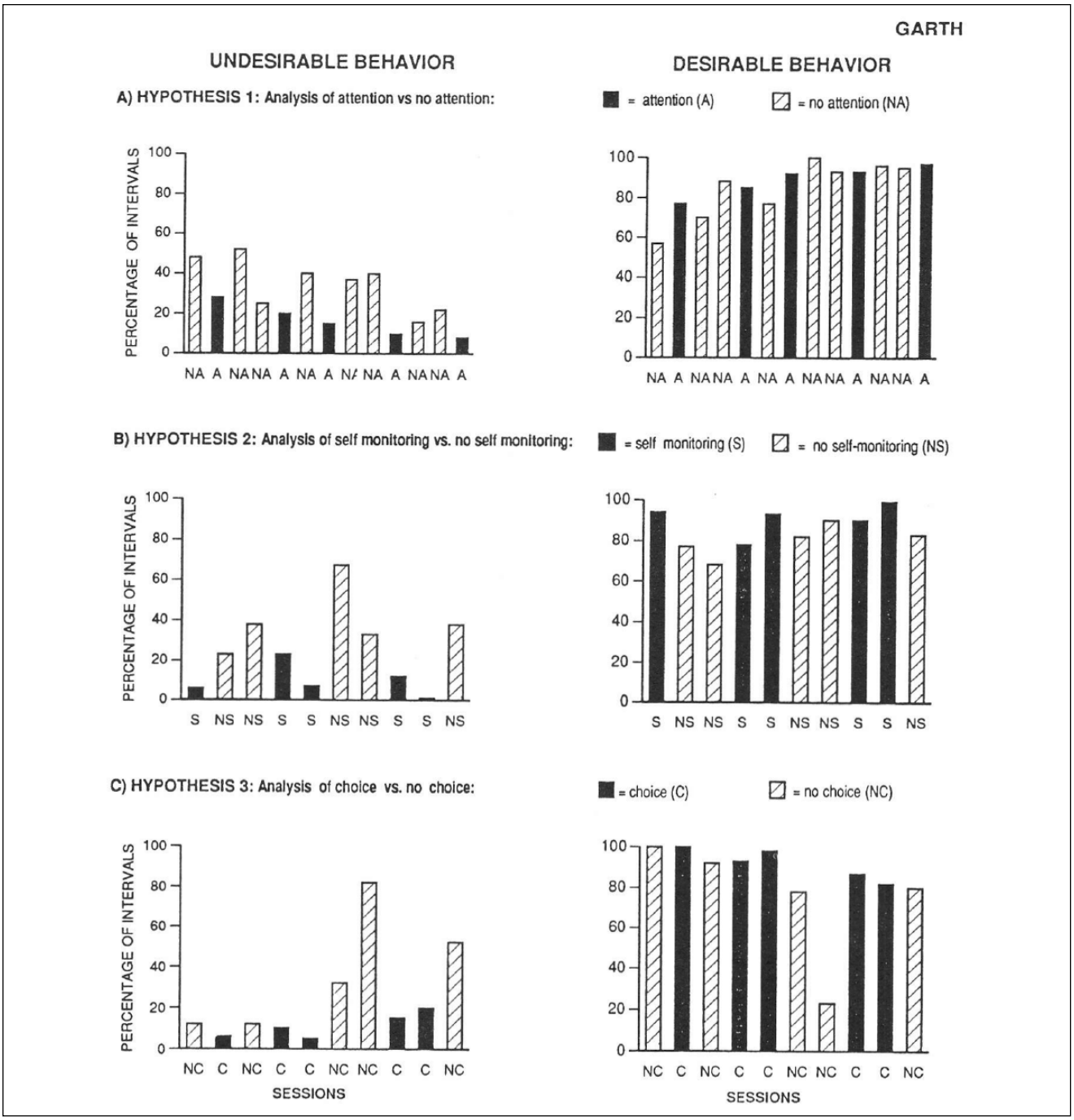

Figure 2. GARTH. Results of the functional analyses for each of the hypotheses that were tested with Garth. Levels of undesirable (left) and desirable (right) behavior are shown for each of the three hypotheses.

self-regulation. In addition, it is important to recognize that variables that improve the responding of some students may have no effect or even be deleterious for other students. For example, close teacher proximity (as in Shaun's case) may be related to higher levels of task engagement and lower levels of undesirable behaviors, while for other students, close teacher proximity may be associated with high levels of disruptive behavior. Without an analysis of individual student responses, such a determination cannot be made with confidence.

