



# Using Direct and Indirect Functional Assessments to Guide the Selection of Individualized Academic Interventions

Tyler-Curtis C. Elliott<sup>1</sup> · Andrea M. Zawoyski<sup>2</sup> · Kevin M. Ayres<sup>1</sup>

Accepted: 5 February 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

When teachers work with students exhibiting academic failure, they may look to factors outside of instruction such as a student's home life or perceived disability as explanations. Placing the locus of control outside of the instructional context becomes a convenient way to escape culpability for unsatisfactory outcomes. A more functional approach to addressing academic deficits allows educators to determine environmental factors responsible for the lack of progress and then create interventions designed to address these functions of academic failure. Although experimental analyses serve as the gold standard for evaluating functional relations between behavior and environment, educators may not always have the ability to systematically test all behavior-environment relations. Indirect assessments provide one means to develop hypotheses about environment-behavior relations that can then be validated with experimental analyses. In this study, researchers developed an indirect tool (Academic Diagnostic Checklist - Beta; ADC-B) based on the function of academic performance deficits (Daly et al. in *School Psychology Review* 26:554, 1997) and validated the use of the ADC-B by comparing interventions that were suggested (indicated) and those non-suggested (contraindicated) by the ADC-B. Researchers used the ADC-B with four participants and found that for three of the four participants, the suggested intervention was the most efficacious at improving accuracy with the target skills. One limitation is that we did not evaluate the full technical adequacy of the ADC-B, which should be a focus of future research.

**Keywords** Academic intervention · Functional relation · Indirect assessment · Brief experimental analysis · Academic deficits

---

✉ Tyler-Curtis C. Elliott  
tce62651@uga.edu

Extended author information available on the last page of the article

## Introduction

The behavior analytic approach to learning is rooted in the philosophy that all students can learn and that the teacher is responsible for adapting the environment to promote said learning (Fredrick et al., 2000). In fact, Skinner (1968, p. 242) suggested that the largest inefficiency in education was the lack of differentiation for students. As demonstrated within the Multi-Tiered Systems of Support (MTSS) framework, environmental adaptations, such as supplementary academic interventions, can lead to significant improvements in academic performance (Greenwood et al., 2011). Despite the wide range of available individualized academic interventions, not every intervention will work for all students or all academic deficits (Eckert et al., 2002; Mellott & Ardoin, 2019; Parker et al., 2012). Thus, individual students' needs may vary and not all interventions will result in academic successes (Daly & Martens, 1997; Maggin et al., 2016). Therefore, researchers should identify for whom and under what conditions interventions are effective (Wolery, 2013).

Given that the reasons for the student's academic deficits are likely idiosyncratic, it is important to select interventions based on diagnostic information rather than based on teacher intuition in order to better match the environmental problem to the intervention (Maggin et al., 2016; Wagner et al., 2017). One diagnostic approach to academic failure that may be successful is a functional approach. In the field of developmental disabilities, researchers have repeatedly demonstrated that determining the function of a problem behavior increases the likelihood of finding an effective and precise intervention for that behavior (Campbell, 2003; Hanley et al., 2003). Even for reducing the problem behavior of students with learning disabilities, where the evidence base for function-based treatments is still emerging (McKenna et al., 2015), function-based interventions are still the ideal approach (Chunta & DePaul, 2022) and required by law for students with a behavior support plan (IDEA, 2004).

Daly et al. (1997) successfully demonstrated that a functional approach to identifying effective interventions could improve academic performance (rather than arbitrary selected or default interventions). In doing so, they proposed five environmental explanations (functions) of learning failure.

### Environmental Causes of Academic Failure

- (1) *The student doesn't want to perform the skill.* This function is likely to be the cause of a performance deficit if there is no reinforcement for responding. Even if the learner received instruction and performed the skill before, without reinforcement, the behavior contacts extinction.
- (2) *The learner has not spent enough time doing it.* Students learn to skillfully engage in a behavior by doing it and contacting reinforcement. The only way to shape a behavior is to evoke responding and allow the responses to be molded by the surrounding contingencies. Thus, the more a student performs a response the more likely teachers can shape it into the target behavior. In schools, teachers capitalize on this by increasing opportunities to respond and active student responding (Haydon et al., 2012; MacSuga-Gage & Simonsen, 2015; Sutherland

- & Webby, 2001). If teachers do not maximize these opportunities to respond (or use them incorrectly), it can become a hindrance to effective instruction.
- (3) *They have not had enough help to do it.* While students may learn by contacting naturally occurring contingencies, this is not the most efficient way to learn. Teachers can accelerate learning by prompting correct responses (Mueller et al., 2007) and providing consequences such as performance feedback. Thus, if the student is not learning and does not have any assistance, this may be the reason for the academic deficit. Also under the umbrella of “help to do it,” there may also be a mismatch in the instructional hierarchy (Daly et al., 1996). According to this model, skills develop in a sequence of stages: acquisition, fluency, generalization, and then adaptation. If assistance is provided, but focuses on the incorrect stage of the instructional hierarchy, this assistance is much less likely to be effective.
  - (4) *They have not had to do it that way before.* Sometimes students are taught to complete a skill in a particular manner, but the skill is assessed in a different format. For example, a student may practice spelling with a word bank yet the teacher assesses the spelling skill without one. Thus, this scenario includes instructional materials that allow the student to obtain the answer without actually utilizing the target skill (Vargas, 1984). If the instructional materials are not designed adequately (e.g., poor alignment with natural stimuli or testing stimuli) the learner may not have sufficient opportunities to respond in an appropriate manner to the target stimulus. Then the learner may struggle to respond correctly during the times in which the response is necessary (during testing or when needed in the natural environment).
  - (5) *The skill is too hard.* Academic skills often build upon one another. In school, students first learn component skills that are eventually synthesized into composite skills (Johnson & Street, 2013). For learners who do not master the component or prerequisite skills fluently, they may struggle to learn higher order, more complex tasks. In this case, the learner may not have the behavioral repertoire necessary to learn the target skill.

It is quite possible, if not probable, that a student’s academic deficit may be a product of several combined functions. Overall, these functions of academic performance deficits provide a model for teachers to identify interventions that are likely to be more effective for remediating individual students’ academic problems (Daly et al., 1997). For a full example of the applications of the functions of academic deficits to reading instruction, see (Gibb & Wilder, 2002).

### **Academic Experimental Analyses**

The data-based decision making model, known as the academic experimental analysis (AEA), provides one way to test these possible functions of academic performance deficits and therefore match a function-based instructional strategy (Baranek et al., 2011). Within academic interventions, the AEA allows researchers and practitioners to compare several environmental conditions in a short

amount of time (Eckert et al., 2000). Researchers can then identify effective interventions and rule out ineffective ones as well (Daly et al., 1997). They can determine a functional relation using single-case design methodology (Eckert et al., 2000) such as a multi-element design (Wolery et al., 2018), withdrawal design (Gast et al., 2018), or modified versions such as brief experimental analysis that contain mini-withdrawal designs (Leford & Gast, 2018). Because academic skills are difficult to “reverse” following the removal of an intervention (Wolery et al., 2018), most AEA focus on the comparison of two or more instructional strategies applied to different sets of materials to determine which intervention is more effective for an individual student (Daly et al., 1997; Barnek et al., 2011).

Wagner et al. (2006) combined the academic function methodology (Daly et al., 1997) with experimental analyses by creating a specific reading test condition for each of the five aforementioned hypotheses for academic deficits. Although the researchers had to develop several iterations of the experimental analysis to find the optimal reading intervention, they eventually found an effective intervention for each participant. This demonstration of academic functional analysis methodology is powerful, yet has limitations. This study highlighted the ability for researchers to use specific academic interventions corresponding to specific environmental problems to test each of Daly et al. (1997) hypothesized functions of academic deficits. Not only did they demonstrate a way to systematically test Daly’s functions, but the results suggest that this functional approach to academic deficits can result in improved academic performance.

Researchers and practitioners in behavioral sciences consider experimental analysis methodology the gold standard because it experimentally demonstrates the environmental variables’ effect on behavior (Oliver et al., 2015). In academic interventions, researchers have suggested that brief AEA takes about the same amount of time (or less) as a standardized and norm-referenced test, yet has the added benefit of providing information about intervention selection (Baranek et al., 2011; Cates et al., 2006). Research suggests that information yielded by AEAs is unique and much more effective at identifying beneficial interventions than traditional teacher-identification methods (Wagner et al., 2017). Concerning problem-solving, teachers can use AEA to determine which function is most likely related to the academic deficit (Wagner et al., 2006).

Previous researchers have used AEA to address various academic topics, including early reading skills (Wagner et al., 2017), reading comprehension (Cates et al., 2006), sight words (Baranek et al., 2011), and oral reading fluency (Daly et al., 1997). When using AEAs, researchers often compare a multitude of different interventions. For example, Baranek et al. (2011) used the AEA to evaluate *eight* different interventions on sight word accuracy. In addition, Cates et al. (2006) evaluated over *seven* interventions to determine their effects on reading fluency and reading comprehension. Finally, Eckert et al. (2000) assessed the impact of *eight* reading interventions on reading fluency. The authors justified this high number of interventions by suggesting that systematically evaluating every possible combination of intervention allows them to determine the best-individualized treatment for each participant (Eckert et al., 2000). However, in academic problem solving, time is a commodity (Daly et al., 1997), and thus, testing a large number of interventions may not be the most efficient method for practitioners.

## Informing Experimental Analyses—Indirect and Direct Assessments

A majority of functional analyses (FA) in the field of behavior analysis are conducted to identify the environmental factors responsible for severe problem behavior. Practitioners working in this domain of severe problem behavior have reported time as a barrier to implementing FAs (Roscoe et al., 2015); thus, schools must balance the need for comprehensive data with efficiency. Because of this need, the current FA methodology in addressing problem behavior is moving toward a more streamlined approach. Instead of testing multiple conditions that may, or may not, be related to the problem behavior (Iwata et al., 1994), researchers and practitioners have long used indirect and direct assessments to guide the selection of individualized conditions for the FA (Iwata et al., 2013; Northup et al., 1991; Paclawskyj et al., 2000). For example, if following an indirect and direct assessment of the behavior it is determined that those in the environment never provide access to a tangible item contingent on problem behavior, then it is not necessary to conduct a tangible condition. Doing so would take more time and may also lead to a false positive outcome (Rooker et al., 2011). Overall, the use of indirect and direct assessment together before conducting a FA allows the creation of an FA that is flexible and individualized (Hanley, 2012) as well as helps to truncate assessment time on the contingencies that are most likely to be influencing problem behavior (Broussard & Northup, 1995; Derby et al., 1992).

