A Guiding Framework for Integrating the Consultation Process and Behavior Analytic Practice in Schools: The Treatment Validation Consultation Model

Melissa N. Andersen, Kristi L. Hofstadter, Sara Kupzyk, Edward J. Daly III
Amanda A. Bleck, Alyssa L. Collaro, Kristin E. Jones,
Lynne F. Clure and Carrie A. Blevins

Abstract

Since behavioral consultation (BC) was introduced over 30 years ago, the core procedures and assumptions of the model have remained largely unchanged in spite of its widespread popularity in practice. This article presents the Treatment Validation Consultation (TVC) model, which addresses limitations of BC and integrates recent technological innovations that are lacking in the traditional BC model. Two detailed cases are described to illustrate the procedures associated with the TVC model. Guidelines for evaluating the consultation process and outcome variables across cases are provided as a framework for consultants to self-evaluate their services as a part of the TVC model. Finally, aggregated data across five consultant trainees are presented to illustrate how to validate consultation services across cases and offer preliminary data on outcomes associated with the model.
Keywords*: behavioral consultation, school interventions, applied behavior analysis, treatment evaluation, consultation evaluation

Introduction

Consultation is the provision of psychological services to clients through formation of problem solving partnerships with important stakeholders who, through collaboration and dissemination of consultant knowledge, serve as agents of behavioral change (Elliott & Sheridan, 1992). Verbal behavior plays a primary role in this form of service delivery, as treatment plans are developed and implemented through a series of verbal interactions (e.g., interviews). Consultation models attempt to structure verbal interactions to achieve a level of consistent implementation among both experienced and novel consultants (Noell & Witt, 1996). While interviews and other verbal behaviors are necessary for the exchange of relevant and necessary information, additional critical activities (e.g., accurate selection of target behaviors and behavioral function, reliability of data collection, treatment integrity) impact the success of the process.

Applied Behavior Analysis (ABA) provides a useful framework for consultation and the development of effective treatments in school settings. Behavioral principles and procedures guide the development and implementation of treatments, as emphasis is placed on operationally defining behaviors, systematically collecting data, designing treatments that address the environmental variables maintaining the target behaviors, and making data-based decisions. As a consultation model, Behavioral Consultation (BC) generally follows these principles and practices and can be used by consultants as a vehicle for facilitating treatment implementation by a third party (i.e., the consultee). However, BC is not as closely aligned with state-of-the-art behavior analytic practices as it could be. Furthermore, behavior analysts have shown little interest to date in how verbal exchanges between a consultant and consultee might be structured to strengthen behavioral treatments (and the evaluation thereof) provided by consultees who may have little knowledge of principles and practices of ABA. The purpose of this article is to describe how BC might be improved upon to more closely align its implementation with basic principles of ABA and incorporate recent technological innovations. The model presented is called the Treatment Validation Consultation (TVC) model. Following an overview of the model, two case studies are presented to illustrate how the TVC can be applied in school settings. Finally, a model for evaluating

outcomes across cases is briefly described and results for five cases carried out by graduate trainees in consultation are presented. The results are intended to provide preliminary data in support of the validity of the TVC model and serve as an example of how practitioners can evaluate their own consultation services.

Behavioral Consultation Model Limitations

Behavioral Consultation (BC) emerged as an alternative to mental health consultation and organizational development consultation in the 1970s and was standardized in 1977 with Bergan's seminal book (Bergan, 1977). Today, BC is the prevailing model of consultation used in school settings. Research indicates that BC is more effective in addressing children's academic and behavioral difficulties than other consultation models (Erchul & Sheridan, 2008), leads to stronger child outcomes, reduces referral rates for special education, and increases the referral-to-placement ratio (DuPaul & Eckert, 1997; Erchul & Martens, 2002; Martens & DiGennaro, 2008).

The BC problem-solving model described by Krotochwill and Bergan (1990) consists of four stages: problem identification, problem analysis, treatment implementation, and treatment evaluation. Problem identification involves specifying the problem by conducting a problem identification interview, identifying procedures to measure the target behavior, and determining whether or not a discrepancy exists between current and desired levels of problem behavior. Problem analysis consists of identifying environmental variables related to the problem behavior and developing a treatment plan to address the problem by conducting a problem analysis interview. During the treatment implementation stage, the consultee implements the treatment and continues data collection while the consultant ensures integrity of implementation through brief checks (i.e., phone calls, direct observation). Lastly, in the treatment evaluation phase, the consultant and consultee meet for a final interview (the treatment evaluation interview) to determine the effectiveness of the treatment based on the data collected by the consultee. For additional information regarding the BC model, readers are referred to Kratochwill and Bergan (1990).

Despite the popularity of BC, researchers have begun to challenge the assumptions of BC and note limitations of the model (Witt, Gresham, & Noell, 1996; Noell & Witt, 1996; Watson, Sterling, & McDade, 1997). Widespread use of BC in schools may have been perpetuated by a high degree of procedural standardization, resulting in stagnation toward addressing the limitations through experimental research. Among the limitations of the current BC model are (a) an overreliance on consultee verbal report, (b) the requirement of long interviews prior to treatment selection, and (c) the underuse of recent technological innovations to inform treatment planning.

A fundamental flaw of BC is the heavy reliance on consultee verbal reports of behavior (Witt, Gresham, & Noell, 1996). Collection of assessment data primarily through verbal interactions significantly increases the likelihood of biased, incomplete results (Noell & Witt, 1996). Thus, decisions regarding problem identification, behavioral function, classroom contingencies, treatment planning, and baseline and outcome data become susceptible to error. In contrast, direct assessment contributes objectivity and validity to the consultation process, and can lead to different conclusions about behavioral function and treatment selection than what might be decided based on verbal report alone (Noell & Witt, 1996).

Problem identification is the most important—and also most complicated—component of the BC process (Witt & Elliott, 1983). If the problem is not adequately defined in specific behavioral terms, plan implementation and problem resolution may suffer. In BC, consultants enter initial interviews with little or no direct knowledge of the client, environmental contingencies influencing client behavior, or typical

behavior displayed by peers in the class. Lack of this essential information places consultants at a disadvantage, and makes application of behavioral knowledge particularly difficult.

In BC, the consultee is presented with the task of target behavior selection. Selecting target behaviors in this manner can be problematic, because it assumes that consultees possess the knowledge and skills necessary to identify the critical problem and that their selection will be consistent over time (Witt, Gresham, & Noell, 1996). Consultees' verbal reports of problem behaviors are influenced by a number of factors (e.g., events of the day, variable levels of tolerance for different behaviors). Careful, data-based target behavior selection increases the likelihood that related problem behaviors will decrease as desired behaviors increase over time, and simultaneously decreases the need for additional interviews. Rather than selecting target behaviors based on consultee report and opinion, selection should be based on objective data, and functional or habilitative criteria (Noell & Witt, 1998).

The BC model also relies on consultees for baseline data collection, as well as reporting the environmental events that may be maintaining the problem behavior. Many consultees do not have the training and experience in data collection necessary to accurately and reliably measure behavior or identify controlling variables. To increase the likelihood of developing a treatment that will result in the desired outcome, decisions should be based on objective data collected through direct observation, functional assessment, and other valid methods for gathering information. Verbal exchanges between the consultant and the consultee certainly play a major in the process (Kratochwill, Bergan, Sheridan, & Elliott, 1996); however, the presentation of multiple sources of objective data gathered by trained personnel prior to consultative meetings can only strengthen the verbal interactions within the consultation process, encourage problem-solving, legitimize the consultant's rationale for target behavior and treatment recommendations, and increase the likelihood of including components in the treatment that will result in positive outcomes (Noell & Witt, 1996).

Another limitation of the BC model is the requirement of two time-consuming interviews *prior* to treatment selection. Teachers and school-based consultants alike have limited time and resources, making engagement in lengthy interviews prior to treatment implementation both tedious and inefficient. Furthermore, when a selected treatment does not solve the problem, the consultant and consultee must return to prior steps (problem identification or problem analysis) to revalidate the target problem and select a new treatment, further increasing time requirements. Consultation processes used in schools must become more streamlined to meet the needs of consultees and clients in an efficient manner.

A final limitation to the BC model is an underutilization of recent technological innovations in the field of BA. Since the inception of BC, the field of ABA has evolved and produced many useful technological innovations that have potential for enhancing the reliability and validity of target selection and treatment planning. Recent technological innovations such as curriculum-based measurement (CBM), stimulus-preference assessments, performance deficit analysis, and brief experimental analysis (BEA), can provide consultants and consultees with better information to strengthen target behavior and treatment selection. Academic measures (e.g., CBM) and forms of analysis of academic performance (e.g., BEA) are important components of assessment because students displaying problematic behavioral excesses often possess academic skill deficits (Reid, Gonzalez, Nordness, Trout, & Epstein, 2004). Consultants may also conduct brief functional analyses to test hypotheses of behavioral function when appropriate.

Treatment Validation Consultation Model

The Treatment Validation Consultation (TVC) model is consistent with many key characteristics of BC. The overarching goal of BC and TVC is one in the same: to bring about change in the client's overt behavior (Kratochwill & Bergan, 1990). It is often the case that this change is brought about by

changing the behaviors of the consultee. Change in client and consultee behavior is produced by treatments developed through the problem-solving process (Kratochwill & Bergan, 1990). The consultant and consultee conjointly operationally define problem behaviors in overt and measurable terms, establish the existence of a discrepancy between the client's current and desired level of performance, identify environmental factors related to the problem behaviors, and select and operationalize research-based behavior management strategies to address the problem behavior. The consultee implements the treatment and repeatedly measures the client's progress on an ongoing basis, while the consultant provides the consultee with necessary training, education, and support to implement the treatment with integrity through brief checks. Finally, the consultant and consultee both evaluate the treatment effects based on the data collected by the consultant and consultee during treatment implementation (Kratochwill & Bergan, 1990).

