Teacher Education and Special Education
2005, Volume 28, No. 2, 000–000

Scoring Rubric for Assessing Students' Performance on Functional Behavior Assessment Cases

Sekhar S. Pindiprolu, Benjamin Lignugaris/Kraft, Sarah Rule, Stephanie Peterson, & Timothy Slocum

Abstract: *This article describes the procedures employed to develop a generic scoring rubric to evaluate students' responses on functional behavior assessment cases. The generic rubric was then used to develop case specific scoring rubrics. The components of the generic scoring rubric, its effectiveness to reliably score students' responses across cases, and its utility in evaluating students' responses for errors are described. The components of the scoring rubric helped in reliably scoring students' responses. The importance of scoring rubrics, in the context of the recent mandates of accreditation agencies like National Council for Accreditation of Teacher Education, the limitations of the scoring rubric, and suggestions for further research are discussed.*

In the recent past there has been an emphasis in employing cases as teaching tools in teacher education (Bronack & Kilbane, 1998; Morine-Dershimer, 1996; Pindiprolu, Peterson, Rule, & Kraft, 2003). Cases are context-based narratives that help students understand the idiosyncrasies associated with real-world practicalities (Bronack & Kilbane, 1998; Pindiprolu et al., 2003) and require students to perform a task or generate their own responses rather than identify a correct response from a set of responses. Cases are primarily employed by instructors as a teaching method to bridge the gap between theory and practice (Knirk, 1991; Pindiprolu et al., 2003).

During the last two decades the focus of behavioral assessments has expanded from defining and measuring problem behaviors to identifying and analyzing the relationship between problem behaviors and their environmental events (Iwata, Vollmer, Zarcone, & Rodgers, 1993; Umbreit, 1995). A variety of methods have been employed to analyze the relationship between problem behaviors and environmental variables associated with them. These methods, which identify and verify events that maintain problem behaviors, are collectively referred to as functional assessments or functional behavioral assessments (Gable, 1996; The Center for Effective Collaboration and Practice, 1998; Umbreit, 1995). Functional behavior assessments are defined as "combining descriptive and experimental methods to determine whether problem behavior is positively rein-

This study was supported by the Center for Persons with Disabilities at Utah State University. This study was part of the dissertation completed by the first author in partial fulfillment of the requirements for a Doctor of Philosophy in Special Education at Utah State University. We would like to thank Hollie Berglof for her assistance in conducting the study and the reviewers of the manuscript for the suggestions.

Address correspondence to Sekhar S. Pindiprolu, Department of Human Learning and Development, Box 70548, East Tennessee State University, Johnson City, TN, 37614. Electronic mail may be sent to pindiprs@etsu.edu.

TESE, **Volume 28, No. 2**
Spring 2005

forced via attention and/or tangibles-or sensory stimulation-or negatively reinforced via escape from either task demands or aversive sensory stimulation" (Umbreit, 1995, p. 267).

Functional assessment procedures are broadly classified into three categories: indirect or informant procedures; direct procedures or descriptive analysis; and functional analysis or experimental analysis (Gable, 1996; Iwata et al., 1993; Ward, 1998). Experimental analysis is generally conducted after indirect and/or direct methods have been undertaken and the hypothesized functions of a problem behavior are generated. During systematic manipulations, the events or variables that maintain or correlate with problem behavior are repeatedly introduced and removed using single subject experimental design tactics. The hypothesized functions are verified by contrasting the occurrence of problem behavior during conditions when the events maintaining problem behavior are introduced with the occurrence of the problem behavior during a suitable control condition in which events (hypothesized as) maintaining problem behavior are removed (Iwata, et al., 1993). Unlike the direct and indirect methods, which are suggestive of the events affecting problem behavior, experimental analysis helps in verifying the role of the events in triggering problem behaviors (Iwata et al., 1993).

Recently, cases were employed to teach functional behavior assessment (FBA) skills to preservice and inservice teachers (Chandler, Dahlquist, Repp, & Feltz, 1999; Pindiprolu et al., 2003). To conduct functional behavior assessments, students must have knowledge in multiple skills areas. They must understand (a) how to conduct indirect and direct functional behavior assessments, (b) the principles of positive and negative reinforcement, (c) hypothesis development and how to design experimental conditions that will examine variables associated with the problem behavior(s), and (d) single subject research designs (Pindiprolu, 2001). Further, as the environmental variables affecting a problem behavior are idiosyncratic (Iwata et al., 1993), school personnel need to have knowledge about modifying conditions of a functional analysis in accordance with the variables being verified (Pindiprolu, 2001). Thus, training in functional assessments and more specifically in functional analysis procedures require teaching tactics that can facilitate logical knowledge (when and how to modify procedure) of functional assessments. The failure to facilitate the logical knowledge of assessment procedures with teachers and other school personnel may inhibit the transfer of learning to new situations (Pindiprolu, 2001). The ideal way to facilitate logical knowledge of functional assessments is to provide preservice teachers multiple opportunities to conduct functional assessments with children who display problem behaviors that serve different functions. However, in the real world it is difficult to provide preservice teachers with such varied experiences due to time constraints. One way to overcome the problem is to use cases as tools for teaching functional assessments.