The subfield of performance management (Daniels & Bailey, 2014) has also reported on the benefits of using indirect and direct tools in their functional approach to performance deficits. Carr and Wilder (2016) identified many different possible interventions related to staff performance problems, including behavior skills training, adjusting staffing, changing materials, increasing supervisor presence, highlighting task outcomes, reducing task effort, and reducing task aversive qualities. However, similar to academic functional assessment, not all interventions will be effective for all staff performance problems, and it may not be time-efficient to test all possible interventions. Therefore, researchers often use tools such as the performance diagnostic checklist (PDC; Austin, 2000) or its iterations such as PDC-human services (PDC-HS; Carr & Wilder, 2016; Carr et al., 2013). With the use of these tools, researchers can isolate possible functionally related interventions for performance deficits and then validate them with experimental analyses (Wilder et al., 2020). As with problem behavior and academics, when practitioners select interventions based on the function of the performance deficit, the interventions are more likely to be effective than when the intervention is chosen arbitrarily (Gravina et al., 2021).

### Purpose

Daly et al. (1997) recommended streamlining the functional approach to academics by testing the parsimonious solutions first and then progressively increasing the intensity of the intervention. Although other research commonly utilizes this hierarchical application (Daly & Martens, 1999; Wagner et al., 2006), this method has clear limitations (Eckert et al., 2000). One other possible method may be to use a direct and indirect tool to narrow down the hypothesized functions for the academic

deficit prior to conducting an AEA. We developed a tool, similar to PDC, called the Academic Diagnostic Checklist-Beta (ADC-B) based on the functions of academic deficits described by Daly et al. (1997) and current research on academic interventions within the fields of education, educational psychology, behavior analysis, and special education. This paper contains 4 experiments that aim to evaluate the preliminary use of this tool and validate its use with experimental analysis. We evaluated the accuracy of the ADC-B in each of the 4 experiments by comparing an intervention not suggested by the ADC-B (functional unmatched) with an intervention recommended by the tool (functionally matched), thereby replicating the validation methods used for the PDC (Carr & Wilder, 2016; Carr et al., 2013).

## General Method

### Tool Development

We developed the ADC-B based on the functions of academic deficits described by Daly et al. (1997). We made adaptations to the functions described by Daly et al. based on new research on academic interventions and also based on the logistics and formatting of the questions. Since Daly et al. (1997) already included an evidence-based foundation for his model, we were heavily influenced by this body of research when developing the tool (references can be found on the ADC-B in Supplementary File 1). During the development of the tool, we consulted with two doctoral level behavior analysts with combined specialities in behavior analysis, special education, educational psychology, and individualized academic interventions. After we developed the ADC-B, we conducted a pilot use of the tool with several teachers who identified a student with an academic deficit resistant to intervention. We used the tool with these teachers to ensure the tool differentially identified functions versus non-functions (rather than the tool identifying all domains or no domains as problematic). The teachers then provided feedback to inform improvement of ADC-B questions. The teachers that made suggestions all focused on the specific wording of questions, to make the intent of each question more clear. After edits based on the trial run and teacher feedback, we fully piloted and validated the tool through this study.

Each section of the tool corresponds to one of Daly et al. (1997) functions. The motivation section corresponded with “the student doesn’t want to perform the skill.” The opportunities to respond section focused on the function “the learner has not spent enough time doing it”. We split the function “they have not had assistance” into two separate subsections for logistical purposes: assistance and instructional hierarchy. When developing the tool, we made all of the questions except the instructional hierarchy (Daly et al., 1996) section focus on “acquisition” struggles. The questions in the instructional hierarchy section look at problems with fluency, generalization, and adaptation (Haring et al., 1978). In addition, we included the “they have not had to do it that way before” section under instructional materials. Lastly, we presented the function “the skill is too hard” in the section titled unmatched difficulty.

The ADC-B contained 32 questions divided into six categories that all could be answered through direct observation, parent interviews, teacher interviews, and observations in contrived situations. Each contained a final “yes” or “no” response based on the information gathered. Each primary question contained several guiding questions to assist the implementer in obtaining information during interviews. Interviewers should not use the questions as scripts, rather they may adapt the questions in a way that the informant can understand or adapt questions to make them specific to the skill at hand. For each question, an answer on the right column (bolded) represented an increased chance that the academic deficit was because of that specific environmental issue (function). Lastly, we added example interventions with literature related to each function of academic deficit on the final page of the tool. For a complete copy of the ADC-B tool, see Supplementary File 1.

### **ADC-B Administration and Selection**

To answer the questions on the ADC-B, we used a multi-method, multi-informant approach to answer the questions (e.g., interviews with teacher, parents, student, review of records, CBM data, permanent products, peer comparisons, attempts at isolating the problem, etc.). If possible, we validated the information with direct observations. During interviews, we adapted questions and provided examples specific to the target problem and used all the information together for our final answer. When reported information was inconsistent (mother reported one thing, teacher reported another) we weighed the validity of the information in the final decision. For example, we always valued direct observation more than interview and trusted the interviewee with the most experience with teaching the client the skill. With Anna (experiment 1), an example of this situation is when mother reported that Anna had not worked on multiplication in a year and both Anna and the teacher reported that she worked on multiplication almost daily. When information was inconsistent, it was always clear who the more valid source of information was. This assertion can be demonstrated by the high procedural fidelity when a second observer also scored the ADC-B tool.

It took us 15–30 min to interview each informant. This information in combination with direct observation and permanent product data was enough for us to answer all questions. After completion, the tool suggested a variety of suggested interventions for each participant. When making selections for the comparison, we prioritized functional domains that were rated at a higher percentage and used clinical judgment to develop intervention packages that were likely to be feasible together. The tool also suggested a variety of non-suggested interventions, and thus for the comparison, we chose non-suggested interventions likely to be used or recommended by teachers in schools.

## Procedural Fidelity and Interobserver Agreement

A second observer collected data on the correct implementation of procedures using direct systematic observation. All observers were master's or doctoral level students studying behavior analysis or special education and behavior analysis. The observers were familiar with data collection and thus only required an explanation of procedures and the target behaviors to achieve high levels of interobserver agreement (IOA) and procedural fidelity.

The observer calculated the percentage of procedural fidelity by taking the number of steps the researcher completed correctly, dividing it by the total steps they should have completed, and multiplying by 100%. The first author created individualized procedural fidelity checklists for each intervention prior to conducting the first intervention session with each participant. We attempted to cover the critical features of each intervention procedure with the checklists (checklists are available from the first author upon request; procedure descriptions are available via Supplementary File 2). Table 1 shows the specific results of the procedural fidelity checks with data collectors obtaining well above the suggested minimum (Ledford et al., 2020) with near perfect procedural fidelity. We did not collect agreement data for procedural fidelity.

The second observer also simultaneously coded sessions alongside the primary data collector to determine IOA. We used the point-by-point method (Ledford et al., 2018) to evaluate the agreement between the two data collectors (problem-by problem for Anna, Chase, and Trent, letter-by-letter for Damon). For each point, we coded if data collectors agreed on the type of response (correct or error) for all problems attempted. Table 2 shows the specific results of the IOA checks with data collectors obtaining well above the suggested minimum percentage of sessions (Ledford et al., 2020) with near perfect agreement.

In order to obtain reliability data on the ADC-B, we provided a second coder with all relevant records from the administration of the ADC-B and had the second coder also complete the ADC-B. We calculated the primary ADC-B and the secondary agreement using a point-by-point formula in which we counted the number of questions with an exact agreement and divided it by the number of total questions to obtain the percentage of agreement. For Anna (experiment 1), we allowed them to listen to the interviews with Anna and Anna's teacher and obtain an inter-assessor

**Table 1** A summary of procedural fidelity data for all participants

Procedural Fidelity	Probes/Baseline	Suggested	Non-suggested	Control
Anna	42.31%; 100%	39.47%; 100%	50%; 100%	60%; 100%
Chase	44.44%; 95%	66.66%; 100%	33.33%; 100%	33.3%; 100%
Damon	40%; 100%	38.88%; 97.14%	44%; 100%	–
Trent	50%; 100%	33.33%; 100%	66.66%; 100%	–

The first number represents the percentage of sessions for which researchers collected procedural fidelity. The second number after the semicolon represents the average steps completed correctly by the researchers



**Table 2** A summary of procedural interobserver agreement data for all participants

IOA	Baseline	Suggested	Non-Suggested	Control
Anna	100%; 95.85%	39.47%; 99.21%	50%; 100%	60%; 100%
Chase	33.33%; 100%	66.66%; 100%	33.33%; 100%	33.33%; 100%
Damon	40%; 98.68%	38.88%; 98.87%	44%; 99.40%	–
Trent	50%; 100%	33.33%; 100%	66.66%; 100%	–

The first number represents the percentage of sessions for which researchers collected interobserver agreement data. The second number after the semicolon represents the average percentage of agreement by the researchers

agreement of 92.31%. For Damon (experiment 3), we had the secondary coders listen to the interviews with Damon and his mother and we obtained 100% inter-assessor agreement.

## Experiment 1

### Method

#### Participant and Setting

Anna was a 10-year-old Caucasian female who spoke English in the home. She attended the local university clinic to receive academic intervention due to difficulty with multiplication. Anna received her 5th grade education in a general education classroom located in a large school located in the suburb with 443 other students with 22% eligible for free lunch. A cognitive assessment (Wechsler Intelligence Scale for Children—Fifth Edition) conducted within the year preceding the current study suggested that Anna demonstrated average intellectual functioning. She demonstrated strong verbal abilities, within the 95th percentile. On the Kaufman Tests of Educational Achievement—Third Edition she scored within the 18th percentile for math skills, suggesting a low average range of mathematical achievement. Anna and her parents reported that she had always struggled with foundational mathematics such as using place value during adding with regrouping. Anna attended all sessions via Zoom from her home.

#### Materials and Data Collection

For Anna, we created three equal-difficulty sets of multiplication facts based on the results of multiplication probes. Set one contained multiplication facts that start with 11, 1, 3, and 6. Set two contained multiplication facts that begin with 9, 2, 10, and 7. Set three contained multiplication facts that start with 12, 4, 5, and 8. We created all worksheets using WorkSheet Genius (n.d.) and displayed them on Anna's screen

using the share screen feature on Zoom. We used the drawing feature on Zoom to write when needed.