Table 1 Problem Validation Objectives and Procedures

Objective	Methods to accomplish objective
Identify appropriate target behavior	▪ Gather information from the referral form and brief meeting with the consultee 　o Client background 　o Brief description of the problem behavior(s) 　o Description of appropriate behavior(s) in which the client should be engaging when exhibiting problem behavior(s) 　o Client schedule highlighting problematic times 　o Previous accommodations to address problem behavior(s) 　o Survey of potential reinforcers consultee is willing to use ▪ Complete relevant assessment activities (see Table 3) ▪ Prioritize behavioral targets in terms of habilitative value to students, construct validity, likelihood of affecting other behaviors (keystone behaviors), and ease of data collection ▪ Conduct a problem validation interview to present assessment data, validate, operationally define the problem, and set a goal
Select a method to measure the target behavior	▪ Consider use of data collection method used during assessments ▪ Consider efficiency and ease of data collection for consultee (i.e., use of permanent products when possible) ▪ Conduct a problem validation interview to finalize data collection method, including what, who, and how often data will be collected
Identify potential treatment components	▪ Based on assessment information consider the type of problem (see Figure 3) and antecedents and consequences to behavior that will be essential to the treatment plan and how they will fit into the classroom environment 　o Antecedents to behavior: prompting, manipulating discriminative stimuli and/or establishing operations, offer choice 　o Consequences to behavior: arrange delivery (e.g., differential reinforcement), schedule (e.g., variable ratio), and type of reinforcement (e.g., social), performance feedback, error correction, punishment (e.g., response cost)
Develop a treatment plan	▪ Conduct a problem validation interview to present essential plan components with a rationale and various ways they can be arranged in the classroom, and develop a treatment plan protocol including steps, who will implement the plan, and when it will be implemented ▪ Plan to support treatment integrity by providing training to the consultee

if needed, providing the consultee with the treatment protocol checklist,
and planning observations and feedback sessions

The TVC model differs from BC in that treatments are selected after a single interview, and data collection is expanded to include direct assessment and functional assessment. The TVC model consists of three phases: (a) problem validation, (b) treatment validation, and (c) treatment evaluation. The purpose of the problem validation phase is to identify a valid target behavior and to develop a functionally relevant treatment plan. The treatment validation phase consists of briefly testing and making data-based judgment of its feasibility and probable effectiveness in extended treatment implementation. Treatment evaluation is then conducted to evaluate the validated or revised treatment plan, and to develop a plan for post-implementation. An overview of the phases of the TVC model is presented in Figure 1. The following is a detailed overview of the objectives and methods of each phase.
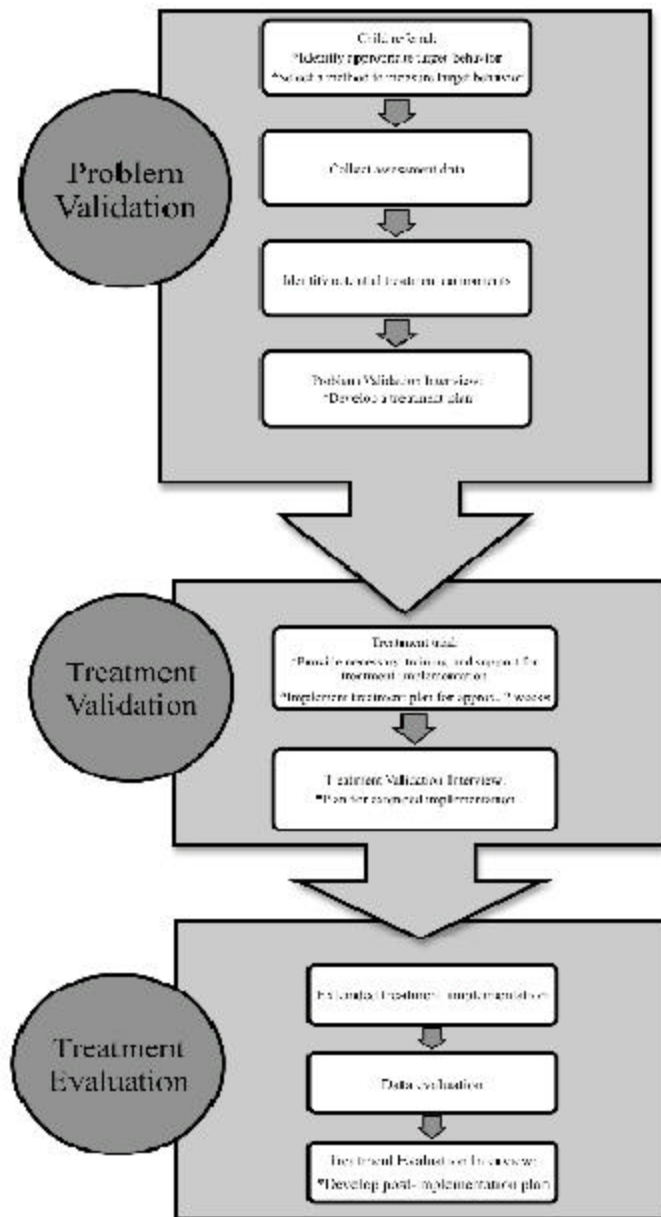

Figure 1. Stages and objectives of the TVC model.

Problem Validation

In the problem validation phase, the primary objectives are to identify the appropriate target behavior in objective and measureable terms, select a method to measure the target behavior, identify potential treatment components based on directly assessed, objective data, and develop a treatment plan. To achieve these objectives, the consultant gathers information from a referral form, collects a variety of data, and conducts the problem validation interview with the consultee. See Table 1 for an overview of the objectives and methods used to achieve each objective. At the outset of each case, the consultee completes a short referral form providing the consultant with important background information, a brief description of the consultee's conceptualization of the problem(s), previous treatments and accommodations attempted with the client, and a survey of incentives the consultee is willing to use. Upon reviewing the referral form, the consultant follows up with the teacher in a brief meeting of approximately 10 minutes to provide the consultee with an overview of the process, explain the roles of the consultant and consultee, and clarify any questions that arise from the information provided in the referral form (see Table 2).

Table 2 Objectives of the Brief Meeting

Objective	Explanation
Provide an overview of the consultation process	▪ Conduct assessments (observations in the classroom, basic academic skills, stimulus preference, etc) ▪ Meet in two weeks to review assessment data and develop plan ▪ Put plan in place for two weeks ▪ Meet to review progress and make any changes if necessary ▪ Meet for a final evaluation meeting to review progress
Explain roles of the consultant	▪ Provide support to the teacher ▪ Conduct assessments ▪ Observe plan implementation and provide feedback
Explain roles of the consultee	▪ Carry out the plan ▪ Participate in meetings
Clarify information from the referral	▪ Summarize the reported problem ▪ Ask any clarifying questions (e.g., what students were supposed to do on a math assignment, what a behavior looks like, etc.) ▪ Create a tentative operational definition of the problem behavior for direct observations

One of the key differences between the TVC model and BC is that the consultant administers a number of direct assessments with the client *prior* to the first interview. Thus, the consultant and consultee enter the initial interview empowered with a more objective understanding of the target behavior, maintaining variables, and potential treatment components (Noell & Witt, 1996). Furthermore, these data may also serve as baseline data, allowing treatment implementation to begin immediately after the first interview. The assessment activities include direct observation of academic engagement and problem behavior, CBM of academic skills, examination of work products, a stimulus preference assessment, and a performance deficit analysis. If the consultant does not gain sufficient information about the problem from the initial set of assessments, a brief functional analysis or brief experimental analysis may also be included. See Table 3 for a brief description of each assessment.

Table 3 Assessment Activities

Assessment activity	Description	Question to guide interpretation
Curriculum based measurement	Brief, standardized tool for measuring the basic skills of oral reading fluency, math computation fluency, spellin g fluency, and writing fluency	Is there a discrepancy in comparison to peers?
Work sample analysis	Collecting and assessing permanent products (e.g., worksheets, assignments) for completion, accuracy, and skill level	Completion problematic? Accuracy problematic?
Direct observation	Systematic observation of problem or target behavior in the environment using interval (provides percentage of intervals student engages in behavior), time-sampling, and/or A-B-C recording (information about variables that occasion and maintain problem behavior)	Is there a discrepancy in comparison to peers? What variables appear to be maintaining the problem behavior?
Performance deficit analysis	Method that differentiates whether the student possesses a skill or performance deficit	Is it a performance or skill deficit?
Stimulus preference assessment	Method for identifying multiple stimuli that may serve as effective reinforcers based on student selections from an array of stimuli	Which stimuli are most likely to be effective reinforcers?
Brief functional analysis	Systematic manipulation of variables to identify behavioral function or potential treatments	Which variables appear to be maintaining the problem behavior? Which treatment(s) resulted in the greatest improvement?
Brief experimental analysis	Tool for testing potential academic treatments for academic deficits	Which component(s) led to the greatest performance improvement?

Selection of the assessment activities to administer is determined based on individual case referral concerns. When a client is referred for problem behavior in the classroom, the consultant conducts direct observations of academic engagement and/or problem behavior, and when relevant, and A-B-C analysis during times in which the behavior(s) typically occur. Additionally, the consultant assesses the skills that the client is expected to demonstrate when he/she typically engages in the problem behavior (e.g., math computation skills). When a client is referred for problem behavior outside the classroom (e.g., halls,

lunchroom, or playground), the consultant conducts direct observations of the problem behavior, an A-B-C analysis, and, if necessary, a brief functional analysis to generate a hypothesized function for the problem behavior. When a client is referred for an academic concern and the teacher has no concern regarding other behavior problems, the consultant assesses the client's academic skills in the area(s) of concern. The following is a description of each assessment activity in which the consultant may engage during the problem validation phase.
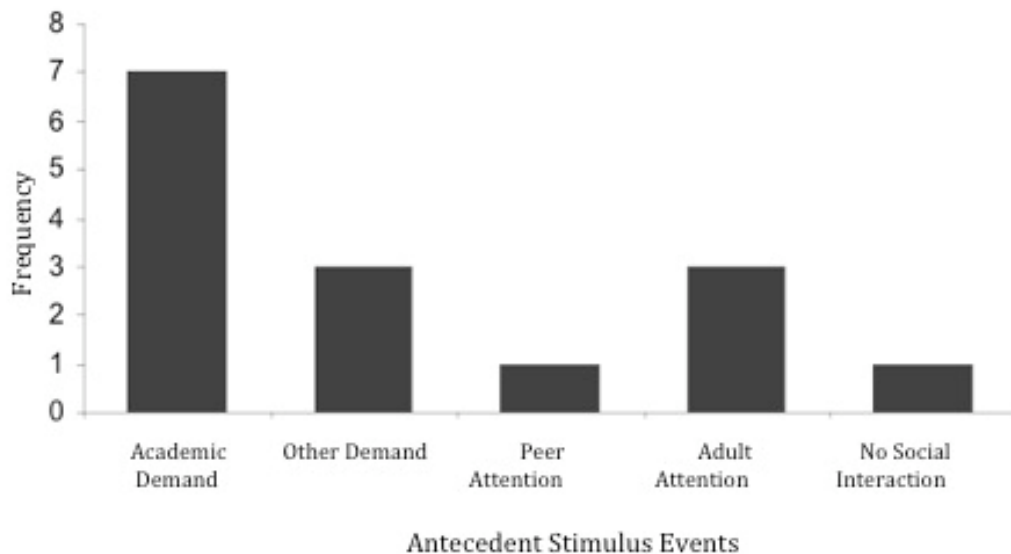

Figure 2. Graphic representation for observed antecedents to a problem behavior.

Direct assessment. Direct assessment methods include CBM, work sample analysis, direct observation, and A-B-C recording. CBM is a brief, standardized tool that provides valid, reliable indicators of a client's academic competence (Fuchs, Fuchs, & Hamlett, 2007). CBM can be used to measure a client's level of performance in basic skills such as reading, math computation, spelling, and written expression (Shinn, 1989). Research supports the use of CBM to evaluate a student's performance in relation to the curriculum and to class peers, to monitor a student's progress over time, and to determine whether a student is in need of additional instruction (Deno, 1985; Deno, 2003; Gettinger & Seibert, 2000).