The functional analyses in this investigation tested hypotheses statements that related the behaviors of individual children to specific events in the classroom. The analyses were not designed as fully controlled evaluations but

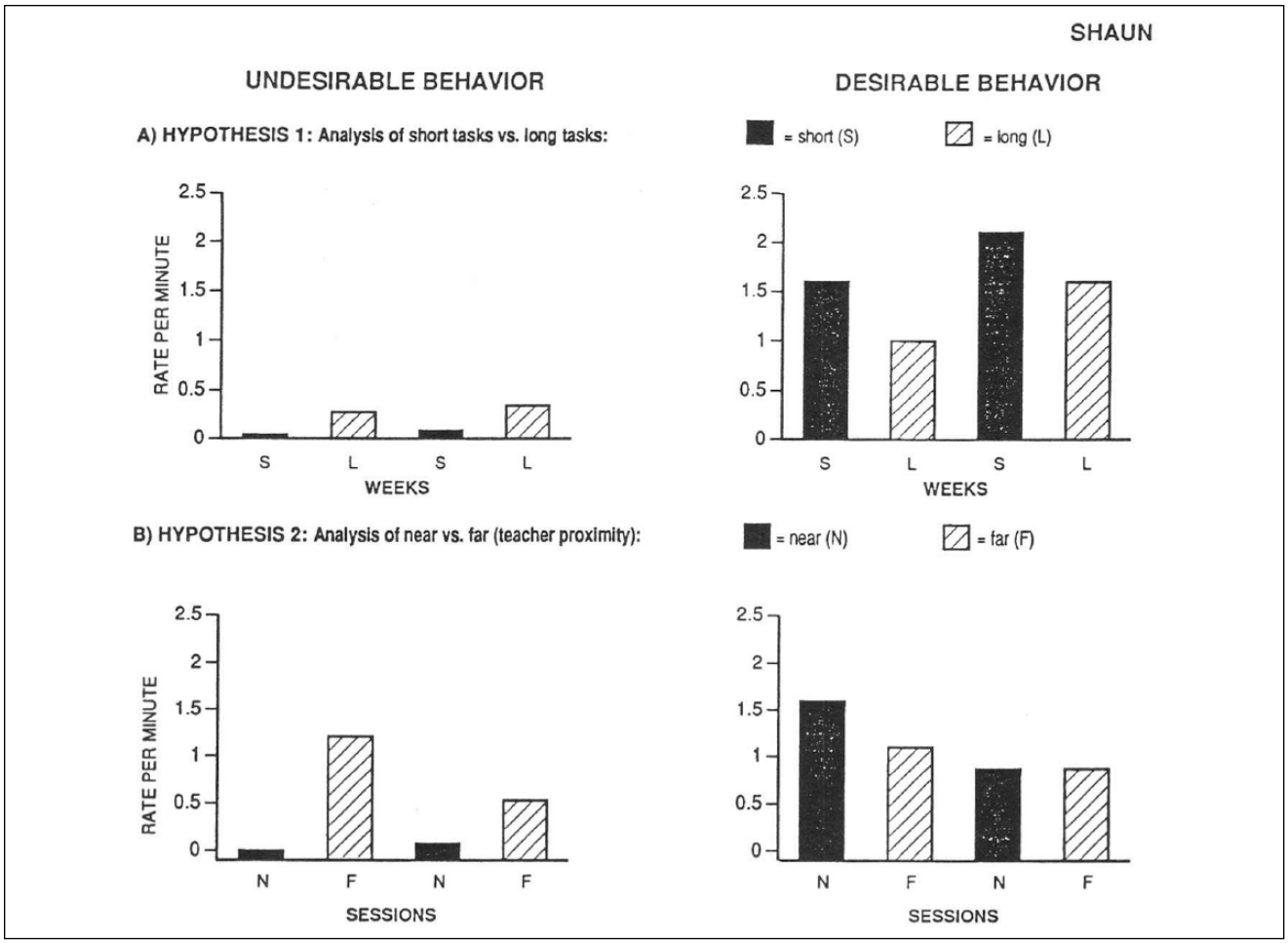

Figure 3. SHAUN Results of the functional analyses for each of the hypotheses that were tested with Shaun. Levels of undesirable (left) and desirable (right) behavior are shown for each of the two hypotheses.

rather as practical trials to assess, in at least a preliminary manner, the validity of the hypotheses and the potential value of the manipulations. Thus, some of the tests can be subjected to alternative interpretations. For example, on the basis of the hypothesis tests alone, Shaun's first manipulations could be related to the incidental attention that was interspersed between the short worksheets; and Sammy's performance on the purposeful activity (hypothesis 3) might have been related to an idiosyncratic preference for crossword puzzles or novelty. From an experimental perspective, such interpretations cannot be discounted; however, the purpose in this case was to conduct an empirical test of hypotheses that were based on considerable data obtained during the hypothesis development phases. As tests of hypotheses, none of the results served to reject the hypotheses and all added confidence that the hypotheses were valid descriptions of functional relationships.

The literature on functional assessment reveals somewhat different objectives across studies. Some investigations have focused 'directly on the operant functions of target behaviors (e.g., Iwata et al., 1982), seeking to determine whether the motivation for the behavior is related to attention, escape, or other identifiable objectives. Other studies have used functional assessment strategies to explore a more broadly defined range of functional relationships (e.g., Cooper et al., 1992; Dunlap et al., 1991). The approach described in this study sought only to identify associations between classroom events and the children's target behaviors (Dunlap & Kern, in press). In some cases, these associations lead to reasonable inferences regarding the operant functions. For example, several of the current hypothesis statements (e.g., Desi's first and third hypotheses, Garth's first hypothesis) clearly implicate praise or attention as a positive reinforcer. In other examples (e.g., Ann's third hypothesis) it is possible to identify the operations of escape responding. However, other hypotheses are not easily related to specific functions or reinforcers (e.g., those having to do with choice or self-monitoring). The relative advantages of these different approaches to functional assessment could be evaluated in future research.