## Design and Measurement

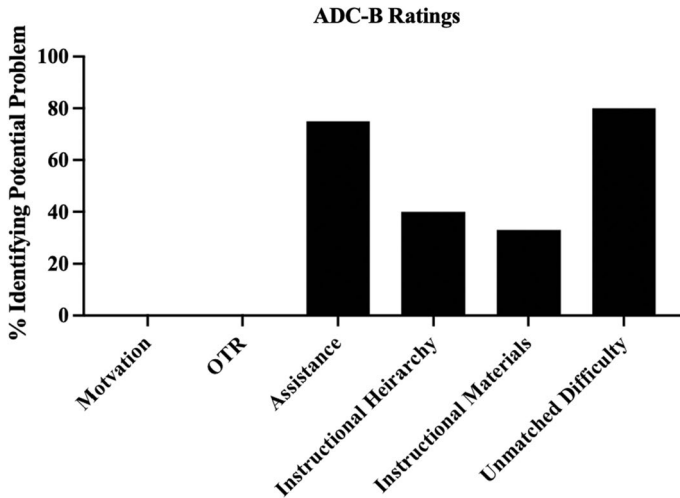
To comparatively evaluate the effects of the suggested and non-suggested interventions, we used an adapted alternating treatment design with a baseline and a control set of stimuli (Wolery et al., 2018). Specifically for Anna, we created three sets of equal-difficulty multiplication facts. After conducting baseline sessions with the three sets to ensure equal difficulty, we used a random number generator to assign the suggested intervention to set 1, the non-suggested intervention to set 2, and used set 3 as a control set. During the comparison phase, we selected the order of sequence for the sets semi-randomly (by using a random number generator while ensuring that we completed a full series before repeating the a set). Once we identified differential responding between the conditions/sets, we applied that intervention to all problem sets.

For Anna's multiplication responses, we defined correct responses as stating the answer to the problem that was the same as the answer written on the answer key. We defined errors as either skipping the problem or stating any solution that did not align with the correct response written on the paper. We collected data on correct responses and errors using a rate of responding (per minute). We conducted visual analysis by looking at changes in trend, level, and variability during condition changes (Horner & Odom, 2014). As a quantification for trend, level, and variability, we report slope, mean, and range of data when appropriate.

## Procedures

*Baseline/Probes.* We presented the relevant worksheet for the target skill (100 multiplication facts) and started by providing the specific rules and the directions. During the two minute probes, we did not help, prompt, model, or provide feedback. The only praise we provided was general praise such as "good job answering all the questions." For Anna specifically, we added to the instructions starting on session 13 that she could not skip any problems due to a high rate of skipping problems.

*ADC-B Administration and Selection.* In order to answer the questions included in the ADC-B with Anna, we conducted a record review. This record review included looking at previous psychological evaluations, a previous Individualized Education Plan (IEP), and current data from her math class. We conducted two interviews following this record review, one with Anna and one with her teacher. The results of the ADC-B (Fig. 1) suggested that lack of motivation (0%) was not a target issue because she chose to work on multiplication over other tasks and did not improve accuracy with the task when incentivized with reinforcement. In addition, lack of opportunities to respond (0%) was not the environmental events responsible for the multiplication deficit because she received daily, complete opportunities to respond. Because of this, we selected both "increasing rates of OTRs" and "ensure complete learning trial" as a treatment package for the non-suggested intervention.



**Fig. 1** Anna's academic diagnostic checklist results. *Note* The results of the Academic Diagnostic Checklist—Beta identifying which environmental factors may have been the function of Anna's multiplication performance deficit

According to the results of the ADC-B, the environmental variables that contributed to Anna's struggles to learn multiplication were that she lacked assistance (75%) informed by the fact that she rarely received prompts for completing multiplication, and there were certain strategies for multiplication that she had not been taught, such as skip counting. In addition, the ADC-B identified that the teaching was occurring within the incorrect instructional hierarchy (40%) because she could answer some multiplication facts correctly but she was not fluent with these facts. The tool also provided evidence that the instructional materials she used were not adequate (30%). Specifically, she used a multiplication chart in school and thus could get the answer to a multiplication fact without having to practice multiplication (she could find the answer with the chart). Lastly, the tool suggested that multiplication facts were too difficult for her current skill level (80%) because she had not learned various prerequisite or component skills for multiplication, such as skip counting numbers. Because of this information provided by the tool, we selected "modeling" strategies for completing multiplication and "teach component skills" as our treatment package for the suggested intervention. Table 3 provides more detailed descriptions of the suggested and non-suggested interventions for Anna and Supplementary File 2 contains operational descriptions of the procedures.

### Social Validity Measure

We developed our own social validity questionnaire (1) to determine how important the goals selected by the participants' parents were to the participants and (2) to compare the client's preference of the two compared interventions. We developed this social validity questionnaire based on widely accepted social validity domains within the field of behavior analysis (Wolf, 1978). Specifically, we looked at the social significance

**Table 3** Summary of interventions for all participants

	Treatment packages
Anna	<p>Non-suggested</p> <p>Increase OTRs: Completed 100 problems. Complete Learning Trial: Received feedback and error correction until correct response</p> <p>Suggested</p> <p>Teach prerequisite skills: Prompt and practice skip counting numbers. Model: Demonstrate using skip counting skills to obtain fact answers</p> <p>Non-suggested</p> <p>Repeated Practice: Repeated the same four problems three times each. Reinforcement: Received 3 min with phone following correct response</p> <p>Suggested</p> <p>Model: Modeled how to complete a problem. Complete Learning Trial: Practiced three problems providing feedback and error correction until correct response</p> <p>Non-suggested</p> <p>Multiple Exemplar Training: Provided various Dolch words with different spelling patterns and taught to accuracy</p> <p>Suggested</p> <p>Direct Instruction: Used lessons from spelling mastery to teach spelling and other prerequisite skills</p> <p>Non-suggested</p> <p>Problem Solving: Prompted to find critical words from the question in the passage and highlight them. *Feedback: Received information on if the answer is correct or not following a response</p> <p>Suggested</p> <p>Instructional Adaptation: Segmented to passage into smaller parts Teach prerequisite skills: Used explicit instruction to teach 5 relevant vocabulary words. *Reinforcement: Received gummy following a correct response</p>
Chase	
Damon	
Trent	

of the goals by asking about the importance of the goal, if it helped in school, and if the goal would help the client in the future. Next, we looked at the social significance of each intervention (suggested and non-suggested) by asking if the client liked each procedure, if the procedures were easy for them to understand, and if the procedure made sense in relation to the target goal. Lastly, we asked questions related to the social importance of the effects of the teaching procedures by asking if the teaching procedures worked well, if they learned the target set well, and if they could complete the target set effortlessly.

## Results

The data presented in baseline suggest that Anna demonstrated accuracy with multiplication facts (14.5–20 facts per minute) but also high rates of errors (9–10.5 per minute) prior to instruction (see Fig. 2). Rates of correct responses per minute for sets one, two, and three were 20, 18, and 14.5, respectively. Rates of errors for sets one, two, and three were 10.5, 9, and 9 per minute. Given the low variability between the different sets, the baseline data provide some face validity to the selection of equal difficulty sets of materials.

During the introduction of the intervention comparison, all three data sets resulted in patterns of differentiated responding. The control set, set three, resulted in the lowest level of correct responding with a mean of 14.5 correct responses per minute (range: 10–20). This set maintained minor variability with a slight increasing trend across time (slope:  $0.544 \times X + 8.41$ ). Set two, which received the non-suggested intervention, resulted in more correct responses than the control set but less than the suggested intervention set, with a mean of 23.4 correct responses per minute (range: 18.5–28). This set maintained minor variability with a moderate increasing trend across time (slope:  $0.549 \times X + 17.5$ ). During the comparison condition, the suggested intervention set, set 1, resulted in the highest response level with a mean of 24.2 correct responses per minute (range: 11–32). This set resulted in a large increasing trend across time (slope:  $1.27 \times X + 10.2$ ) with minimal variability between each data point (range: 11.5–32).

Upon the introduction of the suggested intervention for all sets, we saw an immediate decrease in errors for all sets. Set 1 (suggested) continued to increase in correct responding (slope:  $0.95 \times X + 9.4$ ,  $R^2 = 0.25$ ) until she met mastery criteria in nine sessions. Upon the application of the suggested intervention to all sets, Anna also began to increase her correct responding in set 2 (slope:  $0.43 \times X + 12.7$ ,  $R^2 = 0.421$ ) until she reached mastery criterion within 10 sessions. Although the errors in set 3 began to decrease immediately, we saw a very slow increase in correct responding across sessions (slope:  $0.406 \times X + 4.43$ ,  $R^2 = 0.39$ ) and it took Anna 20 sessions to meet mastery criterion with this set with this set. It is likely that it took longer to master this set because this set had been a control set and thus she had spent the previous sessions repeatedly practicing errors in this set with no feedback or error correction. The results of Anna's social validity probes are shown in Table 4. Anna rated the suggested intervention as more acceptable in both domains of intervention procedures and intervention effects.

**Table 4** Social validity scale results

	Overall	Suggested	Non-suggested
<i>Anna</i>			
Goals	2.66		
Procedures		3.66	3.33
Effects		3.00	2.66
<i>Chase</i>			
Goals	3.00		
Procedures		5.00	2.33
Effects		5.00	2.00
<i>Damon</i>			
Goals	4.66		
Procedures		4.33	4.00
Effects		4.33	4.66

Client responses on a Likert scale (1–5) range of strongly disagree to strongly agree when asked about the goals, procedures, and effects utilized in the study. Higher Likert score indicates a higher acceptability rating

## Discussion

In this experiment, we used the ADC-B with Anna, a 10-year-old 5th grade student, to determine the hypothesized function(s) of her academic performance deficit related to multiplication. Once we determined hypothesized functions with the tool, we tested the functions experimentally by comparatively applying interventions suggested and non-suggested by the ADC-B tool. The interventions suggested by the tool were the most effective at increasing correct responses and reducing rates of errors related to multiplication. In addition, Anna rated the suggested interventions the most favorably on the social validity questionnaire.

One unpredicted finding was that the intervention package selected from the interventions contraindicated by the ADC-B resulted in higher levels of correct responding than the control set (but still not to the levels achieved with the suggested intervention). One possible explanation is that any intervention focused on a target skill could result in some improvements to responding, even if not tied to the function of the academic performance deficit. A more likely explanation is that the increased OTRs (contraindicated) also functioned as repeated practice of multiplication problems (indicated). There were only a finite number of multiplication problems that we practiced; thus, at 100 problems per session, Anna encountered certain math problems repeatedly. There were problems that Anna already responded correctly and thus encountering them frequently during the practice may have helped her to build fluency with those specific problems. This interpretation would explain why the non-suggested intervention resulted in increases in correct responding (fluency), while it did not decrease errors (accuracy). Thus, only the suggested intervention increased both accuracy and fluency.