Curriculum-based oral reading fluency probes are administered to measure the rate at which a student can read words correctly in connected text. Clients are presented with a grade level reading probe consisting of a story containing approximately 100 to 250 words. Clients are asked to read the story aloud for 1 minute. The total number of correctly read words is calculated by subtracting the number of words read incorrectly from the total number of words read within 1 minute. Mispronunciations, omissions, substitutions, or words not read within 3 s are considered errors (Shinn, 1989).

Curriculum-based math probes are administered to measure a client's mathematic s computation skills. Depending on the client's grade level, clients are asked to complete as many addition, subtraction, multiplication or division problems as possible within a 2 to 4 minute time period. The number of digits correct (DC) is calculated by counting the number of digits the client writes correctly (Shinn, 2004).

Curriculum-based writing probes are administered to examine writing accuracy and production. Clients are presented with a written expression probe consisting of a one-sentence story starter, asked to think about what to write for 1 minute, and then write for 3 minutes. The total words written (TWW) and percentage of correct word sequences (% CWS) are calculated for all clients. Correct word sequences are calculated by identifying consecutive, correctly spelled words, which are grammatically correct in the context of the sentence (Videen, Deno, & Marston, 1982). The % CWS is calculated by dividing the number of correct word sequences by the total number of word sequences within each writing sample.

Additionally, the consultant collects peer comparison data to serve as a barometer of the level of the client's functioning compared to his/her peers. Peer comparison data can be acquired from national norms, school benchmark data, district or classroom norm data, or direct administration by the consultant. The consultant creates graphic displays of the client and peer data to facilitate data-based decision-making regarding the discrepancy between current and expected performance.

One convenient and easy method of assessment is to measure by-products of a client's behavior with permanent products (e.g., math worksheets). When a client exhibits problem behavior with work completion, or during particular subject areas that produce work samples, the consultant can arrange for the consultee to gather samples of permanent products to analyze. A consultant can quickly and easily conduct a work sample analysis by measuring the percent of problems completed (number of problems attempted divided by the total number presented), or the percent of problems completed accurately (number of correct responses divided by total number of problems presented that were completed).

Direct observation. In the problem validation phase, the consultant directly observes academic engagement and problem behavior for the client and a peer of the same gender during the client's problematic times. Academic engagement refers to behaviors that indicate active engagement in tasks, including reading aloud, answering or asking an academic question, looking at the teacher during class lecture, writing in response to a teacher's request, silent reading as indicated by scanning eye movements, and/or actively writing responses to math problems. Academic engagement measured because it is associated with higher levels of achievement (Cobb & Hops, 1973; DeBaryshe, Patterson, & Capaldi, 1993; Derevensky, Hart, & Farrell, 1983; Gamoran & Nystrand, 1991). Students with higher academic engagement experience more opportunities to respond to instruction, and are more likely to improve their academic performance. Additionally, academic engagement is incompatible with many problem behaviors, and can serve as a replacement behavior.
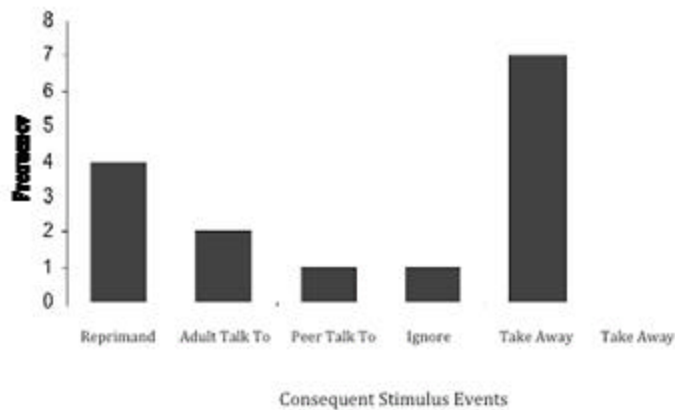
Figures 3. Graphic representation for observed consequences to a problem behavior.

Time-sampling recordings are useful methods for simultaneously measuring academic engagement and problem behavior. Consultants observe and record problem behaviors using 20-second partial interval recording, while simultaneously indicating the presence or absence of academic engagement at the end of each 20-second interval (i.e., momentary time sampling). Additionally a classroom peer is observed every sixth interval to provide peer comparison data.

A-B-C recording is a functional assessment approach that involves directly observing and documenting the environmental events that occur immediately before and after the client engages in a problem behavior (e.g., presentation of an academic task, teacher direction, reprimand, peer attention, etc.). Results obtained from A-B-C analysis provide information about the environmental events that are likely triggering and maintaining the problem behavior and aid in the generation of a hypothesis about the function of the problem behavior (McComas & Mace, 2000). Figures 2 and 3 illustrate antecedents and consequences that immediately preceded and followed a client's engagement in a problem behavior, respectively. In the example, the presentation of an academic demand was the most frequent antecedent, and demand removal contingent on the occurrence of problem behavior was the most frequent consequence. Based on these results, it appears that the function of the problem behavior is escape from academic demands.

Performance deficit analysis. Deficits in academic responding result from lacking either sufficient skill to exhibit the desired behavior or sufficient motivation to perform the desired task (Duhon, Noell, & Witt, 2004). A performance deficit analysis is an assessment method that differentiates whether the client possess a skill deficit, performance deficit, or a combination of both. It entails offering the client a strong incentive for meeting or exceeding a performance goal on a classroom assignment on which the client previously performed below expectations. Based on the client's performance following an incentive, a hypothesis can be generated about the type of deficit the client exhibits. If the client demonstrates a marked improvement, one would hypothesize a performance deficit. If the client fails to meet the performance goal, one would hypothesize that the client lacks necessary skills, and requires additional instruction. (Duhon et al., 2004).
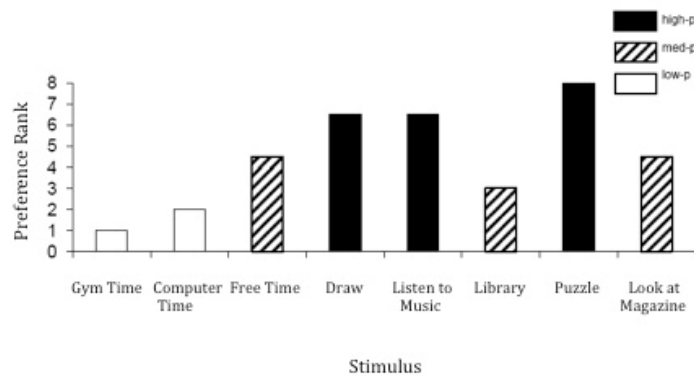
Figure 4. Stimulus preference assessment results.

Stimulus preference assessment. Stimulus preference assessments identify multiple stimuli that may serve as effective reinforcers to improve a child's academic and behavioral performance (Daly et al., in press). Stimulus preference assessment methodology has gathered substantial empirical support for use with individuals with low incidence disabilities, and is now being extended to individuals with high incidence disabilities (e.g., ADHD and behavior disorders; Daly et al., in press; Northup, George, Jones, Broussard, & Vollmer, 1996; Northup, Jones, Broussard, & Jones 1995). Multiple stimulus without replacement (MSWO) is an efficient method for differentiating high, medium, and low preference reinforcers on an individual basis. The MSWO method has identified reinforcers to increase on-task behavior of adolescent students labeled with behavior disorders (Paramore & Higbee, 2005) and math task completion with elementary students with behavior disorders (Daly et al., in press). In our practice and research, we have found common activitie s (e.g., playing basketball in the gym, playing cards with a consultant, having lunch with the teacher) to be desirable to both consultees (who like them because they are easy to deliver and do not cost anything) and clients (who like them because they provide a break from school activities and appear to be privileges). Potential reinforcers can be identified by having the consultee complete a survey indicating a variety of different reinforcers (e.g., activities, privileges, tangibles, and edibles) the consultee is willing to use in the classroom. From the consultee's list of approved reinforcers, eight items are selected and displayed or written on index cards. Actual edibles may be used if available and acceptable to the consultee. The assessment consists of placing the stimulus cards in front of the client and allowing him/her to choose one card he/she would like to work for most. The card chosen is removed from the group, and the cards are repositioned. The client selects the remaining stimulus cards in the same manner, until all cards are chosen. The consultant records the order (1-8) in which each reinforcer was selected. The consultant follows these procedures two additional times (preferably across different days) to identify a median score (1-8) for each reinforcer. Median scores are then reverse scored (i.e., a median score of 1 is given a score of 8) for purposes of data display. Rank order is then graphed for each item, with higher scores indicating greater reinforcing value (Daly et al., in press). Results can be used to develop various reinforcement programs, such as token economies. Figure 4 illustrates results from a stimulus preference assessment.

Additional functional assessment activities. In some cases, a more direct test of function or treatment may be necessary and even more efficient. Brief functional analysis (BFA) is feasible for use in schools and can aid in treatment selection (Cihak, Alberto, & Fredrick, 2007; Harding, Wacker, Cooper, Millard, & Jensen-Kovalan, 1994; Northup et al., 1991; Wallace & Knights, 2003; Wilder, Chen, Atwell,

Pritchard, & Weinstein, 2006). Brief functional analysis consists of administering brief (e.g. 5 minute) trials to test either the possible functions of a problem behavior or the possible treatments for the problem behavior. When testing the function of a behavior, brief sessions are conducted consisting of providing selected consequences (e.g., attention, escape, tangible items) contingent on the occurrence of the problem behavior. For example, in an attention condition of a BFA of the potential functions for disruptive behavior, the client would be provided with adult and/or peer attention immediately following each occurrence of disruptive behavior, while all other consequences (e.g., escape from a task, access to activities or privileges) are withheld to prevent the data from being confounded by the influence of other consequences. When testing possible treatments for a behavior, brief sessions incorporating behavioral treatments for a target behavior are administered sequentially until an effective treatment is identified (Harding, et al., 1994).