FBA cases can help students understand how to design experimental conditions and how to modify those conditions. They also offer multiple advantages to instructors. They provide (a) real-life situations to promote students' interest, (b) multiple examples to facilitate generalization of skills, and (c) connections between theory and practice in a non-threatening environment (Pindiprolu, 2001). Further, students' performance on cases is likely to reveal their understanding and facility in application of what they are learning (Arter & McTighe, 2001).

However, the utility of cases as teaching tools is limited by the quality of the scoring rubrics employed to evaluate students' performance. When scoring rubrics are employed that globally describe important components of a task, student responses (on a case) are evaluated in a subjective way, which could prevent meaningful and consistent feedback to the students on complex tasks like FBAs. A more specific scoring rubric, on the other hand, can transform a subjective evaluation process into a clear, consistent, and verifiable procedure (Arter & McTighe, 2001). A good scoring rubric consists of a (a) clear description of each component of a task; (b) well operationalized definition of each component; (c) clear and objective scoring criteria that reflect current best practices in the field; and (d) reliable and feasible cri-

teria that can be applied to a wide variety of tasks (Herman, Aschbacher, & Winters, 1992).

There are different types of scoring rubrics. Some are holistic and others are analytical. Holistic rubrics provide a single score for a student product or performance (i.e., all the components of the task are evaluated and a single score is assigned to a students' performance/product). In an analytical scoring rubric, the important dimensions of a task are judged separately and assigned a score. The score for each dimension is added to obtain an overall score for a performance or product. Analytical scoring rubrics help in differentiating students' performance on all the important components of a task and in assessing students' errors, which in turn facilitates the instructor's ability to provide appropriate feedback (Arter & McTighe, 2001). In addition, analytic scoring rubrics can help instructors design targeted learning experiences, based on student errors, for specific components of complex tasks such as FBAs. Finally, analytic scoring rubrics help document student performance and learning progress, which are required by the National Council for Accreditation of Teacher Education (NCATE) and other accreditation agencies.

Scoring rubrics are further classified as either task specific or generic. Generic scoring rubrics help in using the same criteria across different tasks that measure the same skills or dimensions. They are generally useful for measuring complex skills (Arter & McTighe, 2001) like FBAs. Task specific rubrics, on the other hand, are applicable only to one task and usually help in assessing particular facts and/or procedures of a particular task. Thus, their utility in evaluating complex skills, such as FBA skills, are limited. However, task specific rubrics help in reliable and consistent scoring of the product. Thus, it is important to develop a generic scoring rubric that delineates objective scoring criteria that reflects current best practices in the field and then to use the generic scoring rubric guidelines to develop task specific scoring rubrics to help assess a student's performance on a case.

Given the NCATE focus on students' performance, which calls for documentation of a student's knowledge, skills, and dispositions and their progress over the years, a need exists in the field of teacher education for the development and dissemination of examples of scoring rubrics that facilitate assessment of students' performance in the essential skill areas. For special education, the essential areas, based on the Council for Exceptional Children Standards, include lesson plans, individualized education programs, individualized family service plans, functional behavior assessments and behavioral plans, transition plans, and delivery of instruction.

The purpose of this paper is to describe (a) the process undertaken to develop an analytic scoring rubric to measure students' performances on FBA cases, (b) the components of an analytic scoring rubric to evaluate students' performance on FBA cases, and (c) the effectiveness of the scoring rubric in objectively evaluating students' responses. Further, implications and suggestions for future research in this area are discussed.

Context of the Study

An experimental pretest-posttest design was used to examine the effects of three case-based teaching tactics on the preservice teachers' knowledge and application of FBA skills. Instruction in FBA was delivered before the students were exposed to teaching tactics. The three teaching tactics were administered with three groups of preservice teachers. The groups were Acropolis text group, Acropolis chat group, and Acropolis chat group with student interactions. The students in the Acropolis text group were given access to all the assessment information on a case and were asked to work on the cases individually. The chat groups had to elicit assessment information by asking questions of the first author. The students in the Acropolis chat group worked on the cases individually whereas the students of the Acropolis chat group with student interactions worked on the cases in a group situation. After all the members in a sub-group completed a case study, correct responses for the five FBA questions on the case were displayed on the Website (each student accessed different pages based on their group ID). The three teaching tactics with cases were

TESE, **Volume 28, No. 2**
Spring 2005

conducted online. At the end of the study, all students were tested individually. The students logged into the chat tool on the Acropolis Website, the first author sent background information on a case, which was a video clip of a teacher asking for help to conduct FBA on a 19 year old student with behavioral problems. The students then elicited information by asking questions and when they had sufficient information answered five FBA questions. The purpose of the study was to examine the effects of the three teaching tactics with cases for facilitating preservice teacher's skills in conducting functional behavioral assessments (i.e., can the students identify relevant information about problem behaviors, identify variables affecting a problem behavior, and develop a plan to test their hypothesis). It is in this context a generic, analytic scoring rubric was developed to assess students' ability to develop a FBA when presented with cases and offered information.