Results for Anna

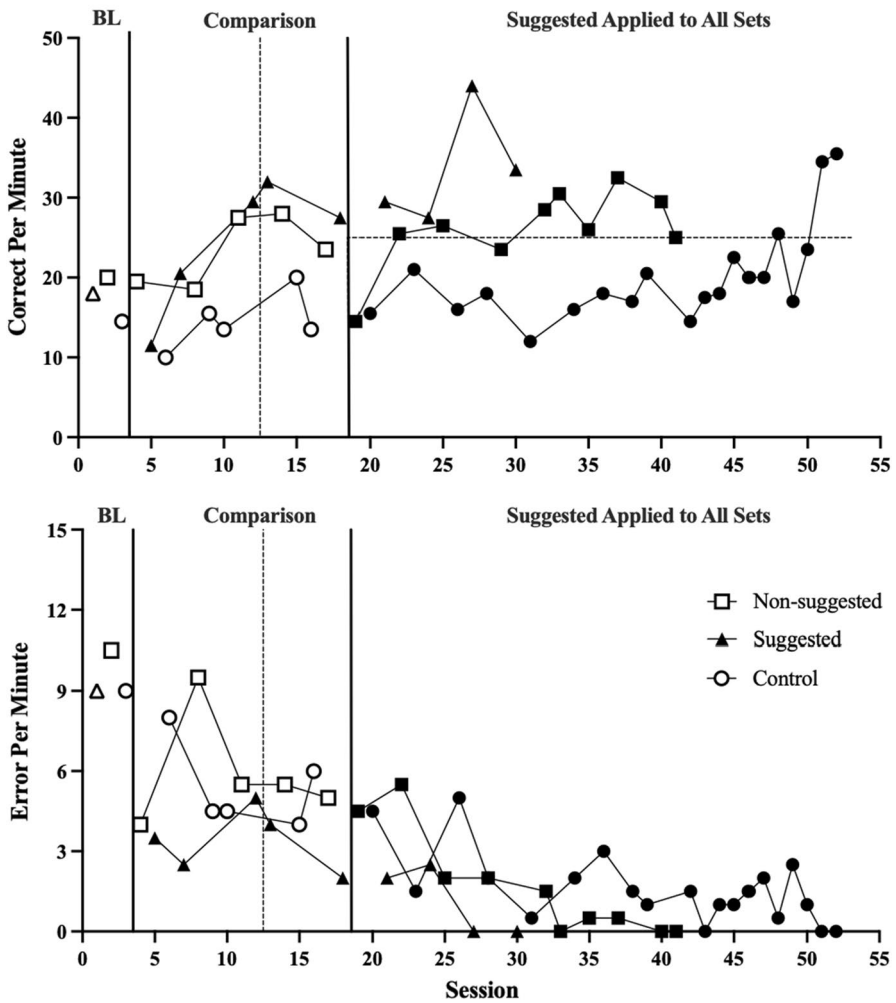


Fig. 2 Results for Anna. Note Data for the multiplication probes during baseline (BL), the condition comparing the suggested and non-suggested interventions (comparison), and the condition where the researchers applied the suggested interventions to all multiplication sets

These results together provide some face validity use of this tool in selected academic interventions. However, only one client makes generalizations about the utility of the ADC-B to other skills, participants, or interventions difficult to make. To extend the external validity, we used this tool again with a different participant, different dependent variable, and different suggested and non-suggested interventions. In experiment two, we applied this tool to 9th grade-level mathematics skills.

## Experiment 2

### Method

#### Participant and Setting

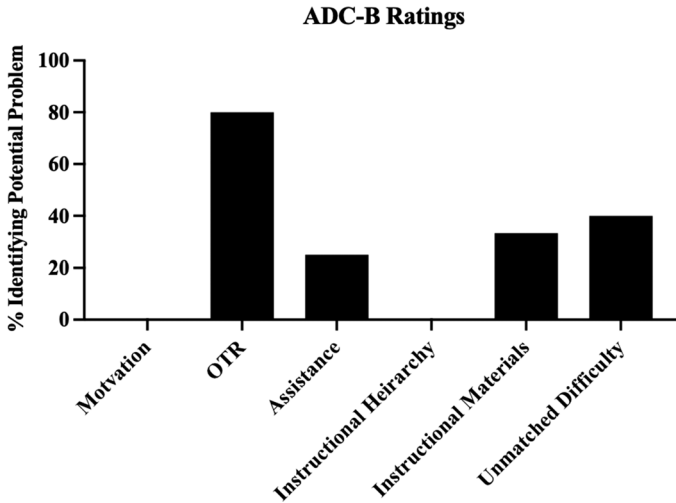
Chase was a 15-year-old Latino male who was in the 10th grade and spoke English as his primary language. He attended a local high school through a virtual program available to students during the COVID-19 pandemic. When Chase attended public school in person, he received special education services under the eligibility categories of Autism Spectrum Disorder and Speech/Language Impairment. In addition to these educational eligibilities, he was diagnosed with Attention-Deficit/Hyperactivity Disorder, combined presentation (ADHD), Oppositional Defiant Disorder (ODD), and Bipolar Disorder. His mother also reported that Chase had dysgraphia and visual-motor challenges. The online program Chase attended utilized modules containing instructional videos, discussion boards, practice worksheets, quizzes, and tests to teach academic content. During school hours, Chase had either his mother, his grandparents, or a local tutor to assist him by reading the questions for him, scribing for him, pointing out relevant discriminative stimuli, and at times, prompting him to obtain the correct answer. Chase's mother reported that Chase still struggled to master foundational skills in mathematics. Chase attended all sessions at a local university classroom.

We created our own worksheets to practice and assess the different types of transformations. Each worksheet contained two ordinate planes on the top half of the sheet, and blank space on the bottom half of the sheet for Chase to use as scratch work. We wrote the instructions for each problem below the ordinate planes. For translations, the instructions read "use the translation  $(x,y) \rightarrow (x \pm \#, y \pm \#)$ ". For dilations, the instructions read "use the dilation factor of  $\#$  and a center of  $(0,0)$ ". For reflections, the instructions read "reflect across the line". All details of the problem including shapes, numbers, number signs, lines, and planes were semi-randomly selected to ensure that the final answer reasonably fit on the ordinate plane. We randomly selected sheets from those created for both the practice problems and probes to ensure the difficulty of the practice and probe problems were not biased in any way.

#### Design and Measurement

With Chase, we also used an adapted alternating treatment design with a baseline and control set (Wolery et al., 2018). We created three sets of equal-difficulty transformation problems (dilation, translation, reflection) and applied the interventions randomly to each set. We applied the non-suggested intervention to the translation set, the suggested intervention to the dilation set, and selected reflections as the control set. During the comparison phase, we selected the order of sequence for the sets semi-randomly by using a random number generator while ensuring that we completed a full series before repeating the a set. With Chase, we did not apply the most





**Fig. 3** Chase's Academic Diagnostic Checklist results. *Note* The results of the Academic Diagnostic Checklist—Beta identifying which environmental factors may have been the function of Chase's transformation performance deficit

effective intervention to other sets. We did this because he wanted to work on other geometry tasks that more closely approximated what he was working on in school (translations was a unit/module form earlier in the semester and he had already completed it). Thus, to maintain social validity of goals, we did not want to ignore the client's request for the benefit of obtaining a more rigorous design.

For Chase's transformations, we defined correct responses as drawing a final shape on the ordinate plane that matched exactly to the shape written on the answer key. We defined errors as drawing any shape on the ordinary plane that did not align with the correct response written on the paper. We collected data on correct responses and errors using a rate of responding (per minute).

## Procedures

*Baseline/Probes.* We presented the transformation worksheets with eight problems and started by providing the specific rules and the directions. During the probes, we did not help, prompt, model, or provide feedback. The only praise we provided was general praise such as "good job answering all the questions." We stopped the session after 20 min or when he finished all eight problems, whichever came first.

*ADC-B Administration and Selection.* We first conducted a record review of a previous IEP. Secondly, we conducted several sessions of direct observation where we watched Chase complete his math schoolwork with his tutor. We also conducted two interviews following this record review, one with Chase and one with Chase's tutor. The results of the ADC-B (Fig. 3) suggested that lack of motivation (0%) was not an environmental cause because Chase reported that he found scores/grades motivating

but it didn't help him with the task and he never reported he found translations aversive. In addition, the tool suggested that the incorrect instructional hierarchy (0%) was not the environmental events responsible for the transformations deficit because Chase's teachers correctly focused on acquisition rather than fluency or generalization. Because of these results, we selected both "repeated practice" and "reinforcement/incentives" as a treatment package for the non-suggested intervention.

According to the results of the ADC-B, the environmental variables that contributed to Chase struggling to learn transformations include that he lacked opportunities to respond (80%) when his parent and tutor completed the problems for him, did not receive assistance (25%) because he had not received modeling on how to complete the skills. In addition, the instructional materials were not adequate (25%) because they did not provide examples and non-examples of different types of problems and rules for completing the problems. Lastly, the tool suggested that the task was too difficult because he often was successful when provided accommodations or modifications. With these results, we selected modeling and complete learning trials as the treatment package for the suggested intervention. Table 3 provides more detailed descriptions of the suggested and non-suggested interventions for Chase and Supplementary File 2 contains operational descriptions of the procedures.

### Social Validity Measure

We used the same social validity measure described in experiment 1 to evaluate the social significance of the goals, interventions, and effects.

### Results

The data collected during baseline validated the assumption that he had not yet developed accuracy or fluency with the transformations as he engaged in near zero levels of correct responding (range: 0–0.07 per minute) and high levels of errors (range: 0.38–0.93 per minute; see Fig. 4). Chase engaged in 0.05, 0.07, and 0 correct responses per minute for reflections, translations, and dilations respectively. Rates of errors for the sets were 0.38, 0.51, and 0.93 per minute. Although there was some moderate variability between the sets, baseline results demonstrated that Chase did not demonstrate mastery over any transformation functions.

The control set (dilations) resulted in an increase in correct responding with a mean of 0.17 correct responses per minute (0.09–0.28). Throughout the comparison conduction, the control set trend continued to improve (slope:  $0.0478 \times X - 0.193$ ) with minor variability. Although there was an increase in correct responding, Chase continued to emit high levels of error responding with a mean of 0.83 per minute (range: 0.63–1.02), no discernable trend (slope:  $0.0289 \times X + 0.613$ ,  $R^2 = 0.09$ ), and minimal variability.