Brief experimental analysis (BEA) methodology has recently emerged as a tool for experimentally testing potential academic treatments for students exhibiting deficits in academic responding (Daly, Andersen, Gortmaker, & Turner, 2006). Since its inception, numerous studies implementing variations to the approach have demonstrated that BEAs can efficiently identify effective academic treatments for individual students (Daly et. al, 2006; Daly, Martens, Dool, & Hintze, 1998; Daly, Martens, Hamler, Dool, & Eckert, 1999; Daly, Murdoch, Lillenstein, Webber, & Lentz, 2002; Gortmaker, Daly, McCurdy, Persampieri, & Hergenrader, 2007; Jones & Wickstrom, 2002). Administration of BEAs consist of identifying several probes of the skill to be remediated (e.g., math computation, reading etc.) at the client's instructional level, establishing a baseline score, and systematically implementing research based instructional strategies (individually and/or in combination) using a different probe for each strategy/set of strategies tested to determine which strategies result in the highest increases in performance relative to baseline. Readers are directed to Daly et al. (2006) for further information on BEA application and methodology.
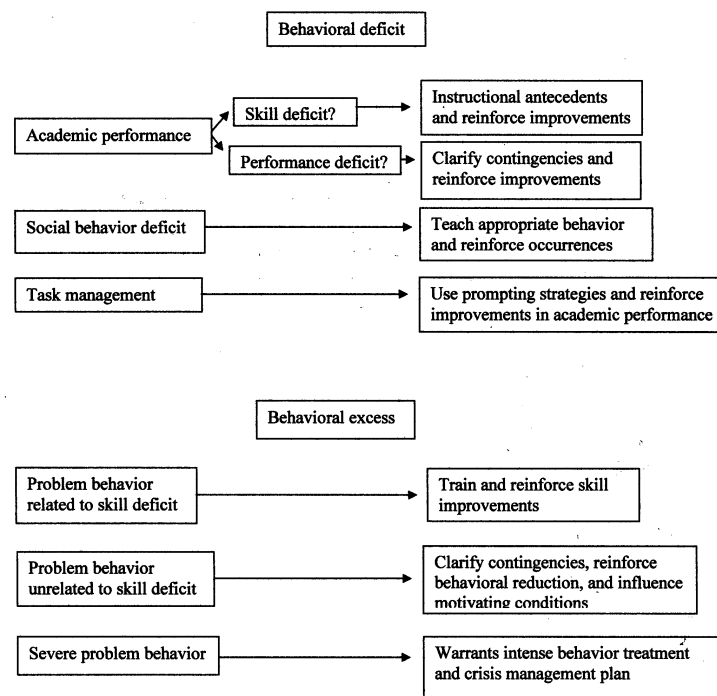

Figure 5. Treatment components by problem behavior

Data integration. Upon completion of all assessment activities, the consultant graphs and analyzes the data to determine whether the referral concern reflects a valid target behavior, the type of problem the client exhibits (e.g., academic skills deficit, behavior excess) and to identify a number of potential research-based behavioral treatment strategies prior to meeting with the teacher. Table 3 provides questions to guide the integration of assessment data in preparation for the teacher interview. Figure 5 provides treatment components to consider based on problem type.

Problem validation interview

The problem validation interview is conducted to validate the consultee's initial concerns based on the data collected, operationally define one or more target behaviors or skills for remediation, set a performance goal, generate a treatment plan, develop a plan for data collection, and create a plan to monitor treatment integrity. The desired outcome of the PVI is a measurement plan for repeatedly measuring the target behavior over time and a treatment plan that specifies the treatment steps, change agent(s), time(s), location(s), and frequency of implementation. The consultant generates a treatment protocol delineating each of the treatment steps. The protocol serves as a prompt for treatment implementation for the consultee and as a tool for monitoring treatment integrity for the consultant.

Table 4 Treatment Validation Objectives and Procedures

Objective	Methods to accomplish objective
Implement the treatment plan	▪ Complete a brief treatment trial (e.g., 2-weeks) ▪ Monitor client's target behavior using data collection plan ▪ Support and monitor treatment integrity by providing training to the consultee if needed, having the consultee complete the treatment protocol checklist, conducting observations, and providing feedback
Validate treatment plan	▪ Conduct a treatment validation interview to review the rationale for each treatment component, review data from the treatment trial (client's target behavior and consultee treatment integrity), and evaluate the effectiveness of the plan
Develop plan for extended treatment implementation	▪ During the treatment validation interview a decision is made to continue, modify, replace, or terminate the plan based on child outcome data and treatment integrity results ▪ Revise treatment plan protocol as necessary

Treatment Validation.

The primary objectives of the treatment validation phase are to complete and evaluate the effects of a brief treatment trial (e.g., two weeks), and to develop of a plan for extended treatment implementation (See Table 4). During the treatment validation phase, the consultant must be careful not to assume that the verbal interchange in the problem validation interview is sufficient to change the consultee's behavior (Noell & Witt, 1996). Therefore, the consultant trains the consultee to implement the plan through modeling and practice, and by giving feedback to the consultee based on the treatment protocol checklist. Although the consultee is responsible for participating in data collection (assisting with gathering permanent products, frequency counts, duration, etc.), the consultant oversees and takes primary responsibility for data collection and often supplements data with direct behavioral observations of academic engagement and problem behavior, when these are not the target behaviors.

After the treatment has been implemented for approximately two weeks, a treatment validation interview (TVI) is conducted to evaluate treatment effectiveness, and develop a plan for extended treatment implementation. The consultant and consultee visually analyze the outcome graphs to determine whether the client has made sufficient progress. The consultant provides the consultee with treatment integrity data and assists the consultee with interpreting the outcome data in light of the extent to which the treatment was followed as planned. Additionally, the consultant solicits the consultee's report of implementation feasibility. Based on this information, the consultant and consultee then determine together whether to continue, modify, replace, or terminate, or extend the treatment to other times and/or locations.

Table 5 Treatment Evaluation Objectives and Procedures

Objective	Methods to accomplish objective
Implement the treatment plan for an extended period	▪ Implement the treatment plan for an extended period (e.g., 4-8 weeks) ▪ Monitor client's target behavior using data collection plan ▪ Continue to support and monitor treatment integrity
Evaluate the treatment plan	▪ Conduct a treatment evaluation interview to review data (client's target behavior and consultee treatment integrity), and evaluate the effectiveness of the plan
Develop a post-implementation plan	▪ During the treatment evaluation interview a decision is made to determine whether to continue, fade, modify, replace, or terminate the plan based on client's outcome data ▪ Revise treatment plan protocol as necessary

Treatment Evaluation

The primary objective of the treatment evaluation phase is to implement the treatment for an extended period of time, evaluate treatment effectiveness, and to develop a plan to continue, modify, fade, or terminate treatment (see Table 5). To achieve these objectives, the consultant works to ensure the treatment is implemented with a high degree of integrity, and the client's progress is continuously monitored in the same manner as in the treatment validation phase.

After the treatment has been implemented for an extended period of time, the treatment evaluation interview (TEI) is conducted. The consultant and consultee evaluate the plan based on visual analysis of the data patterns, goal attainment, and treatment integrity. When inspecting the data, consultants consider the degree to which desired criterion levels of the target behavior have been met (e.g., whether the client's behavior approached the pre-established goal, met the goal inconsistently, or met the goal consistently). Additionally, the consultant inspects the trend and variability of the data. If the client's behavior has reached or exceeded the pre-established goal consistently, and the trend is stable and continues in the desired direction, then the consultant and consultee consider planning to continue treatment, fade treatment, terminate treatment, or extend the treatment to other times and/or locations. If data are variable, the consultant and consultee consider treatment integrity, reinforcement satiation, or other external variables that may be adversely affecting treatment effects. If the client's behavior is indistinguishable from the baseline level and/or trend for a skill-based treatment, then the consultant and consultee consider whether the client has the necessary prerequisite skills to be successful, whether additional antecedent

strategies such as modeling and prompting are necessary, and whether the difficulty level of instructional materials should be adjusted. If the client's behavior is indistinguishable from the level and/or trend for a contingency-based treatment, then the consultant and consultee would consider whether the reinforcement contingencies are actually motivating, whether the rewards target the correct function of behavior, and/or whether the rewards are sufficiently robust to compete with reinforcement for inappropriate behavior. Finally, if the client's behavior demonstrates a trend in the undesired direction, the consultant and consultee consider whether client is exhibiting an extinction burst, whether the treatment has been implemented as planned, whether the treatment includes potentially aversive components to the client, whether the treatment addresses the function of the behavior, and whether the treatment needs to be replaced. Once the data are analyzed, the consultant and consultee develop a plan to continue to meet the client's needs, as well as methods for continuing to support plan implementation and progress monitoring.

Table 6 Procedural Checklist

Procedural Steps

---

    1. Appropriate consents were obtained (parent, teacher, and child).

Problem Validation

    2. Multiple data sources were used to validate the target behavior, including teacher report, direct observation of behavior, basic skills assessment, performance deficit analysis, stimulus preference assessment data, and permanent products of schoolwork.

    3. The consultant and consultee defined the target behavior in objective, observable, and measurable terms, and the reliability and validity of the target behavior were established.

    4. The target behavior was measured repeatedly over time and results are displayed graphically as a baseline.

    5. One or more quantitative goals for client performance (including level of performance and a goal date) were established with the consultee.

Treatment Validation

    6. An empirically supported treatment protocol was developed based on an analysis of the natural setting.

    7. The consultant trained the change agent to implement the treatment protocol.

    8. Direct observations of treatment implementation were conducted (and summarized quantitatively as well as qualitatively), and performance was reviewed with the change agent.

    9. The consultant and consultee met during treatment implementation to discuss and plan for possible modifications to the plan.

Evaluation/Decision Making

    10. There is a specified design and evaluation plan.

    11. Conclusions regarding effectiveness were based on (a) integrity of implementation and (b)

visual inspection of client performance data.

12. Data-based recommendations for future programming were made.

Table 7 Case Examples Overview

Student	Characteristics	Referral Concern	Target Behaviors	Treatment Package
Dan	9 years old 3rd grade	Mrs. Long, indicated that Dan displayed a high rate of off-task behavior and corresponding low rate of academic engagement. She reported that the regular education tasks appeared to be too difficult for Dan, and his academic achievement was discrepant from that of classroom peers and grade-level standards in math and writing.	1. Math performance: Percentage of problems completed on teacher assigned worksheets. 2. Academic engagement: Active behaviors, including reading aloud, answering or asking an academic question, looking at the teacher during class lecture, writing in response to a teacher's request, silent reading as indicated by scanning eye movements, and/or actively writing responses to math problems. 3. Off-task behavior: Failure to comply with commands within five seconds of request, eyes averted from class materials or teacher for three seconds or longer, playing with tangible items, and/or talking to peers.	Treatment components included modeling and guided practice with performance feedback, differential reinforcement, and a plan for structured teacher prompts during math tasks to reinforce academic engagement.
Trey	7 years old 1st grade	Mrs. Burke indicated that Trey displayed tantrums and disruptive outbursts in the classroom, as well as academic performance well below grade level standards.	1. Writing performance: The number of letters written and number of identifiable words. 2. Academic engagement: Active student behaviors such as reading aloud, asking or answering academic questions, writing, and silent reading.	The treatment consisted of modeling and prompting for writing as well as daily goal setting and contingent rewards..

Case Examples

The following cases illustrate the procedures of the TVC model as well as the measures employed for evaluation both within and across cases. Examples include cases conducted by first-year graduate students in a school psychology doctoral program who were being trained to conduct case consultation following training in ABA. Each case was conducted to fulfill a consultation course requirement under the supervision of a university professor (the fourth author). An overview of the procedural checklist followed by each student consultant and pertinent case information are presented in Tables 6 and 7, respectively.