Five cases portraying pupils with problem behaviors were developed for the study. The cases were based on real-life examples of pupils with problem behaviors for whom the functional behavioral assessments were conducted in order to identify functions of the problem behaviors. Each case consisted of background information and assessment information. The background information included the name of the student with problem behavior, age, class he or she is attending, teacher name and his or her concerns, disability information, and types of assessment information available on the student (i.e., student-assisted and/or teacher-assisted interviews). Assessment information included descriptions of the topography of behaviors, their frequency, the conditions under which the behaviors did or did not occur, typical class routines, ecological events, staffing patterns, and past interventions and their effects. The Student-assisted Functional Assessment Interview (Kern, Dunlap, Clarke, & Childs, 1995) format was used to develop the student interview content. The Functional Assessment Interview (O' Neill, Horner, Alnbin, Storey, & Sprague, 1990) format was used as the framework to structure and develop teacher interview content for each case.

The assessment information on each case consisted of necessary information to identify all the possible functions of the problem behavior(s). After the content for each case was developed, the fourth author examined the content for its coherence and completeness (see Pindiprolu, 2001). The cases varied in terms of age of the pupil with problem behaviors, disability conditions, and functions of problem behavior. Out of the five case studies, three were used for intervention purposes and one each for pretest and posttest measures to test application of FBA knowledge of students (see Pindiprolu, 2001). The purpose of the cases was to provide students with practice deciding what kinds of FBA data to collect, determining what hypotheses a particular data set support and developing a strategy to test their hypotheses. The problem solving real-life cases provided opportunities to practice these analytical skills in a non-threatening situation and also to develop some fluency with analytical skills by exposing students to multiple examples of pupils with problem behaviors.

Method

Participants

The sample consisted of 79 students who were enrolled in the SPED 5050: Applied Behavior Analysis course, a semester-long (15 weeks) course at Utah State University. The course was offered in two sections: an on-campus and a distance education section. The on-campus class was comprised primarily of typical undergraduate students (i.e., full-time students) and the distance education class was comprised primarily of students who were working in classrooms (as paraeducators or instructors with temporary licenses) and completing their undergraduate degrees. The sample was selected based on the following criteria: The student (a) had taken the introductory course in Applied Behavior Analysis, (b) had participated in a series of pretests evaluating knowledge of the course content, and (c) was not a teacher at the Utah Schools for the Deaf and Blind. Seven students who wanted to participate but did not meet one or more of the criteria were assigned to a pilot group.

Development of the Scoring Rubric

A literature review was undertaken to identify existing scoring rubrics employed with functional behavioral assessment cases. Computer searches of the ERIC and psychLIT databases revealed one dissertation study (Ward, 1998) that employed a scoring criterion for assessing participant's performance on FBA cases. In this study, inservice teachers worked with a real student with problem behavior(s) and the teacher's reports were evaluated using a checklist for the presence or absence of 12 components: adult and/or student interview; motivation assessment scale; behavioral definition; antecedent or predictor; consequence or function; summary statement/hypothesis; functional direct observation; problem behavior diagram; test hypothesis; problem analysis; intervention plan; and progress monitoring. In addition, participants' case reports were coded for quality on the following seven indicators: (1) behavioral definition, (2) hypothesis statement, (3) functional direct observation, (4) test hypothesis, (5) problem analysis, (6) intervention plan; and (7) progress monitoring. The authors employed a 5-point scale/rubric to evaluate the quality of the case reports on the above listed 7 dimensions (see Ward, 1998).

The scoring rubric used in the present study employed similar dimensions, but focused on the analytical skills to conduct a functional behavior assessment. Specifically, the purpose of this training was to refine preservice teachers' analytical skills. Thus, in the present study the preservice teachers needed to (a) define a pupil's problem behavior; (b) request information about the pupil that would be helpful in determining a hypothesis; (c) formulate a hypothesis about the function of the problem behavior; (d) describe the data that supports the hypothesis; (e) develop a plan for testing the hypothesized function(s) of the pupil's problem behavior (i.e., antecedent or consequence analysis); and (f) develop a graph to display the patterns of data regarding problem behavior for each phase of the functional analysis. In essence the scoring rubric needed to reflect the quality of preservice teachers' thinking about the functional assessment process, what they need to find out about the pupil to derive hypotheses and the strategies they should use to confirm or reject their hypothesis.

Scaling

A three-point scale/rubric with descriptive task specific criteria was developed for each trait (see Table 1). The rubric was then pilot tested with a group of students who were not part of the experimental study. Based on this initial field test, four limitations were identified. First, some of the students' responses did not fit into the categories. For example, if a student identified two problem behaviors and described the third behavior (when there were three target behaviors listed in a case), the task specific scoring rubric did not have a category to accommodate the response (see Table 1). Second, students received the same score on a trait in spite of differences in quality of their answers. For example, students who identified one, two, or three problem behaviors received the same scores. Third, students' answers did not show a relationship across traits. For example, if a student identified three hypothesized functions for a problem behavior(s), in some cases s/he did not provide a plan for testing all three hypotheses. Fourth, the scoring rubric was task specific (the number of problem behaviors and functions of the problem behavior for each case determined the number of correct responses and their spread on the 3-point scale) and hence a separate scoring rubric was required to assess students' responses across cases. For example, a case that had two problem behaviors were more spread out on a 3-point scale than a case with three problem behaviors. For the former, the students received 1 point for identifying the two problem behaviors and 1 point each for describing the problem behaviors. In the latter, the students received 1 point for identifying all problem behaviors and describing a problem behavior, 1 point for describing two problem behaviors and 1 point if all the behaviors were described. Thus the scoring criteria were dependent on the case and the students' scores were not comparable across cases (i.e., a student received 1 point on a case when s/he