On the other hand, the non-suggested intervention set (translations) resulted in lower correct responding and higher rates of errors than the baseline probe. During this condition, Chase emitted an average of zero correct responses per minute

Results for Chase

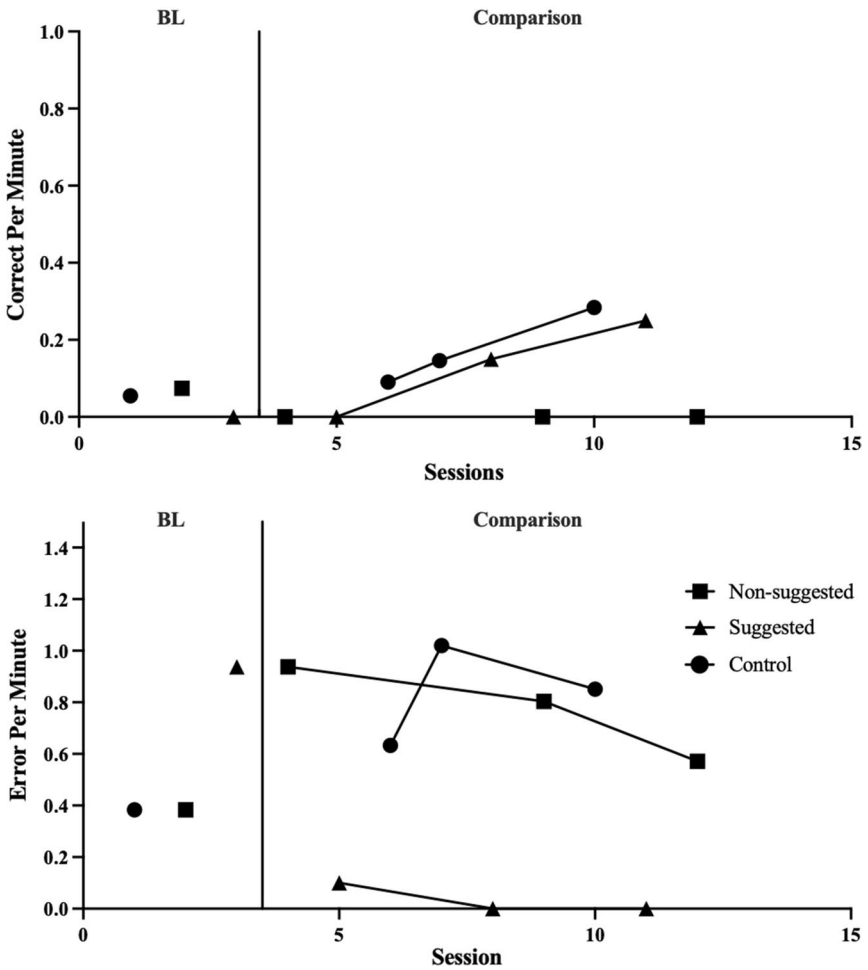


Fig. 4 Results for chase. Note Data for the transformation probes during Chase’s baseline (BL) and the comparison of the suggested and non-suggested interventions (comparison)

and averaged 0.77 error per minute (range: 0.57–0.94). Although the errors across time maintain a slightly decreased trend (slope:  $-0.0439 \times X + 1.14$ ) with low variability, the rate of errors remained much higher than probed during baseline.

Lastly, the suggested intervention set, dilations, resulted in a sudden decrease in errors and a moderate increase in correct responding with a mean of 0.13 correct responses per minute (range: 0–0.25) and a mean of 0.033 errors per minute (range: 0–0.1). Although the increase in correct responding was moderate, the data maintained an increasing trend across time (slope:  $0.0417 \times X - 0.2$ ) and low variability. Errors maintained at low levels once the data reached the floor of zero (slope:  $-0.0167 \times X + 0.167$ ) with almost no variability. The results of Chase’s social validity

probes are shown in Table 4. Chase rated the suggested intervention as more acceptable in both domains of intervention procedures and intervention effects.

## Discussion

In this experiment, we taught a high school student to accurately complete high school level geometry. We used the ADC-B to identify the function(s) of Chase's academic deficits and then used an adapted alternating treatments design to compare interventions suggested by the tool to those interventions not suggested by the tool. The interventions selected from the tool were more effective at both increasing correct responding and decreasing incorrect responding for high school level transformations. In addition, Chase completed a social validity questionnaire extremely favorably. He rated the suggested intervention procedures and effects with the highest score possible.

One limitation of this study is that we did not apply the most effective intervention to other sets. We did this because the client wanted to work on other geometry tasks that more closely approximated what he was working on in math (translations was unit/module form earlier in the semester and he had already completed it). Thus, we were not able to ignore the client's request for the benefit of obtaining a more rigorous design. We have no other reason to question the internal validity of the design; thus, the results can be interpreted as valid even without the replication to another type of transformation.

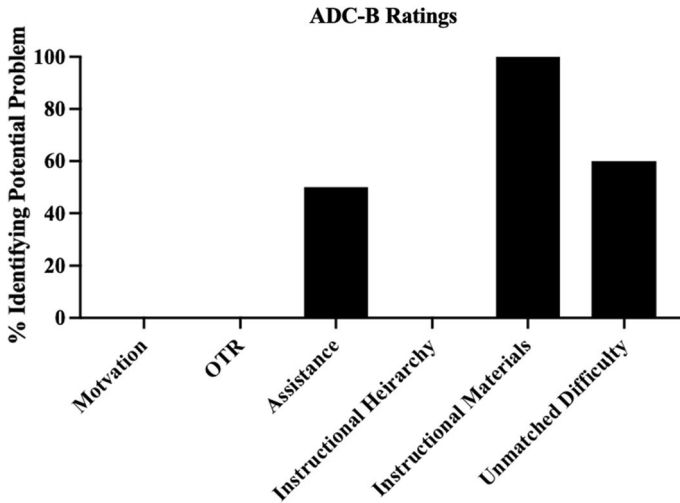
Even with the stated limitation, this experiment systematically replicated the results obtained in Experiment 1; the interventions selected based on the ADC-B were more effective than the non-suggested interventions and a control condition. More replications would further evaluate the validity of the tool. Because of this, we conducted another experiment in which we used the ADC-B to develop interventions for the spelling skills of a 9-year-old boy with Specific Learning Disability.

## Experiment 3

### Method

#### Participant and Setting

Damon was a nine-year-old African American male who was in the 4th grade and spoke English as his primary language. He attended the local virtual university clinic for assistance with spelling. Damon received special education services through the categorical label of Specific Learning Disability and received his education in a general education classroom with pullout resource support for reading and testing with a special education teacher. The school he attended was a Title I school with 391 other students with 38% eligible for free lunch. His difficulties with reading and spelling were apparent as he tended to write letters out of order and struggled to discriminate between different letters. During work time, Damon often bounced in his seat, talked to himself, and made repetitive noises. However, these



**Fig. 5** Damon's Academic Diagnostic Checklist results. *Note* The results of the Academic Diagnostic Checklist—Beta identifying which environmental factors may have been the function of Damon's spelling performance deficit

instances of motor and vocal stereotypy did not distract him from his work during his academic clinic sessions. Damon attended all sessions via Zoom.

For Damon, we used a variety of materials from the “Spelling Mastery” direct instruction curriculum (Dixon et al., 2007) during the suggested interventions and probes. During the non-suggested intervention we used a 4th grade Dolch spelling word list from the Teaching Resource Center website (n.d.).

## Design and Measurement

For Damon, we evaluated the effects using a multitreatment design (ABCBC; Gast et al., 2018) to evaluate the effects of the various interventions. Spelling is considered to be a non-reversible behavior that once learned is not easily forgotten. However, with only one participant and one target skill and one participant, a multiple baseline design was logistically impossible. Thus, we used trend and variability as our primary visual analysis tools rather than level. We hypothesized that if an intervention did not work, that we would see a flat trend and variability. We hypothesized that if the intervention did work, that we would see an increasing trend and minimal variability.

With Damon's spelling mastery probes, we counted the raw frequency of words spelled correctly (WSC) by marking each time Damon wrote the entire word exactly as it should be spelled. In addition, we calculated correct letter sequences (CLS) as a more sensitive measure of spelling improvement and used scoring as described in Hosp et al. (2007).

## Procedures

*Baseline/Probes.* With Damon we presented words one at a time and asked him to repeat the word. If he did not repeat the word correctly we used the word in a sentence then asked him to repeat the word again, until he said the correct word. Finally, we asked him to write the word. We continued to present words until we presented all 10 words in placement A. If Damon made 4 or less errors, we presented the words in placement test B in the same manner.

*ADC-B Administration and Selection.* For Damon, we did not have many records related to his history of spelling instruction. We collected the majority of the information from interviews and direct observations of spelling. For direct observations, we asked Damon to write several grade-level spelling words to look for error patterns. In addition, we probed prerequisite skills such as identifying letter-sound correspondence. We conducted three interviews following this record review: one with Damon, one with Damon's mother, and one with Damon's teacher.

The results of the ADC-B (Fig. 5) suggested that lack of motivation (0%), lack of opportunities to respond (0%), and incorrect instructional hierarchy (0%) were not the environmental events responsible for the spelling deficit. Damon's environment supported motivation through ample reinforcement for academic responding, and received daily, complete opportunities to respond, and his instruction correctly focused on spelling acquisition. Because of this, we selected both "multiple exemplar training" as a treatment package for the non-suggested intervention.

According to the results of the ADC-B, the environmental variables that contributed to Damon's struggles to learn multiplication were that he lacked assistance (50%), that the instructional materials were not adequate (100%) and that the task was too difficult for his current skill level (60%). The tool identified that he rarely received modeling for spelling. His instructional materials allowed him to respond in ways that did not use the skill of spelling and did not teach different examples of spelling rules. Lastly, he did not have the prerequisite skills needed for spelling (segmenting sounds, writing sounds, etc.). Because of this, we selected "direct instruction curriculum" as our treatment for the suggested intervention which also focuses on component/prerequisite skills. Table 3 provides more detailed descriptions of the suggested and non-suggested interventions for Damon and Supplementary File 2 contains operational descriptions of the procedures.

## Social Validity Measure

We used the same social validity measure described in experiment 1 to evaluate the social significance of the goals, interventions, and effects.