Case 1

Dan, a third-grade student, was referred by Mrs. Long, his classroom teacher, to Amanda, the consultant, for concerns regarding off-task behavior.
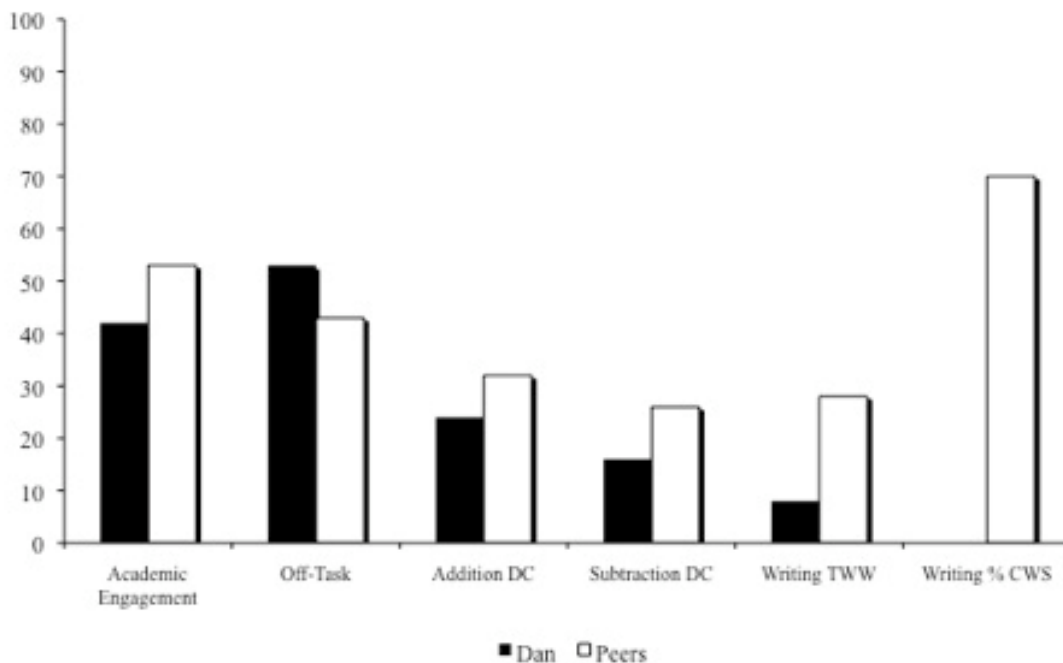

Figure 6. Observation, math, and writing assessment data for Dan and available peer comparisons.

Problem validation. To validate the problem behavior reported by Mrs. Long's, Amanda conducted a number of assessments with Dan, including direct observation during reported problematic time periods, skills assessments using CBM math and writing probes, work samples collected from Mrs. Long, and a stimulus preference assessment. Baseline observation results indicated that Dan was off-task for a mean of 53% of the intervals and academically engaged during a mean of 42% of observed intervals. Dan and four randomly selected peers were administered CBM math and writing probes to examine the potential contribution of skill level to problem behaviors. Dan received a mean score of 24 DC across addition probes, and his peers received a mean of 32 DC across addition assessments. On subtraction probes, Dan received a mean of 16 DC, and his peers received a mean of 26 DC. On writing probes, Dan

received a score of 8 TWW and 0% CWS, compared to the peer mean of 28 TWW and 70% CWS. Figure 6 displays the observation, math, and writing data for Dan and available peer comparisons. Additional math and writing work samples were collected from Mrs. Long and were scored and graphed to examine pre-treatment performance levels across time. A stimulus preference assessment was also conducted, and identified reinforcers were included in an individualized reward menu for Dan.
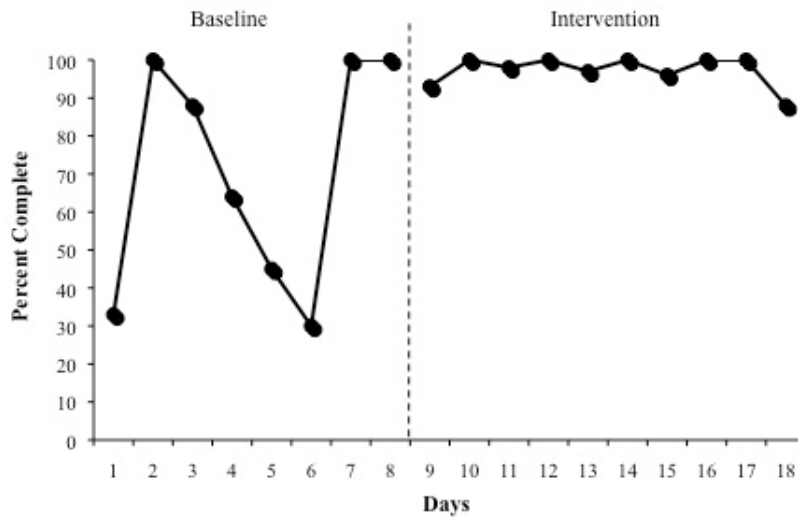

Figure 7. The percentage of math worksheet completion for Dan across baseline and treatment conditions.
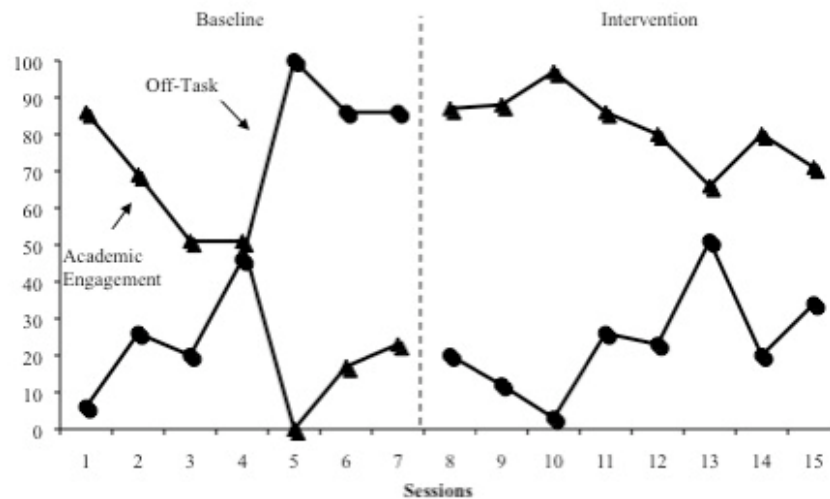

Figure 8. The percentage of intervals academically engaged and off-task for Dan across baseline and treatment conditions.

Examination of data during the problem validation interview suggested that Dan displayed a high rate of off-task behavior and corresponding low rate of academic engagement, primarily during mathematics tasks prior to treatment (Figures 7 and 8). Teacher prompting was frequently observed to

follow off-task behavior, and it was hypothesized that social attention maintained the high rates of off-task behavior. Additionally, Dan's math and writing performance was discrepant from that of peers, indicating that skill deficits may have been contributing to the corresponding low rates of academic engagement. Due to a higher rate of teacher-reported and consultant-observed off-task behavior during mathematics tasks, math performance, academic engagement, and off-task behavior were selected as target behaviors. Based on the hypothesis that Dan's poor academic performance was being maintained by a skill deficit in mathematics and teacher attention for off-task behavior, a treatment was developed to improve mathematics fluency, increase academic engagement, and decrease off-task behavior by employing a reversal of baseline contingencies. Instructional strategies (i.e., modeling, guided practice, performance feedback) were provided to improve math fluency, and academic engagement was differentially reinforced through teacher attention. The treatment protocol is displayed in Table 8.

Table 8 Case 1 Treatment Protocol

Steps

1. Give math assignment and instructions and let him choose which reward he wants to work for that day.

2. Demonstrate how to do the first two math problems.

3. Have Dan do the next two problems under your supervision. Provide feedback and praise responses and effort.

4. Ask him to do a third of the problems on the sheet (specifying how many problems and what number to work to) and to raise his hand to call you over to check his work. Provide feedback and praise.

5. Ask him to do the next set of problems, and each time he calls you over, praise and provide feedback.

Mrs. Long and Amanda set the following goal for Dan: When presented with a math worksheet containing approximately 15 to 20 problems, Dan will complete 100% of problems during the 75-minute math period with a maximum of four teacher directed prompts to do his work. This goal was set with an expected attainment date of 5 weeks from the problem validation interview.

Treatment validation. Following the two-week treatment trial, Mrs. Long indicated that the treatment was easy to implement and seemed to be effective. Examination of outcome data (i.e., A-B graphs) provided validation for the effectiveness of the treatment (see Figures 7 and 8). Integrity data revealed a high level of adherence to treatment steps, and consultant integrity check data revealed a high level of agreement between consultant and consultee evaluation of treatment fidelity. Based on treatment validation data, Mrs. Long and Amanda decided to continue implementation of the current plan without modification.

Treatment evaluation. Collection of outcome and integrity data continued throughout extended implementation of the treatment. The percentage of correctly completed math problems and the

percentage of intervals in which off-task behavior and academic engagement were observed across baseline and treatment conditions is displayed in Figures 7 and 8, respectively. Evaluation of Dan's response to treatment revealed that math performance increased from a mean of 70% complete (range, 30% to 100%) during baseline to a mean of 97% complete (range, 88% to 100%) during treatment implementation. Dan's academic engagement increased from a mean of 42% (range, 0% to 86%) during baseline to a mean of 82% (range, 66% to 97%) during treatment implementation, and Dan's off-task behavior decreased from a mean of 53% (range, 6% to 100%) to a mean of 24% (range, 3% to 51%) in response to the treatment.

Mrs. Long indicated that treatment implementation had been easy and successful, and examination of Dan's math performance revealed that his goal had been attained. Based on the effects achieved with Dan, Mrs. Long stated that she planned to use the treatment with another student in her class. She and the consultant also planned for generalization of the treatment to other academic areas that were problematic for Dan (e.g., writing tasks). Amanda provided Mrs. Long with a post-treatment measure of acceptability, and the mean item rating was 5.8 (out of a possible 6), suggesting highly acceptable procedures.