TESE, **Volume 28, No. 2**
Spring 2005

Table 1. Initial Analytic, Task Specific Scoring Criteria

Traits	3 Points	2 Points	1 Point
Defining Problem Behavior	All three behaviors identified and described	One or two problem behaviors described	Problem behaviors identified (one or two or three)
Hypothesized Function	Both functions listed with A-B-C components	One function listed with A-B-C components	Only correct function or functions mentioned
Sources of Data and Specific Information	Both functions are listed and supporting data from interviews provided for each function	Both functions are listed and supporting data provided for one function	The supporting data is mentioned as obtained from interviews
Plan to Test the Functions	Consequence analysis: Free play, escape, and tangible conditions are listed and described	Consequence analysis: Only two conditions identified and described	Consequences analysis: Only conditions mentioned
	Antecedent analysis: Conditions for both functions mentioned and described correctly.	Antecendent analysis: Conditions for one function is correctly described	Antecedent analysis: Conditions for one function partially correct
Graphs	All the conditions for all functions listed, data reflects hypothesized function, number of data points are three for each phase, axes information correct.	All conditions for one function given, data reflects hypothesized function; data points are three for each phase, axes information correct.	Only the number of data points and axes labels are correct. Data display is incorrect

might have described only one of the three problem behaviors and 1 point on another case when s/he only listed the problem behaviors). This hindered the documentation of students' progress and the analysis of students' errors across cases.

Based on these initial data, the *delineation of sources of data and specific information considered for formulating the hypothesized functions* trait was left intact. For the remaining traits or questions, (behavioral definition of the pupil's problem behavior(s), formulating a hypothesis regarding the function(s) of the pupil's problem behavior(s), a plan for testing the hypothesized function(s) of the pupil's problem behavior, developing graph(s) to display the patterns of data regarding problem behavior for each phase of functional analysis) a generic scoring rubric was developed that included a separate scale score for quality and quantity dimensions of each trait (see Table 2).

The operationalized definition of each component of the revised rubric and the objective scoring criteria employed to score the responses are described with the help of a case that the participants in the study responded to as part of the intervention. The participants in the study, depending on which group they are assigned, were either given information or elicited the information on the case by asking questions of the first author (see Pindiprolu, et al., 2003). The salient information about the case is presented below.

Case Example

B.J. is a 4-year-old male diagnosed with communication delays and is currently attending the self-contained classroom for preschoolers at an elementary school. The preschool unit serves three and four year olds with special needs. B.J. has attended this class for the past year. The class teacher, Mrs. Heidi, is worried about B.J.'s non-compliance and aggressive behaviors. According to Mrs. Heidi, the problem behaviors of concern are B.J.'s non-compliance and aggressive behaviors. B.J.'s non-compliance behaviors include saying "no" and/or "I can't do it", yelling above conversation level, stamping his feet on the ground and running away. Aggressive behaviors include pushing and/or hitting other children. These behaviors generally occur when B.J. has to stand in the

Table 2. Components of Generic, Analytic Scoring Rubric

Traits	Quantity	Quality
Defining Problem Behavior	One point for each correct identification of problem behavior.	One point for each description of a problem behavior.
Hypothesized Functions	One point for correct identification of each function.	One point for correct description of the A-B-C components for each correct function.
Source of Data and Specific Information	Half a point for each piece of information for each function. The information is obtained from student or teacher interview data/records (maximum four points).	
Plan to Test Hypothesized Function	A point for mention of each correct contrast condition needed to test each correct function.	One point each for the correct description of manipulation of variables in contrast conditions for each correct function. Behavior ignored or reinforced, easy and hard tasks mentioned, and one variable manipulated while others are kept constant (the varible depends on the function).
Hypothesized Graph	A point for data correctly displayed in control and contrast conditions for each correct function.	A point for graph for each correct function and axes labeled and underlined, minimum of three data points for each phase, and data points within each phase connected.

line to access books in the library or go to his bus, when requests are made to participate in a small group activity, and while playing at the toy center.

According to Mrs. Heidi, B.J. is diagnosed has having delays in expressive communication and, more specifically, with articulation of speech sounds. Mrs. Heidi also reported that B.J. sleeps on the bus almost every day before coming to school, but not in the classroom. Mrs. Heidi doesn't think that the problem behaviors are associated with or a result of his sleep patterns and/or his eating routines. B.J.'s typical schedule includes, large group activities from 8:30 to 9:00 a.m., small group activities from 9:05 to 9:30 a.m., and recess at 9:30. After recess the class has snack at 9:45 a.m., free play from 10:00 to 10:30 a.m. and wind up activities from 10:30 to 10:45 a.m. His daily activities are very predictable. At school, B.J. has the opportunity to choose classroom jobs, snack items, puzzles, blocks, and books. At home, B.J. is with his mother and receives one-to-one attention from her. At school, he is with four adults and approximately 13 other children. According to Mrs. Heidi, noise and confusion seem to make B.J. more aggressive. Training of the staff or their social interactions doesn't seem to impact B.J.'s problem behavior.