## Results

The results of WSC baseline suggest a minimal increasing trend (slope:  $0.2 \times X + 2.2$ ) (see Fig. 6). Baseline resulted with low levels of correct responding with a mean of 2.8 WSC per session. Lastly, baseline resulted in moderate levels

Results for Damon

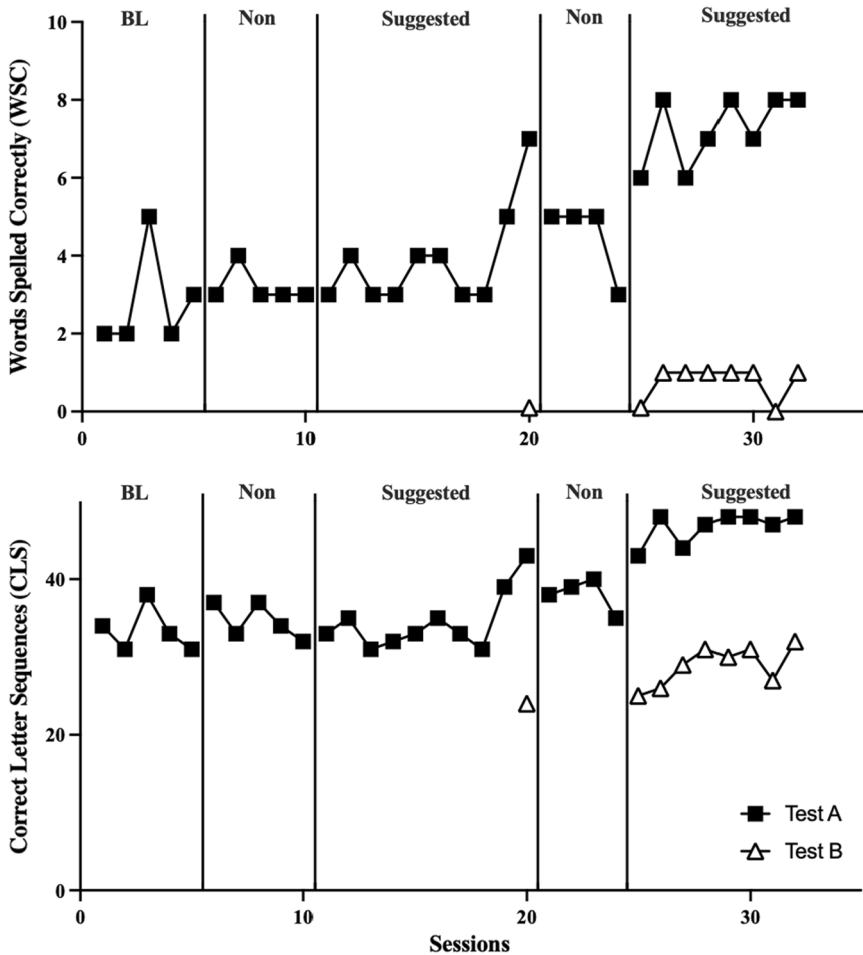


Fig. 6 Results for Damon. Note Data for the spelling probes during Damon’s baseline (BL), the non-suggested intervention (non), and the suggested interventions (suggest) evaluated through a ABCBC design

of variability (range: 2–5). The more sensitive measure, CLS, also suggests similar results during baseline except the CLS baseline data had minimal decreasing trend ( $-0.4 \times X + 34.6$ ), steady correct responding (range: 31–38) with a mean of 33.40 CLS per session.

Once we began with the non-suggested intervention, we noticed no significant change via visual analysis through WSC. We saw a small increase in the average WSC, with a mean of 3.2 per session. There was a decrease in variability (range: 3–4) and a small decreasing trend (slope:  $-0.1 \times X + 45$ ). Visual analysis of CLS suggests the same. There was no increase in level with a mean of 34.60 CLS. In addition, the last three data points resulted in a decrease in CLS overtime (slope:

$-0.9 \times X + 41.8$ ) with minimal variability (range: 32–37). Overall, the non-suggested intervention seemed to have no effect on spelling skills, so we introduced the suggested intervention.

Once we introduced the suggested intervention, there was a small increase in level by about one word for eight sessions. During sessions nine and ten, there was an abrupt increase in trend representing a delay in effect of the intervention. Overall, we saw an increase in WSC with a mean of 3.9. We saw stability across sessions (range: 3–7) with a large increase in trend of the last two sessions (slope:  $0.261 \times X + .139$ ). A similar pattern can be identified with CLS. With an overall mean of 34.5, we saw a steep increase in trend the last 2 sessions (slope:  $0.745 \times X + 22.9$ ). CLS also identified stability across sessions (range: 31–43). In the last session of the suggested intervention, Damon met criteria to receive the placement test B spelling words which is only available if the student makes 0–4 errors on test A. Damon correctly wrote 24 CLS on this last probe for test B.

After researchers returned to the non-suggested intervention, the WSC dropped in level with a mean of 4.5 per session. In addition, response continued to decrease overtime during this condition (slope:  $-0.6 \times X + 18$ ) with minimal variability (range: 3–5). The CLS data suggest the same pattern. We saw an immediate decrease in CLS with a mean of 38. The CLS data continued to decrease overtime (slope:  $-.8 \times X + 56$ ,  $R^2 = 0.25$ ) with minimal variability (range: 35–40). During these probes, Damon never met the criterion to receive the test B word list.

Finally, upon the second introduction of the suggested intervention, Damon's correct responding increased in level to a mean responding of 7.25 WSC with an increasing trend across sessions (slope:  $0.2 \times X - 2.8$ ) (slope:  $0.2 \times x - 2.8$ ). Results of the CLS analysis show similar results with an increase in responding to 46.625 with an increase across sessions (slope:  $0.512 \times X + 32$ ). Damon met criteria to receive the set B test for every session in the second suggested intervention phase. We saw a steady increase in responding for set B CLS across time ( $0.624 \times X + 11.1$ ) with an average of 28.875 CLS (range: 25–32). The results of Damon's social validity probes are shown in Table 4. Damon rated the suggested intervention as more acceptable in the area of procedures but rated non-suggested intervention as better in the domain of effects.

## Discussion

In this experiment, we replicated the results of the previous experiments by using the ADC-B to determine possible interventions to remediate the spelling deficits of a nine-year-old. We validated the results of the ADC-B by using a single case design to compare the effects of an intervention suggested by the ADC-B to a tool non-suggested by the ADC-B. The non-suggested intervention was ineffective at increasing spelling performance for Damon while the suggested intervention resulted in spelling performance improvements.

We expected a delay in intervention due to the nature of the direct instruction program. During the first lessons of spelling mastery, the lessons focus on segmentation, scanning, identifying spelled words, and identifying letters of the alphabet.



Although these skills are critical pivotal skills for a generalizable spelling repertoire, they are not likely to immediately result in increases in spelling immediately. Damon needed to first master the tool and component skills before he could put them together into the more complex composite skill of spelling.

We used the spelling mastery placement test to evaluate the effects of spelling mastery. Thus, one concern with this experiment is that the dependent and independent measures were too interconnected and thus posed a threat to the external validity of the experiment. Some could suggest that the improvement on the placement test of spelling mastery is due to the placement of test words being the same or similar to those practiced during the spelling mastery lessons. Although this is true, the generalization of the spelling repertoire Damon obtained during the suggested intervention (spelling mastery) can be seen through the improvement in test B performance. The test B placement test contains a completely new set of words that are more difficult and correspond to the next book in the spelling mastery sequence. That is, the instruction from spelling mastery A lessons generalized to more difficult words without any explicit training on those words. These data provide preliminary evidence that the suggested intervention provided Damon with a generative spelling repertoire, which is one of the components of direct instruction (Watkins & Slocum, 2004).

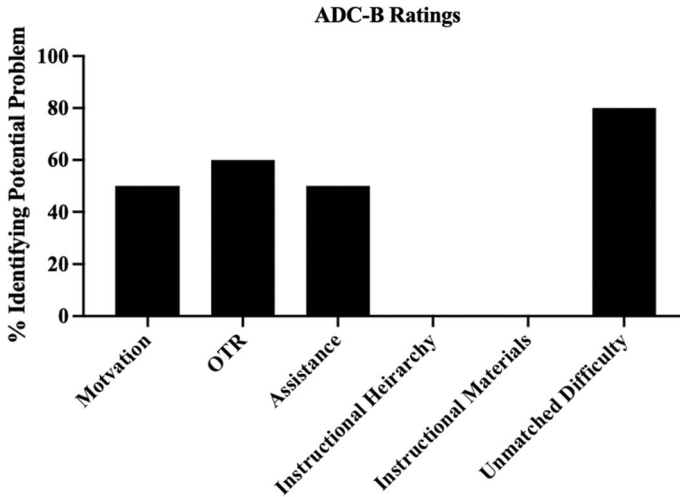
In order to continue to evaluate the generalizability of this tool, we conducted a final experiment in which we used the ADC-B to develop interventions for the reading comprehension skills of a ten-year-old boy with autism and a speech and language impairment.

## **Experiment 4**

### **Method**

#### **Participant and Setting**

Trent was a 10-year-old Caucasian male who was in 3rd grade and spoke English as his primary language. He was seen at the local university clinic to address deficits in reading comprehension. He received special education services in a self-contained/adapted curriculum classroom under the categorical label of Autism Spectrum Disorder and Speech/Language Impairment. Trent also received consultative occupational therapy in the classroom as well as 1.5 h a week of small group speech and language services. He attended a rural elementary school with 572 students, with 11% qualifying for free lunch. Trent inconsistently communicated with two to three word phrases or small sentences to meet his basic wants and needs. Results of a speech and language evaluation conducted within the last two years report that he had significant delays in language skills, with scores from formal assessments falling in the 2nd percentile for receptive language and in the 1st percentile for expressive language. In addition, Trent completed an assessment within the last two years utilizing the Verbal-Behavior Milestones and



**Fig. 7** Trent's Academic Diagnostic Checklist results. *Note* The results of the Academic Diagnostic Checklist—Beta identifying which environmental factors may have been the function of Trent's reading comprehension performance deficit

Placement Program (VB-Mapp). The results of this assessment indicate that he fell in the “Level 2” range of the VB-Mapp, indicating a developmental bracket of 18–30 months. Although Trent came to the clinic for assistance with reading comprehension, parent reports (and direct observation) suggest that reading was a personal strength for him. Specifically, he learned to decode and read fluently at a young age without any formal instruction. During the first observations in the clinic, Trent could read with a high percentage of accuracy and a high rate of fluency and scored 5/5 on both the reading and writing sections of the VB-Mapp assessment. Trent attended all sessions at a local university classroom.

For Trent, we used two 2nd grade-level proficient reading passages from the website easyCBM (n.d.) called “Feeding the Birds” (2.1) and “The Tea Party” (2.2). The passages consisted of narrative fictions approximately 500 words in length. At the end of each passage, there were 12 multiple choice (a–c) comprehension questions relating to the story. Seven of the questions intend to test “literal comprehension” while the other five test their “inferential comprehension.” In order to ensure the passages were of equal difficulty, we used the Readability Estimate Formulas offered through Intervention Central (n.d.) and demonstrated that both passages were similar in difficulty.

### Design and Measurement

For Trent, we used an alternating treatments design (Wolery et al., 2018) with baseline to evaluate the effects of the two interventions. We assigned each intervention to one passage to evaluate the cumulative effects of the intervention on those passages. Thus, each data point represents an additional exposure to the passage using the stated intervention.

Lastly, for Trent's reading comprehension, we counted a correct response as circling an answer that aligned with the answer key provided by easyCBM (n.d.) and an error as circling any other response. Using these responses, we calculated a percentage accuracy for both inferential and literal comprehension questions.

## Procedures

*Baseline/Probes.* For Trent we presented easyCBM passage and questions. We started by providing the specific rules and the directions. We used prompts to ensure he continued to read the passage but did not provide prompts or assistance for answering the comprehension problems correctly. We allowed him to take three to five minute breaks with toys after every one to three paragraphs to prevent problem behavior evoked by long task durations.