Case 2

Trey, a first-grade student, was referred by Mrs. Burke, his classroom teacher, to Lynne, the consultant, for concerns regarding academic performance and high rates of disruptive and aggressive behavior.
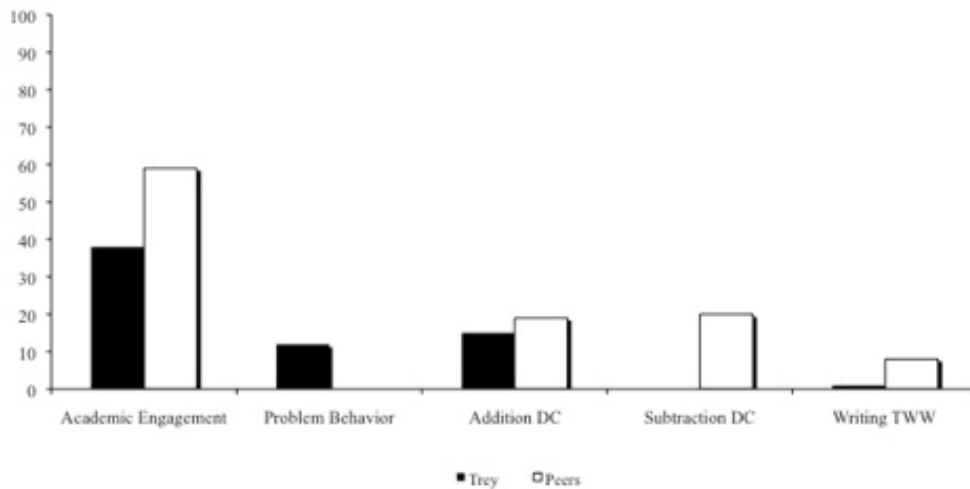

Figure 9. Observation, reading, math, and writing assessment data for Trey and peers.

Problem validation. Lynne conducted systematic direct observations to validate Mrs. Burke's reported concerns and determine the rate of Trey's academic engagement and problem behavior. Results indicated that Trey was academically engaged for a mean of 38% of intervals, compared to his peers, who were engaged a mean of 59% of the intervals. Trey displayed problem behavior for a mean of 12% of observation intervals, compared to 0% for his peers. Curriculum-based measures (CBM) of mathematics and written expression indicated that Trey scored a mean of 1 TWW across writing probes, 15 DC on addition problems and 0 DC on subtraction probes, compared to classroom peer means of 8 TWW, 19 DC on addition problems, and 20 DC on subtraction problems. Direct observation, math, and writing data are displayed in Figure 9.
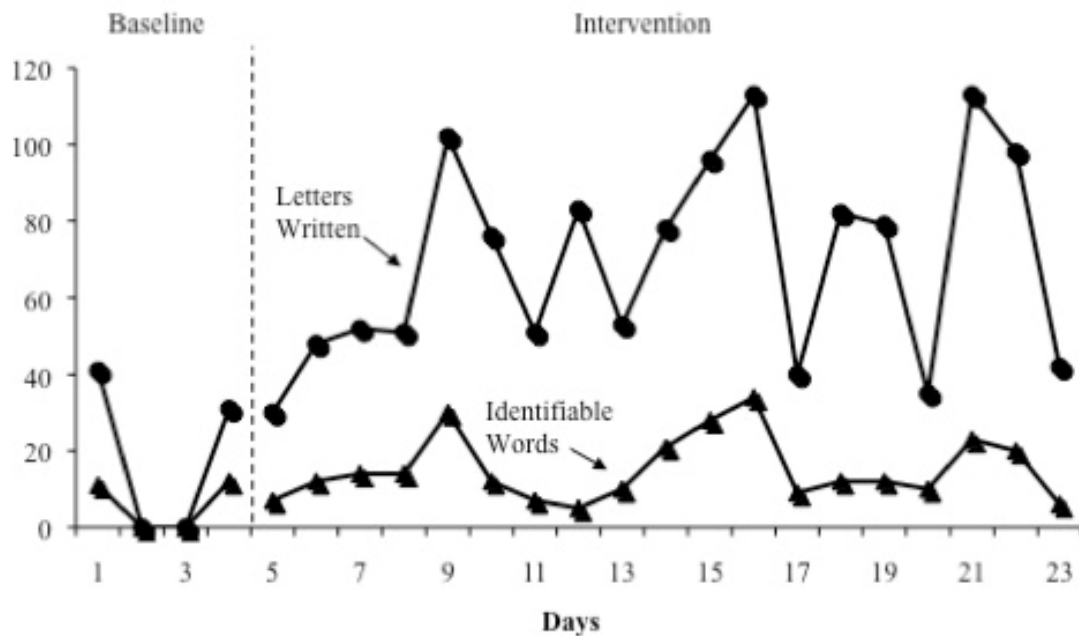
Figure 10. The number of letters and identifiable words written for Trey across baseline and treatment conditions.

A review of data during the problem validation interview provided support for the concern that Trey's academic achievement was discrepant from that of peers and grade level standards. According to Mrs. Burke's documentation, Trey's problem behavior occurred at least 13 times over a 2.5 week period, approximately 4 times per week. Trey's work samples also indicated that his writing production was consistently below expected levels on daily writing tasks (Figure 10). Furthermore, Trey's academic engagement was below expected levels of performance (Figure 11). Based on these data, the target behaviors identified were academic engagement and writing performance. It was hypothesized that Trey's deficits in academic engagement and writing were maintained by both escape from difficult tasks (writing) and teacher attention for inappropriate behavior, as competing problem behaviors often resulted in both escape from academic tasks and teacher attention in the form of reprimands. Thus, the treatment consisted of instructional components (i.e., modeling and prompting) to remediate Trey's skill deficits and programmed antecedents and consequences (i.e., goal setting and contingent reinforcement in the form of attention) to increase desired responding (i.e., writing and academic engagement). The consultant and consultee set the following goal: When prompted to write for 15 minutes, Trey will write an average of 40 letters and 12 identifiable words, and his academic engagement will increase to 75% during writing tasks. The goal date was set for 5 weeks after the problem validation meeting.
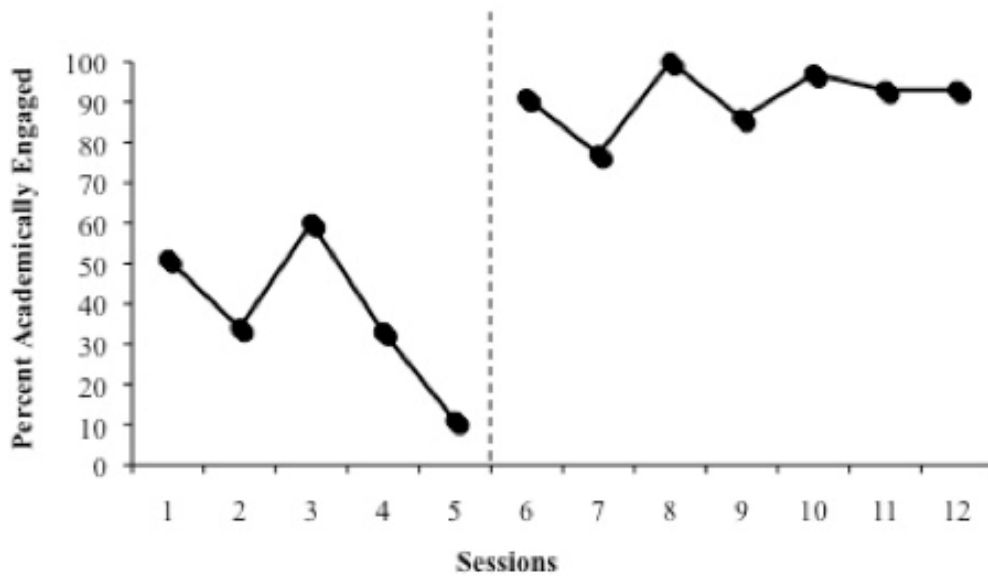
Figure 11. The percentage of intervals academically engaged for Trey across baseline and treatment conditions.

Treatment validation. After the 2-week treatment trial, a review of the data provided validation for treatment effectiveness. A high level of treatment adherence was reported by the teacher and substantiated by the results of treatment integrity checks by the consultant. Mrs. Burke and Lynne decided to continue implementation of the current treatment without modification.

Treatment evaluation. The treatment evaluation interview included an examination of writing performance and academic engagement using visual inspection of A-B graphs. Final results for the number of letters and identifiable words written as well as the percentage of intervals in which academic engagement was recorded across baseline and treatment conditions are displayed in Figures 10 and 11, respectively. Evaluation of Trey's response to treatment revealed that the number of letters written increased from a mean of 18 (range, 0 to 41) during baseline to a mean of 70 (range, 30 to 113) in response to treatment. Trey's academic engagement increased from a mean of 38% of intervals (range, 11% to 60%) during baseline to a mean of 91% (range, 77% to 100%) during treatment implementation.

Mrs. Burke and Lynne examined Trey's graphs and found that he made some progress toward the goal, although goal levels of performance were not fully attained. Mrs. Burke reported plans to send the treatment plan to Trey's second grade teacher for implementation during the following academic year. Lynne provided Mrs. Burke with a post-treatment measure of acceptability, and the average item rating was 4.6 (out of a possible 6), suggesting moderate acceptability of procedures.

Consultation Model Evaluation

It is particularly difficult to untangle the myriad of variables impacting consultation outcomes, making the task of experimental validation complex and challenging. The positive outcomes associated with consultation (Erchul & Sheridan, 2008) have not systematically ruled out the possible influence of other variables. For example, the treatment itself is a confounding variable in attempts to validate the contribution of the verbal interchange between consultant and consultee to client or consultee outcomes (Witt, 1997). However, the current impetus toward accountability necessitates measures designed to

evaluate all educational practices, whether they meet rigorous experimental criteria or not. In the consultation literature, both process and outcome variables are typically examined in an effort to determine whether the model employed contributed to positive outcomes for consultation consumers. Early evaluation research focused on consultation training and primarily measured consultant skill acquisition (Brown, Kratochwill, & Bergan, 1982; Duley, Cancelli, Kratochwill, Bergan, & Meredith, 1983). While consultant skill level is an important component necessary to the success of the consultation process, additional variables, such as child outcomes and consumer satisfaction, have been explored more recently (Barnett et al., 1999; Bonner & Barnett, 2004; Kratochwill, Elliott, & Busse, 1995; Sheridan, 1992).

Table 9 Outcome Measures Across Cases

Case	Procedural Adherence	Measure	Structured Criteria	Effect Size	Percentage of Nonoverlapping Data	Goal Attainment Scaling	Social Validity
1	100%	Math	Not Significant	.88	0%	+1	5.8
		Academic Engagement	Significant	1.29	38%	+2	
		Off-task	Significant	.78	13%	+2	
2	100%	Class Academic Engagement	Significant	3.0	100%	+1	5.6
		Academic Engagement	N/A	1.15	75%	+1	
3	100%	Letters	Significant	2.5	84%	+1	4.6
		Words	Not Significant	1.4	50%	0	
		Academic Engagement	Significant	2.8	100%	+2	
4	100%	Academic Engagement	Significant	.99	0%	0	3.6
		Total Words	N/A	2.5	0%	0	
5	100%	Elopement Intensity	Significant	.88	N/A	+1	6

Evaluating process and treatment outcomes across cases has served as the foundation for outcome-based accountability models for training in applied settings (Barnett et al., 1999; Bonner & Barnett, 2004; Kratochwill, Elliot, & Busse, 1995). While previous research has delineated methods for

evaluating outcomes, it has become essential for practitioners to extend these methods to practice to demonstrate the value of services provided (Bonner & Barnett, 2004). Various methods from prior research, including measures of procedural fidelity to critical elements and steps of the consultation model, treatment outcomes and consumer acceptability, are employed in conjunction with the TVC model. The following measures not only provide means for evaluating process and outcome variables of the consultation model employed, but they also provide effective strategies for summarizing data across cases. Summarized data included here have been used to evaluate training in consultation; however, these methods also provide useful metrics for consultant self-evaluation of services across cases.