Mrs. Heidi reported that B.J.'s problem behaviors are most likely to happen throughout the time he is in preschool. The problem behaviors occur in the classroom, in the hall, or outside the classroom. The problem behaviors occur in the presence of all the staff persons. The activities that are most likely to produce the behaviors are transitions from large group to small group, being required to wait to access books/to go to the bus, coming inside after outside play, and when he is not given his first preference to choose a toy. The behaviors are least likely to occur when he is allowed to have his first preference, during the large group activities, and at snack time. Another situation that seems to "set off" the behaviors is when there is noise or confusion in the classroom. Mrs. Heidi reported making demands such as, "you need to go a particular table/small group activity" would most likely make the undesirable behaviors occur. Mrs. Heidi reported that when she asked B.J. to perform a difficult task (e.g.,

TESE, **Volume 28, No. 2**
Spring 2005

trace his name), he would say, "I don't want to!" She also reported that if a desired activity was interrupted (e.g., when told to come inside), B.J. would ignore her/adult, run away, or not move at all. Mrs. Heidi was not sure how B.J. would react if the typical routines were changed or if B.J. was not given attention for 15 mins. However, she reported that B.J. likes one-on-one play and/or reading activities.

According to Mrs. Heidi, some things that might be done to improve the likelihood that a teaching session will go well with B.J. are ignoring inappropriate behavior, keeping activity and noise level down, and structuring transitions. She felt that not giving him a turn and confronting him would escalate the problem behaviors. According to the teacher, B.J. likes candy, bread, cookies, rolls, puppets, puzzles, dinosaur books, someone reading books to him, bus rides, and going to the park.

Quantity Dimensions

For the *behavioral definition of the pupil's problem behavior(s)* trait, 1 point was awarded for correct identification of each problem behavior. For the above case, 1 point was awarded for correct identification of each behavior, i.e., non-compliance, aggression, etc. For *formulating a hypothesis regarding the function(s) of the pupil's problem behavior(s)* trait, 1 point was awarded for correctly identifying each function. For B.J..'s case, the hypothesized function of non-compliance was to "escape" from tasks. The participants received 1 point for identifying the correct function "escape". Participants were not penalized for commissions of errors for a problem behavior. That is, if the participants identified an extra function or a wrong function such as attention for non-compliance, s/he did not earn or lose any points.

For the *plan for testing the hypothesized function(s) of the pupil's problem behavior* trait, 1 point was awarded if a student's response included reference to the control and experimental (or contrast) conditions needed to test the correct function of a problem behavior(s). For B.J..'s example, if a student identified "escape" as the function of the non-compliance behavior, the mention of "high demand task" vs. "low demand task" conditions for an antecedent analysis or a "free play" and "escape from high demand tasks" conditions for a consequence analysis was required in the response to receive a point. Similarly, for the *developing hypothesized graph showing the data patterns in different phases* trait, 1 point was awarded for correct display of data patterns in the control and contrast conditions. For example, for a behavior with an escape function, hypothesized graphs that depicted behavioral problems occurring at low rates during a "low demand" condition and at high rates during the "high demand" condition (for an antecedent analysis) received a point.

For the *delineation of sources and specific information considered for formulating the hypothesized function(s) of the pupil's problem behavior* trait, ½ point was awarded for each correct information piece that was listed in the student/teacher interview portion of the assessment information that supported the hypothesized function. The maximum points awarded were 4 points. For B.J..'s case, a student's response should have included that (1) the problem behavior did not occur during snack time or large group activities that B.J. liked, (2) or time out employed by the teacher did not help decrease the problem behavior, etc.

Quality Dimensions

For the *behavioral definition of the pupil's problem behavior* trait, 1 point was awarded for correct behavioral description of a problem behavior(s). For B.J..'s case, a student's response indicating that the non-compliance behavior involves saying "no" and/or I can't do it, yelling above conversation level, stamping feet on the ground and running away received 1 point.

For the *formulating a hypothesis regarding the function(s) of the pupil's problem behavior* trait, a student received a point for correctly describing the antecedents, behavior, and consequence of each correct behavior/function. For B.J..'s example, a student's response received 1 point if it contained the following information: When Mrs. Heidi asks B.J. to trace his name or go to a particular group activity, B.J. exhibits non-compliance behav-

ior and avoids tracing his name or going to the group activity. The hypothesized function of B.J..'s behavior is to escape from tasks.