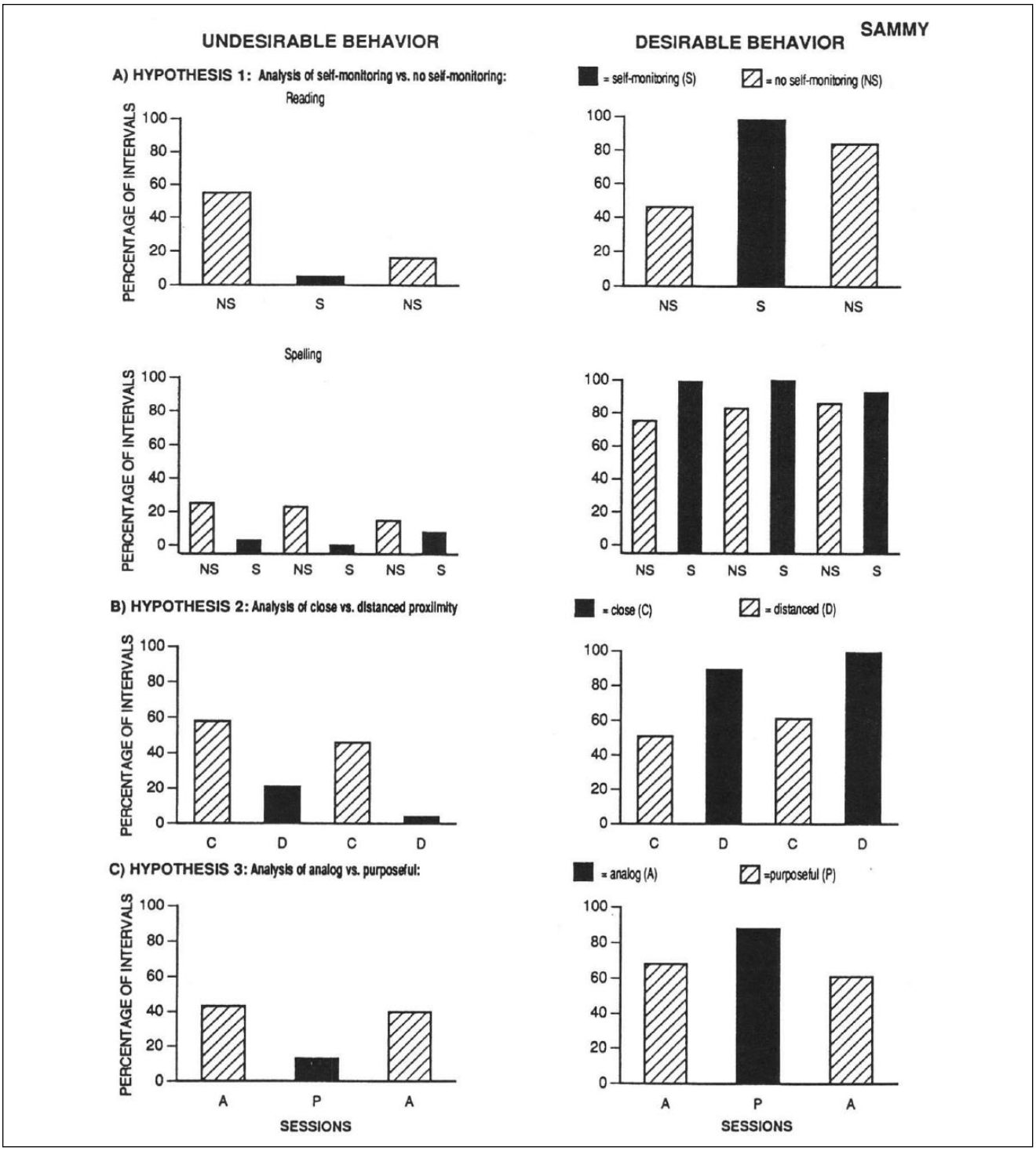

Figure 4. SAMMY. Results of the functional analyses for each of the hypotheses that were tested with Sammy. Levels of undesirable (left) and desirable (right) behavior are shown for each of the three hypotheses. The first hypothesis was tested in both reading and spelling classes.

The methods involved in conducting a functional analysis will continue to be refined. Each of the analyses described here were completed within classroom settings, enhancing the validity of both the processes and the results. Identifying methods for teachers and school personnel to conduct such analyses in an efficient manner is an