Measures

      Procedural integrity. Adherence to TVC model procedures is evaluated using a procedural checklist (see Table 6), which outlines the key steps within each phase of the process. The critical components of each phase—problem validation, treatment validation, and evaluation/decision making—are monitored throughout service delivery for each case. The purpose of the checklist is twofold. First, the procedural checklist outlines and guides the consultation process, providing consultants with a roadmap of services while simultaneously facilitating explanation of procedures to consumers. Additionally, the checklist allows for examination and evaluation of consultant fidelity to the primary objectives of the consultation model (Bonner & Barnett, 2004). The checklist is completed by each consultant initially and reviewed by the supervisor who examines permanent products and meets regularly with the trainee to supervise and provide guidance. Procedural integrity is measured as the percentage of objectives successfully fulfilled across all consultation steps for each case. To accomplish these goals, a checklist composed of the critical features of service delivery should be created and monitored. The percentage of objectives met can be calculated for each case, and repeatedly omitted objectives can provide information regarding potential needs for service revision or professional development.

      Structured criteria for visual inspection. A-B accountability designs, in which the dependent variable is measured continuously across baseline and treatment phases, serve as the foundation for evaluation of service delivery (Barnett, Daly, Jones, & Lentz, 2004). Visual analysis of graphed data (based on changes in level, trend, and variability, both within and across phases) is the cornerstone for the interpretation of behavior analytic treatment data (Parsonson & Baer, 1986; 1992). Therefore, consultants and stakeholders will continue to interpret outcomes on a case-by-case basis by examining graphs with client data. However, when one wants to arrive at conclusions about a treatment model, summarizing effects across cases that may differ in terms of target behaviors, treatments, or other variables is more difficult based on an accumulation of graphs.

      Fisher, Kelley, and Lomas (2003) developed visual aid techniques to improve the reliability and validity of visual inspection of A-B graphs and to train visual analysis skills. Their method, which results in a binary outcome of "significant" or "not significant" treatment effect for each graph, can also be used to summarize success rate as a percentage of significant effects out of all cases. The visual aids are created based on two criteria: (1) A pre-determined number of treatment data points (calculated based on the binomial distribution) must exceed the baseline mean line in the desired direction, and (2) the same number of treatment data points must exceed the baseline trend line in the desired direction. Thus, two visual aids, the baseline mean and trend lines are superimposed across the treatment phase to facilitate detection of treatment effects. Use of these structured criteria has been found to keep Type I and Type II errors at reasonable levels and also increase reliability of decisions across visual inspectors (Fisher, Kelley, & Lomas, 2003).

      Baseline and treatment data are entered into pre-formulated Microsoft Excel® spreadsheets to aid visual inspection. The number of data points needed above or below the two lines for an acquisition or

reduction program, respectively, are displayed within the spreadsheet, and the program reports the number of treatment data points actually exceeding both lines, facilitating the decision of whether the effect is "significant" or not. The program also provides a graph of the data on which lines are superimposed from baseline to the treatment phase to make visual inspection of effects easier. The pre-formulated spreadsheets are available for use by consultants or other service providers to determine significance of treatments both within and across cases (available for download at http://www.unmc.edu/dept/mmi/index.cfm?L2_ID=82&L1_ID=29&L3_ID=89&CONREF=97).

Effect size and percentage of nonoverlapping data. Effect size reflects the degree of change in measured behavior, and is expressed in standard deviation units. For example, an effect of 1.0 indicates that the average treatment data point was 1 standard deviation above baseline. Effect sizes are calculated using the "no assumptions" method outlined by Busk & Serlin (1992), in which the mean difference between treatment and baseline (treatment mean – baseline mean) is divided by the baseline standard deviation.

Percentage of nonoverlapping data (PND), a meta-analytic technique, provides a summary statistic for single-case design outcomes by analyzing overlapping data points. PND is calculated for TVC cases by determining the proportion of treatment data points exceeding the most extreme baseline data point in the desired direction. Although effect size and PND are relatively controversial procedures for summarizing individual cases, they provide a meaningful and objective basis for aggregating individual cases or making comparisons between cases (Parker et al., 2005; Scruggs & Mastropieri, 1998).

Goal attainment scaling. Goal attainment scaling (GAS; Kiresuk, Smith, & Cardillo, 1994) provides an indicator of direction and degree of behavior change in reference to a pre-determined goal and is a common measure in consultation research (Erchul & Sheridan, 2008). GAS is conducted by having impartial evaluators rate case outcomes on a scale from -2 (behavior significantly worse) to +2 (goal fully attained). Independent raters score each TVC case for goal attainment, considering level, trend and variability of graphic data. Consultants can utilize GAS by having consumers or impartial raters provide ratings on an anchored scale of -2 to +2. Analysis of these scores across cases provides a powerful indicator of the magnitude and direction of outcomes associated with service delivery.

Social validity. Social validity refers to an evaluation of treatment goals, procedures, and effects by important consumers (Gresham & Lopez, 1996). Treatment acceptability is assessed upon completion of each TVC case by asking the primary classroom teacher to complete five items derived from the *Intervention Rating Profile-15* (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985). The five selected items are those that loaded highest on the "acceptability" factor in factor analytic research on the scale. Teachers are asked to rate the five statements on a Likert-type scale of 1 (*strongly disagree*) to 6 (*strongly agree*), and a mean item rating (out of 6) is derived for each case. Practitioners can obtain ratings of acceptability through administration of scales, such as the IRP-15, or structured interviews with relevant consumers.

Treatment integrity. Fidelity of treatment implementation is examined for each case using a treatment checklist prepared by the consultant to reflect the treatment plan decided upon with the consultee. Treatment components and procedures are a natural outcome of the PVI, and the consultant develops a formal checklist based on information obtained through this meeting. The checklist is then provided to the consultee during training, prior to treatment implementation. The treatment checklist serves three important purposes. First, the checklist provides a script for the teachers, guiding treatment implementation. Second, the checklist serves as a means for self-evaluation of adherence to treatment procedures. Finally, the checklist functions as a permanent product of integrity to treatment steps, which can be examined after implementation and compared to corresponding treatment data. Change agents (e.g., classroom teachers) are instructed to complete the checklist daily and return the forms to the

consultant. Consultants also conduct integrity checks, in which they independently and simultaneously complete the checklist to examine interrater agreement. The percentage of steps completed is calculated across implementation sessions. Agreement is calculated by dividing the number of agreements on completion or incompletion of treatment steps by the total number of treatment steps, and multiplying that figure by 100 (Hartmann, 1977). Independent consultants can easily provide checklists, which delineate the steps of treatment, to relevant consumers. This practice can not only aid consultee training and implementation, but it can also provide valuable feedback to consultants regarding feasibility and acceptability of specific steps.

Summarized Case Outcomes

The TVC model was applied to a diverse range of cases by five first-year graduate students participating in a school psychology consultation course. The five clients ranged in age from 7 to 11 years of age. All clients had special education designations and were served in elementary schools, with grade levels ranging from 1st to 5th grade. Clients were referred by their respective teachers for an array of concerns, including aggressive behavior, off-task behavior, low academic performance, and elopement. Although cases varied considerably along a number of dimensions (e.g., target behaviors, treatments), adherence to TVC model procedures, structured criteria for visual inspection, ES, PND, GAS, and social validity ratings were obtained for all cases. Results are displayed in Table 9. Based on the aggregated data, it appears as though a high degree of procedural integrity was achieved, positive outcomes were obtained for clients, and treatments employed were moderately to highly acceptable to consultation consumers. The summarized outcome data provide preliminary support for the efficacy of the TVC model and procedures. The success rate for TVC cases is relatively high, as 78% of treatment effects were significant based on the conservative structured criteria for visual inspection (Fisher, Kelly, Lomas, 2003). Additionally, effect sizes were moderate to high across cases, with a median effect size of 1.29 (range, .88 to 3.0). GAS scores reflected predominantly positive ratings by independent raters regarding the magnitude of behavior change. Although PND was highly variable across cases, with a median score of 44% (range, 0% to 100%), closer inspection of relevant graphs reveal extreme baseline data points, which obscured PND scores. The average social validity rating was 5.12 (range, 3.6 to 6). Overall, the data presented provide evidence that novice consultants achieved positive outcomes across a variety of cases addressing a variety of concerns, suggesting that the model is feasible and effective. The wide range of behaviors and settings addressed by the TVC model across cases demonstrates the versatility of application.

Using a combination of the described methods, consultants can work to validate their own service delivery across a broad range of cases and concerns in the way outlined in this paper. Measurement of both procedural and outcome variables within and across cases equips practitioners with research supported tools, which can serve several purposes. Evaluating one's own consultative practice can (a) guide practice and professional development, (b) inform consultation consumers and relevant stakeholders regarding outcomes of services delivered, and (c) provide local, empirical support for employed treatments. Summarization of data across cases strengthens confidence in the consultation model's effectiveness and points out areas of needed revision.

Conclusion

BC was first designed by Bergan (1977) as an entirely indirect model of service delivery based on the idea that any direct contact with the client would cost valuable time and thereby reduce consultant availability (Watson et al., 1997). However, researchers in consultation have begun to recognize that the consultation process could actually benefit from the integration of direct assessment procedures beyond the verbal reports of consultees (Noell, Gresham, & Duhon, 1998; Noell & Witt, 1996). From a behavior

analytic standpoint, clients' needs are not met because BC procedures are meticulously followed, but because all relevant behaviors are directly and precisely measured to enable the careful selection of target behaviors, identification of behavioral function, data-based selection of all necessary treatment components, and confirmation of treatment implementation integrity (Watson et al., 1997).

The TVC model aids in the refinement of the BC process by integrating direct assessment using innovative behavior analytic technology. Thus, the likelihood of effectiveness is increased while the process is simultaneously streamlined. Contrary to the assumption that any direct contact with the client would cost the process efficiency, the data presented in the case examples and aggregated data across cases provide preliminary evidence that comprehensive data collection and direct validation of verbal report can not only strengthen the efficacy of the consultation process but can also be incorporated without sacrificing the efficiency of the process. It is our hope that this article will serve as a springboard for the further refinement and validation of the process and procedures of behavioral consultation.

References

Barnett, D. W., Daly III, E. J., Hampshire, E. M., Hines, N. R., Maples, K. A., Ostrom, J. K., & Van Buren, A. E. (1999). Meeting performance-based training demands: Accountability in an intervention-based practicum. *School Psychology Quarterly, 14,* 357-379.