For a *plan for testing the hypothesized function(s) of the student's problem behavior* trait, 1 point was awarded if the response for each function indicated (a) whether the problem behavior is reinforced or ignored during contrast phases, (b) a list of environmental variables held constant during different experimental phases, and (c) specific environmental variables that were manipulated for testing a function. For example, to receive 1 point for a plan to test an escape function using an antecedent analysis, a response should indicate that (a) all the problem behaviors will be ignored throughout the analysis, (b) a high demand task such as tracing name should be used during the high demand condition, and (c) a low demand task, such as playing with puzzles should be used during the contrasting low-demand condition. To receive 1 point for a consequence analysis, a response should have indicated (a) a description of the free play condition (100% teacher's attention, easy task such as playing with his favorite toy) and (b) a description of the escape condition (100% attention, the task is a demanding task like tracing his name, and break is provided contingent upon B.J. displaying non-compliance).

For the *developing hypothesized graph showing data patterns in different phases* trait, 1 point was awarded if each of the following three criteria present in the student graphs for each function: (a) labeled axes and a legend to describe the data paths, (b) a minimum of three data points in each phase, and (c) connected data points in each phase.

Relationship

The students' responses were also scored for consistency of information on a function among the following four questions: (a) *formulating a hypothesis regarding the function(s) of the pupil's problem behavior,* (b) *delineation of sources of data and specific information for formulating the hypothesized function (s),* (c) *a plan to test the hypothesized functions(s) of the pupil's problem behavior,* and (d) *developing graphs to display the data patterns of problem behavior in different phases of experimental manipulations.* For example, if preservice teachers identified a hypothesized function but did not develop a plan to test the function, they received a low relationship score. If the students' responses had information on all four questions/traits for all the functions they identified (even if the functions identified were incorrect) they received a full score of 4 points on the relationship component. The relationship score was included in an effort to ensure that the students were not simply "guessing" the function of behavior and also to highlight the importance of developing a plan that was cohesive. That is, there was a clear relationship among the hypotheses, the sources of supporting data, the plan to test the hypotheses, and the expected data patterns if the hypotheses were confirmed (see Table 3). For B.J..'s example, two functions were reasonable given the information provided in the case: attention and/or escape for the problem behavior (these are indicated with A and E in the Table 3). If a student identified both hypothesized functions, the student's response should then indicate sources of data and specific information considered for *both* functions, a plan to test *both* the functions, and hypothesized graphs that displayed data patterns of problem behavior for *both* functions to receive a full score of 4 points. If the student did not provide information on one of the four questions for a function, s/he received 3 points. If a student failed to provide a response to the same trait/question across two hypothesized functions or to one question when s/he identified one hypothesized function, s/he received 2 points. If a student identified two hypothesized functions and failed to provide a response on different traits/questions for different hypothesized functions or for two questions on a hypothesized function, s/he received 1 point. All other responses did not receive points on this dimension for this case (see Table 3).

Results

Reliability

The first author scored all five cases using the scoring criteria developed. Using the

TESE, **Volume 28, No. 2**
Spring 2005

Table 3. Criteria for Relationship Scores

Points	Criteria
Four	When one or more functions identified (correct or incorrect), a student's response provides information for all four questions. Examples: One function: attention (A) or escape (E) A[a]+A+A+A or E[b]+E+E+E Two functions: attention (A) and escape (E) A+A+A+A E+E+E+E
Three	When two functions are identified, a student's response lacks information on one question (O) for one of the functions. Examples: Two functions: attention (A) and escape (E) A+A+A+O[c] or A+O+A+A or A+A+O+A E+E+E+E or E+E+E+E or E+E+E+E Or A+A+A+A or A+A+A+A or A+A+A+A E+E+E+O or E+E+O+E or E+O+E+E
Two	When one of two functions are identified, a student's response lacks information (O) on a particular question. Examples: One function: attention (A) or escape (E) A+A+O+A or A+A+A+O or A+O+A+A Or E+E+O+E or E+E+E+O or E+O+E+E Two functions: attention (A) and escape (E) A+O+A+A or A+A+O+A or A+A+A+O E+O+E+E or E+E+O+E or E+E+E+O
One	When two functions are identified, a student's response lacks information (O) on two questions for a function or two different questions for each function. Examples: Two functions: attention (A) and escape (E) A+A+A+O or A+A+O+A or A+O+A+A E+O+E+E or E+E+E+O or E+E+O+E Or A+A+A+A or A+A+O+O or A+O+A+O E+E+O+O or E+E+E+E or E+E+E+E

[a] Indicates the presence of information on a question for the function "attention"
[b] Indicates the presence of information on a question for the function "escape"
[c] Indicates the absence of information for a question in a student's response

pilot group's responses, a research assistant was trained in the scoring procedures. After 90% agreement was achieved between the first author and the research assistant on 25 practice case responses, the research assistant scored 20% of the students' responses on all five cases. Two types of interobserver agreements were calculated: (a) Interobserver agreement calculated by dividing the number of agreements on the total score with the number of agreements plus disagreements and the quotient was multiplied by 100 to obtain a percentage score and (b) interobserver agreement calculated by dividing the number of agreements on each trait with the number of agreements plus disagreements and quotient multiplied by 100 to obtain a percentage score. The interobserver agreement was 100%, 86%, 79%, 83%, and 92% respectively on the total scores and 100%, 98%, 96%, 96% and 99% on the traits. Most of the disagreements occurred with the relationship scores and on the quality dimensions of traits across cases.