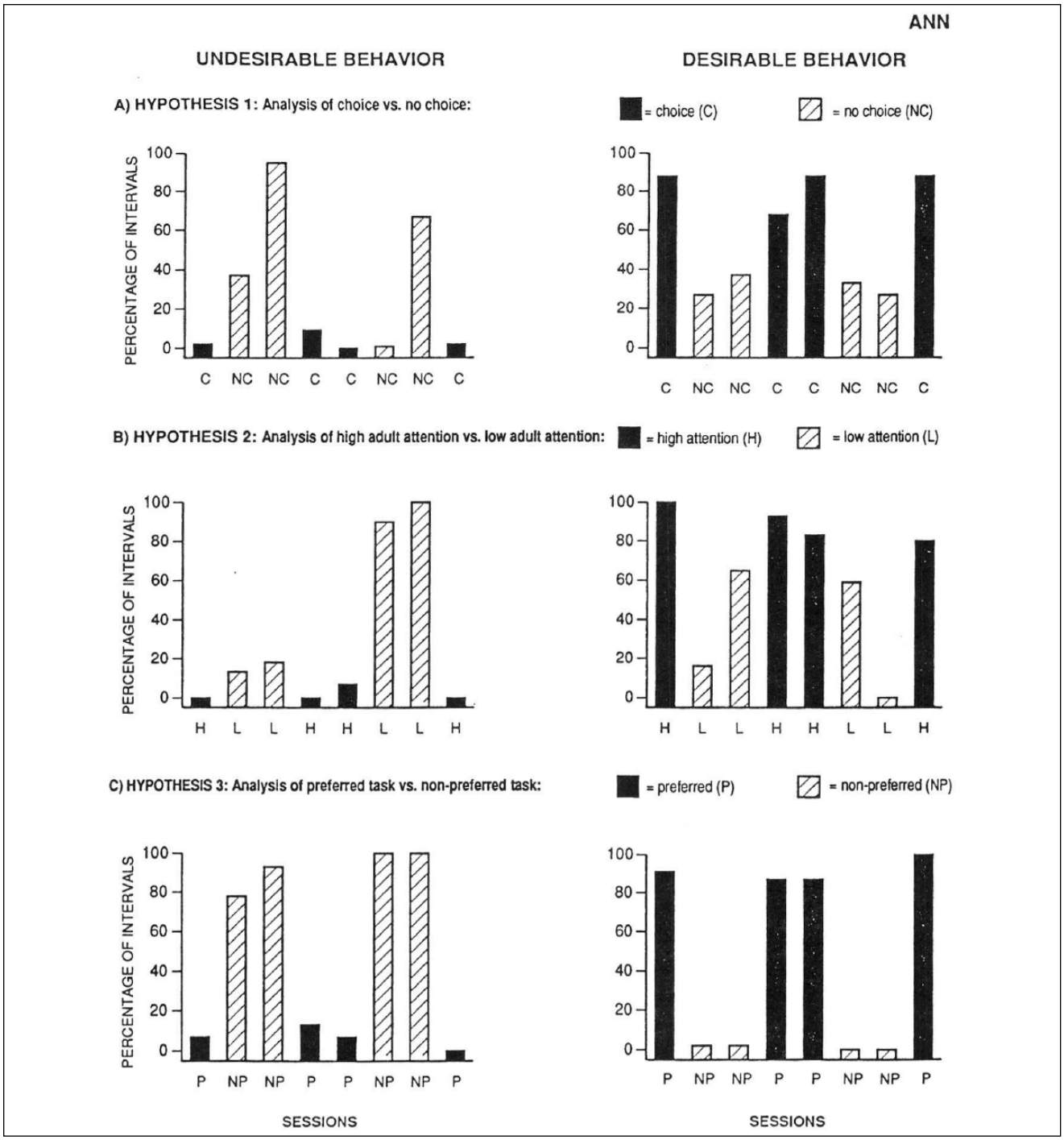

Figure 5. ANN. Results of the functional analyses for each of the hypotheses that were tested with Ann. Levels of undesirable (left) and desirable (right) behavior are shown for each of the three hypotheses.

important objective because it will facilitate understanding and the development of both effective and individualized programs of educational support.

The empirical demonstration of a functional relationship between classroom variables and an individual student's behavioral challenges allows a teacher and/or consultant to de sign an intervention logically related to the child's specific needs. This is especially critical in the development of individualized educational programs for students within specialized classrooms because each student's challenges are governed by a relatively complex and multifaceted set of environmental variables. Not only can multiple

Table 4. Mean Level of Undesirable and Desirable Behavior During the Hypothesis Testing for Each of the Five Participants.