Barnett, D. W., Daly, E. J., III, Jones, K. M., & Lentz, F. E., Jr. (2004). Empirically based special service decisions from increasing and decreasing intensity single case designs. *Journal of Special Education, 38,* 66-79.

Bergan, J. (1977). *Behavioral Consultation*. Columbus, OH: Merrill.

Bonner, M., & Barnett, D. W. (2004). Intervention-based school psychology services: Training for child-level accountability; preparing for program-level accountability. *Journal of School Psychology, 42,* 23-43.

Brown, D. K., Kratochwill, T. R., & Bergan, J. R. (1982). Teaching interview skills for problem identification: An analogue study. *Behavioral Assessment, 4,* 63-73.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: Applications in psychology and education* (pp. 187-212). Hillsdale, NJ: Erlbaum.

Cihak, D., Alberto, P. A., & Fredrick, L. D. (2007). Use of brief functional analysis and intervention evaluation in public settings. *Journal of Positive Behavior Interventions, 9,* 80-93.

Cobb, J. A., & Hops, H. (1973). Effects of academic skill training on low achieving first graders. Journal of Educational Research, 63, 74–80.

Daly, E. J., III, Andersen, M., Gortmaker, V., & Turner, A. (2006). Using Experimental Analysis to Identify Reading Fluency Interventions: Connecting the Dots. *The Behavior Analyst Today, 7,* 133-150.

Daly, E. J., III, Martens, B. K., Dool, E. J., & Hintze, J. M. (1998). Using brief functional analysis to select interventions for oral reading. *Journal of Behavioral Education, 8,* 203-218.

Daly, E. J., III, Martens, B. K., Hamler, K., R., Dool, E. J., & Eckert, T. L. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis, 32,* 83-94.

Daly, E. J., III, Murdoch, A., Lillenstein, L., Webber, L., & Lentz, F. E. (2002). An examination of methods for testing treatments: Conducting experimental analyses of the effects of instructional components on oral reading fluency. *Education and Treatment of Children, 25,* 288-316.

Daly, E. J., III, Wells, N. J., Swanger-Gagné, M. S., Carr, J. E., Kunz, G. M., & Taylor, A. M. (in press). Evaluation of the multiple-stimulus without replacement stimulus preference assessment method with students with behavior disorders using activities as stimulus events.

DeBaryshe, B. D., Patterson, G. R., & Capaldi, D. M. (1993). A performance model for academic achievement in early adolescent boys. *Developmental Psychology, 29,* 795-804.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*, 184-192.

Derevensky, J. L., Hart, S., & Farrell, M. (1983). An examination of achievement-related behaviors of high- and low-achieving inner-city pupils. Psychology in the Schools, 20, 328–336.

Duhon, G. J., Noell, G. Hl, & Witt, J. C. (2004). Identifying academic skill and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review, 33,* 429-443.

Duley, S. M., Cancelli, A. A., Kratochwill, T. R., Bergan, J. R., & Meredith, K. E. (1983). Training and generalization of motivational analysis interview assessment skills. *Behavioral Assessment, 5,* 281-293.

Elliott, S., & Sheridan, S. (1992). Consultation and teaming: Problem solving among educators, parents, and support personnel. *The Elementary School Journal, 92*, 315-338.

Erchul, W., & Martens, B. (2006) *School Consultation: Conceptual and Empirical Bases of Practice* (2nd ed.). New York: Springer.

Erchul, W. P., & Sheridan, S. M. (Eds.). (2008). *Handbook of research in school consultation.* New York, NY: Lawrence Erlbaum Associates.

Fisher, W. W., Kelley, M. E., and Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behavior Analysis, 36,* 387-406.

Fuchs, D., & Fuchs, L. (1989). Exploring effective and efficient prereferral interventions: A component analysis of behavioral consultation. *School Psychology Review, 18*, 260-279.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (2007). Using curriculum-based measurement to inform reading instruction. *Reading and Writing, 20*, 553-567.

Gamoran, A., & Nystrand, M. (1991). Background and instructional effects on achievement in eighth-grade English and social studies. Journal of Research on Adolescence, 1, 277–300.

Gettinger, M., & Seibert, J. K. (2000). Analogue assessment: Research and practice in evaluating academic skills problems. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.). New York: The Guilford Press.

Gortmaker, V. J., Daly, E. J., III, McCurdy, M., Persampieri, M. J., & Hergenrader, M. (2007). Improving reading outcomes for children with learning disabilities: Using brief experimental analysis to develop parent-tutoring interventions. *Journal of Applied Behavior Analysis, 40,* 203-221.

Gresham, F. M., & Lopez, M. F. (1996). Social validation: A unifying construct for school-based consultation research and practice. *School Psychology Quarterly, 11,* 204-227.

Harding, J., Wacker, D. P., Cooper, L. J., Millard, T., & Jensen-Kovalan, P. (1994). Brief hierarchical assessment of potential treatment components with children in an outpatient clinic. *Journal of Applied Behavior Analysis, 27,* 291-300.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis, 10,* 103-116.

Jones, K. M., & Wickstrom, K. F. (2002). Done in sixty seconds: Further analysis of the brief assessment model for academic problems. *School Psychology Review, 31,* 554-568.

Kiresuk, T. J., Smith, A., & Cardillo, J. E. (Eds.) (1994). *Goal attainment scaling: Applications, theory, and measurement.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Kratochwill, T., & Bergan, T. (1990). *Behavioral Consultation in Applied Settings: An Individual Guide.* New York: Plenum.

Kratochwill, T. R., Bergan, J. R., Sheridan, S. M., & Elliott, S. N. (1998). Assumptions of behavioral consultation: After all is said and done more has been done than said. *School Psychology Quarterly, 13,* 63-80.

Kratochwill, T. R., Elliott, S. N., & Busse, R. T. (1995). Behavior consultation: A five-year evaluation of consultant and client outcomes. *School Psychology Quarterly, 10,* 87-117.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice, 16(2),* 191-198.

Martin, G., & Pear, J. (2003). *Behavior modification: What it is and how to do it.* New Jersey: Prentice Hall.

McComas, J. J., & Mace, F. C. (2000). Theory and practice in conducting functional analysis. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Theory, research, and clinical foundations* (2nd ed.). New York: The Guilford Press.

Noell, G. H., Gresham, F. M., & Duhon, G. (1998). Fundamental agreements and epistemological differences in differentiating what was said from what was done in behavioral consultation. *School Psychology Quarterly, 13,* 81-88.

Noell, G., & Witt, J. (1996). A critical re-evaluation of five fundamental assumptions underlying behavioral consultation. *School Psychology Quarterly, 11*, 189-203.

Noell, G., & Witt, J. Toward a behavior analytic approach to consultation. In S. Watson & F. Gresham (ed.), *Handbook of Child Behavior Therapy* (pp. 41-57). New York: Plenum.

Northup, George, Jones, Broussard, & Vollmer, 1996)

Northup, Jones, Broussard, & Jones 1995).

Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., et al. (1991). Brief functional analysis of aggressive and alternative behavior in an outpatient setting. *Journal of Applied Behavior Analysis, 24*, 509–522.

Parker, R. I., Brossart, D. F., Vannest, K. J., Long, J. R., Garcia de Alba, R., Baugh, F. G., & Sullivan, J. R. (2005). Effect sizes in single case research: how large is large? *School Psychology Review, 34,* 116-132.

Parsonson, B. S., & Baer, D. M. (1986). The graphic analysis of data. In A. Poling and R. W. Fuqua (Eds.), *Research methods in applied behavior analysis: Issues and advances* (pp. 157-186). New York, NY: Plenum Press.

Parsonson, B. S., & Baer, D. M. (1992).The visual analysis of data, and current research into the stimuli controlling it. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis* (pp. 15-40). Hillsdale, NJ: Lawrence Erlbaum Associates.

Paramore, N. W. & Higbee, T. S. (2005). An evaluation of a brief multiple-stimulus preference assessment with adolescents with emotional-behavioral disorders in an educational setting. *Journal of Applied Behavior Analysis, 38*, 399-403.

Reid, R., Gonzalez, J. E., Nordness, P. D., Trout, A., & Epstein, M. H. (2004). A meta-analysis of the academic status of students with emotional/behavioral disturbance. *The Journal of Special Education, 38,* 130-143.

Scruggs, T. E. & Mastropieri, M. A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification, 22,* 221-242.

Shapiro, E. S., Angello, L. M., & Eckert, T. L. (2004). Has curriculum-based assessment become a staple of school psychology practice? An update and extension of knowledge, use, and attitudes from 1990 to 2000. *School Psychology Review, 33,* 249-257.

Sheridan, S. (1992). Consultant and client outcomes of competency-based behavioral consultation training. *School Psychology Quarterly, 7,* 245-270.

Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children.* New York: Guilford Press.

Shinn, M. R. (2004). *Administration and scoring of mathematics computation curriculum-based measurement (M-CBM) and math fact probes for use with AIMSweb.* Eden Prarie, MN: Edformation Inc.

Videen, J., Deno, S. L., & Marston, D. (1982). *Correct word sequences: A valid indicator of proficiency in written expression* (Research Report No. 84). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

Wallace, M. D., & Knights, D. J. (2003). An evaluation of a brief functional analysis format within a vocational setting. *Journal of Applied Behavior Analysis, 36*, 125-128.

Watson, S., Sterling, H., & McDade, A. (1997). Demythifying behavioral consultation. *School Psychology Review, 26*, 467-475.

Wilder, D. A., Chen, L., Atwell, J., Pritchard, J., & Weinstein, P. (2006). Brief functional analysis and treatment of tantrums associated with transitions in preschool children. *Journal of Applied Behavior Analysis, 39,* 103-107.

Witt, J. C. (1997). Talk is not cheap. *School Psychology Quarterly, 12,* 281-292.

Witt, J., & Elliott, S. (1983). Assessment in behavioral consultation: The initial interview. *School Psychology Review, 12*, 42-49.

Author Contact Information:

Melissa N. Andersen
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Phone: (402) 472-2207
Email: melissandersen@gmail.com

Kristi L. Hofstadter,
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Email: khofstadter@yahoo.com

Sara Kupzyk,
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Email: sarak@bigred.unl.edu

Edward J. Daly III

33 Teachers College
Lincoln, NE 68588-0345
Phone: (402) 472-5923
Fax: (402) 472-8319
Email: edaly2@unl.edu

Amanda A. Bleck,
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Email: blecky85@hotmail.com

Alyssa L. Collaro,
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Email: AlyssaLCollaro@hotmail.com

Kristin E. Jones,
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Email: kjones39@bigred.unl.edu

Lynne F. Clure,
114 Teachers College Hall
University of Nebraska - Lincoln
Lincoln, NE 68588-0345
Email: lclure1@bigred.unl.edu

Carrie A. Blevins
114 Teachers College Hall
University of Nebraska – Lincoln
Lincoln, NE 68588-0345
Email: cblevin1@bigred.unl.edu