Analysis of Student Errors

The purpose of the revised scoring rubric was to examine the students' errors across cases and the types of errors students made so as to understand the strengths and limitations of interventions employed in the experimental study. As part of the experimental study, the student responses were examined for the consistency of mistakes across cases and the types of errors made on a posttest case. In the experimental study students' performances on the posttest FBA cases revealed a low score on the relationship trait and also a low score on the *plan for testing the hypothesized function(s)* of the student' problem behavior (i.e., antecedent or consequence analysis) trait. Further examination (across cases and on the posttest case) revealed that sometimes the errors were on the quality dimensions and sometimes on the quantity dimensions (see Pindiprolu, 2001). For example, some of the error patterns were (a) some students made the mistakes in displaying data trends in graphs that was similar across cases, (b) students had graphs that did not have a control or experimental condition, (c) students provided intervention plans instead of developing experimental conditions to confirm the hypothesis, and (d) the plans to test the hypothesized functions (quality dimension) were poor (see Pindiprolu, 2001). The revised scoring rubric thus helped in differentiating students' errors on quantity, quality, and relational dimensions, which would not have been possible with the initial 3 point scoring rubric.

Discussion

This study extends the limited published literature on scoring rubrics for assessing students' performance on FBA cases. Initially, a specific analytic scoring rubric was developed that contained five FBA traits and was pilot tested. Based on the limitations found with this scoring rubric, a generic scoring rubric was developed that guided development of case specific scoring rubrics/responses. The generic scoring rubric assessed students' responses on three dimensions: (a) quantity, (b) quality, and (c) relationship. The relationship trait measured if the students' responses were consistent in providing information on a function across four FBA traits. The results of interobserver agreement indicate that the scoring rubric facilitated consistent judgments of students' performance on FBA cases.

The practice of functional assessments or the methods to identify the functions of a behavior have been quite varied (Choi & Kim, 1998; Gable, 1996). The Amendments of IDEA do not mandate any specific methods or procedures for identifying the functions of a behavior (The Center for Effective Collaboration, 1998). However, it is generally accepted in the field of behavior analysis that multiple methods or sources of information provide more accurate information on a problem behavior than a single source or method of information. As each method has its own advantage over others. For example, indirect methods may provide information on the history of the problem behavior and past interventions that the direct and experimental analyses do not; in that case, a multi step or multi stage functional assessment is recommended (Gable, 1996; Mace, Lalli, Lalli, & Shea, 1993). The multiple steps of a functional assessment include, (a) operationally defining a problem behavior, (b) conducting indirect and/or descriptive analysis of natural conditions, (c) developing a hypothesis of possible functions the problem behavior serves, and (d) conducting experimental manipulations (Gable, 1996; Mace et al., 1993). We agree with the recommended practice. Even though it may not be necessary or feasible to conduct functional analysis in all the situations, students should be prepared to conduct them when required during their job-roles. In particular, we believe students should be given an opportunity to develop and practice analytical skills required to gather information needed to develop a hypothesis and skills to plan and test their hypothesis. As students do not get multiple real life opportunities to practice analytical skills needed to develop a hypothesis and verify their hypothesis due to time constraints, we recommend that instructors employ FBA cases (and scoring rubric) to facilitate students' preparedness to conduct multi-step functional behavioral assessments.

When FBA cases are employed as a teaching tool to teach multi-step FBA skills,

TESE, **Volume 28, No. 2**
Spring 2005

a generic scoring rubric is necessary to assess students' responses across cases. The scoring rubric should assess students' responses on all the important traits and dimensions of the task and in an objective/reliable manner. Further, the generic analytic nature of the rubric should help in pinpointing areas to target in future interventions/instruction (to instructors). Such data-based analyses, based on scores students' received on various dimensions of a generic analytic scoring rubric; can help instructors focus on the individual needs of the students and to design additional practice activities. They also can help instructors document the growth of student's skills over a period of time. This being an experimental study, the students were not given feedback or access to the scoring rubric. Also, no specific training was undertaken based on students' errors. However, instructors can employ the FBA scoring rubric that we developed as a teaching tool to highlight the important parts of a task/response and students can use the scoring rubric as a guide to solve FBA cases. Such practice may help students understand and learn complex tasks such as FBAs more efficiently.

Possible Changes that Might be Made to the Scoring Rubric

The objective of the scoring rubric developed was to measure students' analytical FBA skills, discover patterns in students' performances overtime, and examine the differential effects of the three teaching tactics. Due to experimental nature of the study the scoring scale and weights for different traits were kept uniform. However, this need not be the case and certain modifications might better suit the needs of teacher trainers. First, the FBA traits were given equal weight and this need not be the case. Teacher trainers may weigh various traits of a task differently. For example, the students' ability to describe the behavior or identify the function of a behavior or a detailed plan (quality) to test the behavior may be more important than identifying the different experimental conditions (quantity) for testing the function of a behavior. These aspects may be given more weight than the functional analysis portion (quantity) of the exercise.