*ADC-B Administration and Selection.* We first interviewed Trent's mother, teacher, and speech therapist, and conducted a record review consisting of work materials, assessment results, and his IEP. In addition, we directly observed his reading on various tasks and levels in the clinic prior to baseline. The results of the ADC-B (Fig. 7) suggested that incorrect instructional hierarchy (0%) was not a problem because Trent's teachers focused on the acquisition phase of the instructional hierarchy, which is where he needed practice. In addition the instructional materials (0%) were not the environmental events responsible for the comprehension deficit because they provided various examples and non-examples and allowed him various ways to respond that paralleled reading comprehension in the real world. In other words, Trent's teachers used materials that were well designed to require the use of reading comprehension. Because of the results of this tool, we selected both "problem solving" as a treatment for the non-suggested intervention, which focused on the incorrect sequence of the instructional hierarchy (adaptation).

The results suggested that lack of motivation (50%) was a problem, because he avoided reading comprehension. In addition, lack of opportunities to respond (60%) and lack of assistance (50%) were two areas of concern because Trent did not complete learning trials and did not receive prompts. Lastly, the tool suggested unmatched difficulty as a major factor (80%) for his academic deficit because he did not have the prerequisite language skills. Based on these results, we chose to address the "unmatched difficulty" by teaching prerequisite (language) skills and breaking up the readings into smaller sections.

Following minimal progress with either intervention package, we added "contingent reinforcement" to the suggested intervention package. In order to balance out any effects of receiving feedback (reinforcement when correct), we also enhanced the non-suggested intervention by including "feedback" as part of the package. Table 3 provides more detailed descriptions of the suggested and non-suggested interventions for Trent and Supplementary File 2 contains operational descriptions of the procedures.

## Social Validity Measure

We used the same social validity measure described in experiment 1 to evaluate the social significance of the goals, interventions, and effects.

## Results

The data for Trent's reading comprehension performance are shown in Fig. 8. During baseline with the non-suggested passage, Trent answered 57.41% of the literal questions correctly and only 40% of the inferential questions correctly. The baseline data for the suggested passage were reversed with Trent answering 28.57% of the literal questions and 60% of the inferential questions correctly. The baseline data suggest that even though the passages were of equal difficulty, Trent's baseline performance with those passages differed. This difference was likely due to the placement of the correct answers as Trent tended to circle "A" when he guessed. Regardless, Trent answered neither of the passages at a mastery level and thus the results of both passages represented areas of improvement for Trent's reading comprehension.

Upon introduction of the suggested and non-suggested interventions, we saw a small decrease in performance in literal questions for both the non-suggested and suggested interventions. Trent's performance on the passage assigned to the non-suggested intervention dropped to an average of 47.62% (range: 42.86–57.14%). Similarly his performance on the passage corresponding to the suggested intervention dropped to an average of 19.05% (range: 14.29–28.57%). Although the non-suggested intervention data were higher in level than the suggested, both dropped in performance from baseline. The data for inferential questions show the same results. Trent's performance on the non-suggested passage dropped to an average of 46.67% (range: 20–60%) and his performance on the suggested passage dropped to an average of 46.67% (range: 40–60%). The overlap between the suggested and non-suggested inferential questions was high with no discernible difference.

Because we saw no improvement for either intervention, we enhanced both the suggested and non-suggested intervention and then continued the evaluation. There was no discernible change in correct inferential responding following the enhancement of the treatments. The data for both the suggested and non-suggested continued to be lower than baseline with high variability. Specifically, the suggested intervention resulted in an average of 40% correct responding (range: 20–60%), while the non-suggested intervention resulted in 35% correct responding (range: 20–50%). For the literal questions, we saw a 4.76% increase in performance for the non-suggested intervention to an average of 61.90% (range: 57.14–66.67%). These data should be interpreted with caution because the first non-suggested session evoked problem behavior when we told Trent "that's not the right answer" that was extreme enough to warrant ending the session early (after 8 problems instead of 12). It is unclear if the results would still be higher had Trent finished the last four problems.

Because we saw no significant increases following 5 exposures to two different teaching methods we decided to conduct an error analysis of the results. It became

Results for Trent

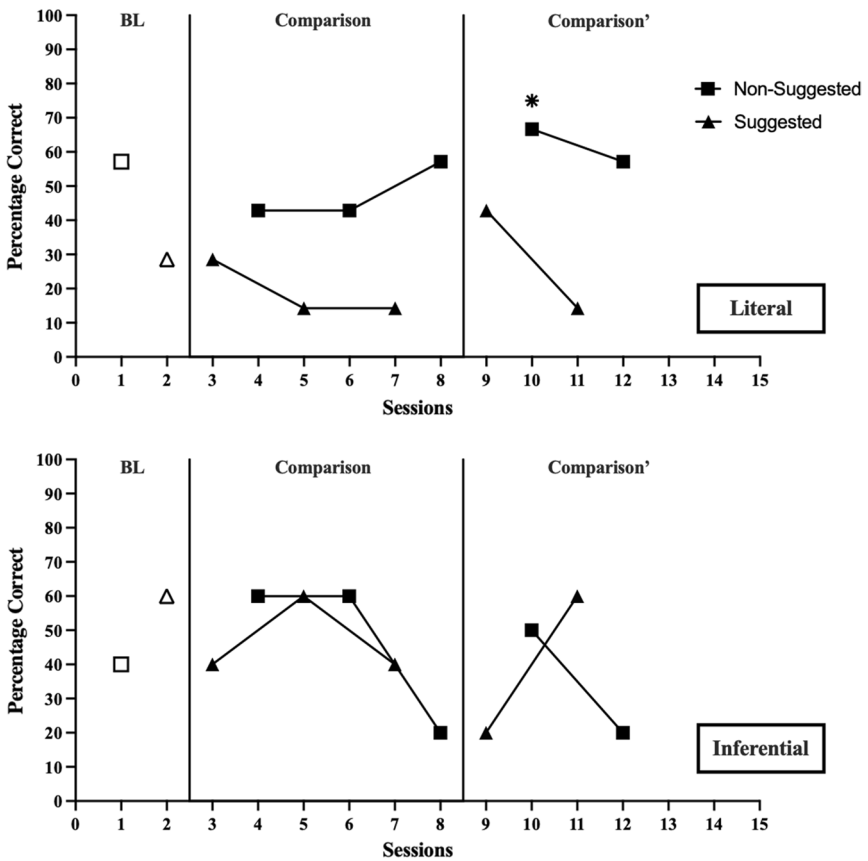


Fig. 8 Results for Trent. Note Data for the literal and inferential comprehension probes during Trent’s baseline (BL), the non-suggested intervention (non), and the suggested interventions (suggest), and an enhanced suggested intervention evaluated through a multi-element design

clear from the error analysis that Trent was selecting the letter A the majority of the time. Across baseline, the suggested intervention, and the non-suggested intervention, Trent selected the letter “A” as his answer choice 76.64% of the time. Because of the results and the error analysis suggested little progress with Trent throughout the comparison and because one of the sessions evoked problem behavior, we stopped the comparison and chose to instead spend our clinical time focusing on developing language skills.

**Discussion**

The results of Trent’s assessment suggest that neither the suggested nor the non-suggested intervention resulted in great improvements in reading comprehension.

Therefore, it is possible that the ADC-B did not provide valid intervention suggestions to address the reading comprehension problem. One could argue that this provides evidence against using this indirect assessment for individualized intervention selection. We suggest that a more likely explanation can be found in the nature of the academic deficit.

In this case, the tool identified that Trent did not have the language repertoire necessary to be able to succeed in reading comprehension. Thus, we selected an intervention intended to teach relevant vocabulary in an attempt to address the lack of prerequisite skills. Standardized assessment results determined that Trent was in the 1st–2nd percentile for language skills. In addition, he had received speech and language therapy services for the past eight years. It is unlikely that such a severe speech and language deficit, resistant to years of speech and language intervention would be susceptible to improvement during a brief adapted alternating treatment design. Thus, it is possible that the tool did identify the correct function of Trent's reading comprehension deficit, but that the deficit was not susceptible to change during a brief experimental analysis.

## General Discussion

We developed the Academic Diagnostic Checklist-Beta (ADC-B) in the hopes that this tool could serve as a direct and indirect assessment for practitioners to be able to select functionally matched individualized academic interventions. In order to evaluate the validity of the tool for identifying the correct function and correspondingly effective interventions, we validated information provided by the tool by comparing interventions suggested by the tool (indicated and functionally matched) to the interventions not suggested by the tool (contraindicated and functionally mismatched). If the suggested interventions increase performance more than the non-suggested intervention, we could assume that the tool correctly identified the function of the academic deficit. This model of tool validation through comparing interventions has been used with other assessments focused on functional relations between behavior and environment (Wilder et al., 2020).

In this study, we evaluated both the (1) process validity (ability to differentiate environmental factors responsible for the academic deficit) and (2) treatment utility (ability to suggest interventions that help to remediate the academic deficit). If these two requirements were met, this direct and indirect assessment could serve as a powerful tool for teachers and academic specialists for remediating children's academic deficits in school.

Firstly, it is clear that the tool was able to (1) differentially identify areas of environmental concern. The tool did not suggest all six domains as an area of environmental concern for any participant. In addition, the tool was able to identify at least one area of concern for every participant. Lastly, for all four participants, the tool identified one *primary* area of concern. Thus, it is clear that the tool is able to differentially suggest areas of environmental concern.

The tool also proved useful in (2) suggesting interventions to help remediate specific academic interventions. When we comparatively evaluated interventions

suggested and non-suggested by the tool, the interventions suggested by the tool were the most effective intervention for three of the four participants. For Trent, the fourth participant, neither intervention (suggested or non-suggested) was successful at remediating his reading comprehension deficit. In addition, both Anna and Chase rated the suggested intervention as more preferable in both domains of intervention procedures and intervention effects. Damon rated the suggested intervention procedures as more acceptable. Although he rated the effects of the non-suggested intervention better, the data suggest that the suggested intervention was more effective.

## Limitations

The use of experimental design as our only measure of tool validation is one limitation of this study. Not all academic skills will increase in the context of brief experimental manipulation and may require much longer durations of intervention prior to performance increasing. For example, the results of Trent's assessment suggest that neither the suggested nor the non-suggested intervention resulted in great improvements in reading comprehension. In reality, it is possible that the tool correctly identified the environmental cause of the academic deficit: that he did not have the prerequisite language skills required for reading comprehension. Standardized assessment results determined that Trent was in the 1st–2nd percentile for expressive and receptive language. In addition, he had received speech and language therapy services for the past eight years. It is unlikely that such a severe speech and language deficit, resistant to years of speech and language intervention would be susceptible to improvement during a brief adapted alternating treatment design. Thus, although the suggested intervention (nor non-suggested) did not work within the context of the experimental design that does not mean it may not have worked in a longer application of the intervention (across months or years).

Another limitation is that we did not evaluate the technical adequacy (McIntosh et al., 2008) of the tool beyond process validity, treatment utility, and IOA. Without evaluation of test-retest reliability, interrater reliability, content validity, convergent validity with brief experimental analyses, and social validity of the tool itself, it is difficult to determine the true reliability and validity of the ADC-B.