	Undesirable Behavior		Desirable Behavior	
Name	A[1]	B[2]	A[1]	B[2]
Desi				
Hypothesis 1	.9/min	.3/min	.87/min	2.43/min
Hypothesis 2	.72/min	0/min	.33/min	.58/min
Hypothesis 3	1.75/min	.35/min	.35/min	.52/min
Garth				
Hypothesis 1	35%	16%	85%	88%
Hypothesis 2	39%	10%	81%	90%
Hypothesis 3	38%	11%	74%	92%
Shaun				
Hypothesis 1	.31/min	.06/min	1.3/min	1.9/min
Hypothesis 2	.86/min	.04/min	.99/min	1.2/min
Sammy				
Hypothesis 1 (Reading)	36%	5%	65%	98%
Hypothesis 1 (Spelling)	21%	3%	81%	97%
Hypothesis 2	52%	13%	56%	94%
Hypothesis 3	42%	13%	65%	88%
Ann				
Hypothesis 1	50%	3%	31%	83%
Hypothesis 2	55%	2%	35%	89%
Hypothesis 3	93%	7%	1%	91%

[1]A refers to those conditions that represented ongoing classroom practice and are shown on the figures by the hatched bars.
[2]B refers to those conditions that were hypothesized to produce reduced levels of undesirable behavior and are depicted on the figures by the solid bars.

determinants for behavior problems exist but they may be very different in various settings or at various times. The functional assessment process allows for a systematic investigation of these conditions and does so in a manner which yields detailed information about a student's specific challenges.

Although the current study focused on assessment and did not address intervention per se, the process did serve to identify variables and specific classroom manipulations that could very well fit within a comprehensive plan of behavioral support. In this regard, it is worth noting that all of the hypotheses that were developed can be viewed as proactive strategies that serve to prevent rather than suppress undesirable behaviors. This may be particularly noteworthy given recent findings showing the coercive and overly controlling aspects of many behavior management programs in classrooms for this population of children (Knitzer, Steinberg, & Fleisch, 1990).

Another important feature of the current manipulations is that they represent alternatives that appear to be both effective and feasible. Because the hypotheses and the tests were devised collaboratively by teachers and consultants, there is a high likelihood that the variables would be used in the context of interventions and not forgotten at the level of assessment. Nevertheless, feasibility and utilization are essential concerns and should be the subject of future investigations. For example, research should evaluate teachers' satisfaction with the process as well as the extent to which they use the assessment procedures over extended periods of time.

Admittedly, some of the individualized manipulations might be a bit difficult to administer in the context of a classroom with one teacher and one aide. For example, the rate of attention that was provided for Ann during the testing of hypothesis 2 required individualized supervision. However, such levels of attention are sometimes necessary to reduce disruptive behavior and it is usually possible to fade the level of assistance so that it becomes more feasible and ecologically appropriate. In Ann's case, the teachers and consultants agreed that initially high rates of attention were far preferable to intrusive disciplinary alternatives such as timeout.

Finally, further research will need to address the extent to which functional assessment and analysis lead to interventions that are related to meaningful outcomes and subsequently are associated with increased generalization and maintenance of treatment gains. The present study emphasized the assessment and analysis processes rather than the development and implementation of interventions. While the treatment strategies are implied by the results of the experimental manipulations, the effects of their implementation as components within intervention packages will be investigated later. It is expected that interventions based on empirically

validated functional relationships will result in effective and long-lasting treatment outcomes. For students experiencing severely disruptive and recalcitrant behaviors, such analyses and interventions can play significant roles in individualized and effective programs of educational support.

Acknowledgments

The authors are grateful for the cooperation and assistance of Bobbie Pittman, Dawn Bobo, Doug Hoover, Catherine McClain, Amy Weinstein, and Peggy Middleton. Important contributions in data collection and data summary were provided by Suzanne Wright, Arcadia Gomez, Susan Pearsall, and Vicki Dominick. Requests for reprints should be addressed to the first author.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This investigation was supported by Field-initiated Research Grant No. H023C10102 from the U.S. Department of Education, Office of Special Education Programs. Additional support for data collection and manuscript preparation was provided by the Florida Mental Health Institute and by Cooperative Agreements No. G0087C0234 and H133B20004 from the U.S. Department of Education, National Institute on Disability and Rehabilitation Research. However, opinions expressed in this manuscript are those of the authors and no endorsement by supporting agencies should be inferred.

References

Bailey, J. S., & Bostow, D. E. (1979). *Research methods in applied behavior analysis*. Tallahassee, FL: Copy Grafix.