Second, the scoring scale for each dimension (for the rubric developed for the study) was artificially limited. The criteria under each quality dimension might be further task analyzed and assigned points to highlight all the critical aspects. For example, when testing hypothesized functions, assigning 1 point each for (a) indicating whether the problem behavior is reinforced or ignored, (b) a list of environmental variables held constant during different experimental phases, and (c) specific environmental variables that were manipulated for testing a hypothesized function (rather than assigning 1 point for the presence of all three components). The scoring scales may also vary to accommodate the complexities of a case (cases with one problem behavior and one function vs. cases with multiple problem behaviors with multiple functions). The increase in range of the scale scores may be critical to document slight changes in student progress and performance over time.

Third, the scoring rubric has not been used as a teaching tool to improve student achievement (in that the students were not exposed to the scoring rubric). So the effects of the scoring rubric in facilitating student learning are not known. Further research needs to be undertaken to (a) identify the effects of scoring rubric in facilitating preservice teachers learning and (b) study the feasibility and usefulness of such scoring rubrics in practice.

Conclusion

With increasing focus on performance-based assessments for preservice teachers, an emphasis on employing cases to prepare preservice teachers for the complex tasks for their jobs, and the requirement to document preservice teachers' progress for NCATE and other accreditation agencies, a need exists for the development and dissemination of effective generic scoring rubrics. FBA includes a complex set of tasks and is an essential skill that special education teachers need to learn. However, it is difficult to provide enough live learning opportunities to preservice teachers to conduct FBAs during their coursework. One way to overcome the limited experiences is to provide the preservice teachers with

multiple examples of cases that they can solve in a non-threatening environment. For instructors to employ cases and document student's progress, they need to use clear scoring rubrics. This study contributes to the literature on scoring rubrics for FBA cases by (a) describing the components of a generic FBA scoring rubric and (b) providing a model for using the scoring rubric to objectively score student responses and analyze students' errors overtime.

References

Arter, J., & McTighe, J. (2001). *Using rubrics in the classroom: Using performance criteria for assessing and improving student performance*. California: Corwin Press, Inc.

Bronack, S. C., & Kilbane, C. R. (1998, March). CaseNET: *Teaching decisions via a web-based learning environment.* Paper presented at the Society for Information Technology & Teacher Educational International Conference, Washington DC.

Chandler, L. K., Dahlquist, C. M., Repp, A. C., & Feltz, C. (1999). The effects of team-based functional assessment on the behavior of students in classroom settings. *Exceptional Children, 66*(1), 101–122.

Choi, H., & Kim, U. (1998). Research digest: Functional assessment for individuals with problem behaviors. *Diagnostique, 23*(4), 227–242.

Gable, R. A. (1996). A critical analysis of functional assessment: Issues for researchers and practitioners. *Behavioral Disorders, 22*(1), 36–40.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Iwata, B. A., Vollmer, T. R., Zarcone, J. R., & Rodgers, T. A. (1993). Treatment classification and selection based on behavioral function. In R.V. Houton & S. Axelrod (Eds.), *Behavior analysis and treatment* (pp. 101–125). NewYork: Plenum Press.

Kern, L., Dunlap, G., Clarke, S., & Childs, K. E. (1995). Student-assisted functional assessment interview. *Diagnostique, 19,* 29–39.

Knirk, F. G. (1991). Case materials: Research and practice. *Performance Improvement Quarterly, 4*(1), 73–81.

Mace, F. C., Lalli, J. S., Lalli, E. P., & Shea, M. C. (1993). Functional analysis and treatment of abrrrant behavior. In R.V. Houton & S. Axelrod (Eds.), *Behavior analysis and treatment* (pp. 101–125). New York: Plenum Press.

O'Neill, R. E., Horner, R. H., Alnbin, R. W., Storey, K., & Sprague, J. R. (1990). The functional analysis interview. In R. H. Horner, J. L. Anderson, E. G. Carr, G. Dunlap, R.L. Loegel, & W. Sailor (Eds.), *Functional analysis: A practical assessment guide* (pp. 10–13). Eugene: University of Oregon Press.

Pindiprolu, S. S. (2001). *The effects of three case-based teaching tactics on the generalization of functional assessment skills with preservice teachers.* Unpublished doctoral dissertation, Utah State University, Logan.

Pindiprolu, S. S., Peterson, S.M., Rule, S., & Lignugaris/Kraft, B. (2003). Using web-mediated experiential case-based instruction to teach functional behavioral assessment skills. *Teacher Education and Special Education, 26*(1), 1–16.

The Center for Effective Collaboration and Practice. (1998). *Addressing student problem behavior: An IEP team's introduction to functional behavioral assessment and behavior intervention plans.* [On-line]. Available: http://www.air.org/cecp/fba/problembehavior/main.htm

Umbreit, J. (1995). Functional assessment and intervention in a regular classroom setting for the disruptive behavior of a student with attention deficit hyperactive disorder. *Behavioral Disorder, 20*(4), 267–278.

Ward, S. M. (1998*). Functional behavioral assessment: An evaluation of continuing education variables related to support staff use in school settings.* Unpublished doctoral dissertation, Iowa State University, Ames.

Sekhar S. Pindiprolu, East Tennessee State University; Benjamin Lignugaris/Kraft, Utah State University; Sarah Rule, Utah State University; Stephanie Peterson, The Ohio State University; & Timothy Slocum, Utah State University.