Bailey, J. S., & Pyles, D. A. M. (1989). Behavioral diagnostics. In E. Cipani (Ed.), *The treatment of severe behavior problems* (pp. 85-106). Washington, DC: American Association on Mental Retardation.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis*, *1*, 175-191.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, *18*, 111-126.

Carr, E. G., Robinson, S., & Palumbo, L. W. (1990). The wrong issue: Aversive versus nonaversive treatment. The right issue: Functional versus nonfunctional treatment. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp. 362-379). Sycamore, IL: Sycamore Press.

Cooper, L. J., Wacker, D. P., Thursby, D., Plagmann, L. A., Harding, J., Millard, T., & Derby, M. (1992). Analysis of the effects of task preferences, task demands, and adult attention on child behavior in outpatient and classroom settings. *Journal of Applied Behavior Analysis*, *25*, 823-840.

Dunlap, G., & Kern, L. (in press). Assessment and intervention for children within the instructional curriculum. In J. Reichle & D. Wacker (Eds.), *Communicative approaches to the management of challenging behavior problems*. Baltimore, MD: Brookes.

Dunlap, G., Kern-Dunlap, L., Clarke, S., & Robbins, F. R. (1991). Functional assessment, curricular revision, and severe behavior problems. *Journal of Applied Behavior Analysis*, *24*, 387-397.

Durand, V. M., & Crimmins, D. B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders*, *18*, 99-117.

Dyer, K. (1987). The competition of autistic stereotyped behavior with usual and specially assessed reinforcers. *Research in Developmental Disabilities*, *8*, 607-626.

Dyer, K., Dunlap, G., & Winterling, V. (1990). The effects of choice making on the serious problem behaviors of students with developmental disabilities. *Journal of Applied Behavior Analysis*, *23*, 515- 524.

Horner, R. H., Dunlap, G., Koegel, R. L., Carr, E. G., Sailor, W., Anderson, J., Albin, R. W., & O'Neill, R. E. (1990). Toward a technology of "nonaversive" behavioral support. *Journal of the Association for Persons with Severe Handicaps*, *15*, 125-132.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities*, *2*, 3-20.

Iwata, B. A., Vollmer, T. R., & Zarcone, J. H. (1990). The experimental (functional) analysis of behavior disorders: Methodology, applications, and limitations. In A. C. Repp & N. N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities* (pp.301-330). Sycamore, IL: Sycamore Press.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.

Kern, L., Childs, K. E., Dunlap, G., Clarke, S., & Falk, G. D. (1993). *Using assessment-based curricular intervention to improve the classroom behavior of a student with emotional and behavioral challenges*. Manuscript submitted for publication.

Knitzer, J., Steinberg, Z., & Fleisch, B. (1990). *At the schoolhouse door: An examination of programs and policies for children with behavioral and emotional problems*. New York: Bank Street College of Education.

Lalli, J. S., Browder, D. M., Mace, C. F., & Brown, D. K. (in press). Teacher use of descriptive analysis data to implement interventions to decrease students' problem behaviors. *Journal of Applied Behavior Analysis*.

Lennox, D. B., & Miltenberger, R. G. (1989). Conducting a functional assessment of problem behavior in applied settings. *Journal of the Association for Persons with Severe Handicaps*, *14*, 304-311.

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

Mace, F. C., Yankanich, M. A., & West, B. J. (1989). Toward a methodology of experimental analysis and treatment of

aberrant classroom behaviors. *Special Services in the Schools*, *4*, 71-88.

O'Neill, R. E., Horner, R. H., Albin, R. W., Storey, K., & Sprague, J. R. (1990). *Functional analysis of problem behavior: A practical assessment guide*. Sycamore, IL: Sycamore Publishing.

Repp, A. C., Felce, D., & Barton, L. E. (1988). Basing the treatment of stereotypic and self-injurious behaviors on hypotheses of their causes. *Journal of Applied Behavior Analysis*, *21*, 281-289.

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