Lastly, the scope of this evaluation may be too broad to yield specific information regarding what types of students/domains for which the tool is most likely to suggest matched interventions. Because we evaluated four different domains, this study provides limited opportunities for inter-participant replication within the selected domains. For example, although we found a suggested intervention for Anna's multiplication deficit, it is unclear if for another student focusing on multiplication, the tool might recommend the same or similar intervention. Had we evaluated one dependent variable with various students (i.e., decoding CVC words), we would be better equipped to evaluate the predictive validity of the tool at determining functionally matched interventions.

## Future Directions

The process of experimentally testing each function individually (Wagner et al., 2006) can be time-consuming and cumbersome. Just as indirect assessments have helped to enhance functional analyses in other areas such as problem behavior (Hanley, 2012) and staff performance deficits (Carr & Wilder, 2016; Carr et al., 2013), we believe direct and indirect assessments such as the ADC-B can help to make the selection of individualized academic deficits more efficient.

Future researchers should continue to systematically replicate this tool evaluation. One example may be to conduct a brief experimental analysis of all of the different functions and compare the results to the suggestions provided by the tool. In addition, future researchers should evaluate if there are differences in the efficacy of the tool with different academic domains. For example, it is possible that the tool may be better equipped to identify interventions for basic math interventions than reading interventions. If researchers were to conduct the ADC-B with several students while focusing on one dependent variable, this information would help to better evaluate the predictive validity of the tool. In addition, although we conducted reliability checks with high agreement between raters, we did not conduct full psychometric analyses. Future research should continue to evaluate the psychometric properties and technical adequacy of the ADC-B.

Pending future successful replications, this tool could provide a method for which to embed more targeted intervention into MTSS prior to recommendation for special education for students who are struggling academically. For example, when a child is engaging in problematic behavior that is not remediated through Tier 1 interventions, functionally arbitrary interventions such as check-in-check-out or behavioral contracts are utilized for Tier 2 (Dunlap et al., 2010). Should these Tier 2 strategies be unsuccessful, a functional behavior assessment is implemented to create a more individualized program for the child prior to referring to special education (Dunlap et al., 2010). This tool allows a parallel for academic instruction: Should Tier 1 strategies be ineffective, schools can attempt functionless Tier 2 strategies such as small group instruction with evidence based programs (Brown-Chidsey & Steege, 2010). If those strategies are ineffective, then teachers can run a functional assessment of the academic deficit using a combination of this tool and AEA for one last attempt at remediation prior to referring a student for special education services.

Applying the ADC-B as a later step in an MTSS model is consistent with the aims of MTSS, as a tiered system of service delivery functions by concentrating the highest intensity resources for the students with the greatest need. For example, not all students can receive a FA of problem behavior due the constraints of the school resources, thus it is reserved for those who have not demonstrated progress with Tier 1 and Tier 2 strategies. The same case can be made for academic failure; not all students can receive an academic experimental analysis and thus it may be used for those who have not shown progress with Tier 1 and Tier 2 strategies.

We do not believe that the ADC-B is *the only* assessment that can be used to determine environment-learning relations responsible for academic deficits. In fact, although they do not use the word function, we know of some other assessment models that look at the “functions” of academic deficits. For example, the



Functional Assessment of Academic Behavior (Ysseldyke & Christenson, 2002) is one tool that, with major adaptations, could identify possible factors responsible for academic deficits. Although, it has not yet been validated for this purpose. Other practitioners and researchers may see the ADC-B and identify areas of adaptation. Similar to the manner in which the PDC contains several iterations (human service version, parent version), we believe that there could be certain skills (such as reading) that may benefit from its own academic diagnostic checklist. We encourage researchers to systematically replicate the ADC-B including modification that may enhance its applicability.

All students, even those whose deficits are resistant to common educational interventions, have a right to effective education (Barrett et al., 1991). This study contains a tool that recommends individualized academic interventions based on the likely environmental causes for the academic deficit. Although Daly et al. (1997) created the “functions of academic deficits,” this work has largely been untouched by practicing behavior analysts and school psychologists since its introduction. We believe this tool extends the work of Daly et al. and pushes the field one step closer to functional approaches to academics.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10864-023-09511-x>.

**Acknowledgements** We would like to thank Kadiah Quinland, Joseph Hart, Stephanie Liollo, Scott Ardoin, and Rose Morlino for assistance in various ways. Thank you to Geege Taylor for your comments on a previous version of the manuscript.

**Author Contributions** The first author developed the idea for the paper, conducted the research, and wrote the paper. The second and third author made significant contributions to the adaptation of the manuscript.

**Funding** No funds, grants, or other support was received. The authors have no relevant financial or non-financial interests to disclose.

## Declarations

**Conflict of interest** The authors have no conflict of interests to report.

**Informed consent** Researchers received approval from the University IRB and parents of all participants provided informed consent to conduct and publish the study.

## References

- Austin, J. (2000). Performance analysis and performance diagnostics. In J. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 321–349). Context Press/New Harbinger Publications.
- Barrett, B. H., Beck, R., Binder, C., Cook, D. A., Engelmann, S., Greer, R. D., Kyrklund, S. J., Johnson, K. R., Maloney, M., McCorkle, N., Vargas, J. S., & Watkins, C. L. (1991). The right to effective education. *The Behavior Analyst, 14*(1), 79. <https://doi.org/10.1007/BF03392556>

- Baranek, A., Fienup, D. M., & Pace, G. (2011). Brief experimental analysis of sight word interventions: A comparison of acquisition and maintenance of detected interventions. *Behavior Modification*, 35(1), 78–94. <https://doi.org/10.1177/0145445510391242>
- Broussard, C. D., & Northup, J. (1995). An approach to functional assessment and analysis of disruptive behavior in regular education classrooms. *School Psychology Quarterly*, 10(2), 151. <https://doi.org/10.1037/h0088301>
- Brown-Chidsey, R., & Steege, M. W. (2010). *Response to intervention: Principles and strategies for effective practice* (2nd ed.). The Guilford Press.
- Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities*, 24(2), 120–138. [https://doi.org/10.1016/s0891-4222\(03\)00014-3](https://doi.org/10.1016/s0891-4222(03)00014-3)
- Carr, J. E., Wilder, D. A., Majdalany, L., Mathisen, D., & Strain, L. A. (2013). An assessment-based solution to a human-service employee performance problem. *Behavior Analysis in Practice*, 6(1), 16–32. <https://doi.org/10.1007/BF03391789>
- Carr, J. E., & Wilder, D. A. (2016). The performance diagnostic checklist—human services: A correction. *Behavior Analysis in Practice*, 9(1), 63–63. <https://doi.org/10.1007/s40617-015-0099-3>
- Cates, G. L., Thomason, K., Havey, M., & McCormick, C. (2006). A preliminary investigation of the effects of reading fluency interventions on comprehension: Using brief experimental analysis to select reading interventions. *Journal of Applied School Psychology*, 23(1), 133–154. [https://doi.org/10.1300/J370v23n01\\_07](https://doi.org/10.1300/J370v23n01_07)
- Chunta, A. M., & DuPaul, G. J. (2022). Educational diagnostic label and teacher self-efficacy: Impact on classroom intervention choice. *School Psychology*, 37(4), 298–308. <https://doi.org/10.1037/spq0000509>
- Daly, E. J., III., Lentz, F. E., Jr., & Boyer, J. (1996). The Instructional Hierarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly*, 11(4), 369. <https://doi.org/10.1037/h0088941>
- Daly, E. J., III., Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review*, 26(4), 554. <https://doi.org/10.1080/02796015.1997.12085886>
- Daly, E. J., III., & Martens, B. K. (1999). A brief experimental analysis for identifying instructional components needed to improve oral reading fluency. *Journal of Applied Behavior Analysis*, 32(1), 83. <https://doi.org/10.1901/jaba.1999.32-83>
- Daniels, A., & Bailey, J. S. (2014). *Performance management: Changing behavior that drives organizational effectiveness* (5th ed.). Performance Management Publications.
- Derby, K. M., Wacker, D. P., Sasso, G., Steege, M., Northup, J., Cigrand, K., & Asmus, J. (1992). Brief functional assessment techniques to evaluate aberrant behavior in an outpatient setting: A summary of 79 cases. *Journal of Applied Behavior Analysis*, 25(3), 713–721. <https://doi.org/10.1901/jaba.1992.25-713>
- Dixon, R., Engelmann, S., & Bauer, M. M. (2007). *Spelling mastery*. UK: SRA/McGraw Hill.
- Dunlap, K., Goodman, S., McEvoy, C., & Paris, F. (2010). *School-wide positive behavioral interventions and supports: Implementation guide*. Lansing, MI: Michigan Department of Education.
- EasyCBM. (n.d.). *Proficient reading*. Easy CBM. <https://app.easycbm.com/training/?action=measures#B--RiInDY8Y>.
- Eckert, T. L., Ardoin, S. P., Daisey, D. M., & Scarola, M. D. (2000). Empirically evaluating the effectiveness of reading interventions: The use of brief experimental analysis and single case designs. *Psychology in the Schools*, 37(5), 463–474. [https://doi.org/10.1002/1520-6807\(200009\)37:5%3c463::AID-PITS6%3e3.0.CO;2-X](https://doi.org/10.1002/1520-6807(200009)37:5%3c463::AID-PITS6%3e3.0.CO;2-X)
- Eckert, T. L., Ardoin, S. P., Daly, E. J., III., & Martens, B. K. (2002). Improving oral reading fluency: A brief experimental analysis of combining an antecedent intervention with consequences. *Journal of Applied Behavior Analysis*, 35(3), 271–281. <https://doi.org/10.1901/jaba.2002.35-271>
- Fredrick, L. D., Deitz, S. M., Bryceland, J. A., & Hummel, J. H. (2000). *Behavior analysis, education, and effective schooling*. Context Press.
- Gast, D. L., Ledford, J. R., & Severini, K. E. (2018). Withdrawal and reversal designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (3rd ed., pp. 215–238). Routledge.

- Gibb, G. S., & Wilder, L. K. (2002). Using functional analysis to improve reading instruction for students with learning disabilities and emotional/behavioral disorders. *Preventing School Failure: Alternative Education for Children and Youth*, 46(4), 152–157. <https://doi.org/10.1080/10459880209604414>
- Gravina, N., Nastasi, J., & Austin, J. (2021). Assessment of employee performance. *Journal of Organizational Behavior Management*, 41(2), 124–149. <https://doi.org/10.1080/01608061.2020.1869136>
- Greenwood, C. R., Bradfield, T., Kaminski, R., Linas, M., Carta, J. J., & Nylander, D. (2011). The response to intervention (RTI) approach in early childhood. *Focus on Exceptional Children*, 43(9), 1–22. <https://doi.org/10.17161/foec.v43i9.6912>
- Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis*, 36(2), 147–185. <https://doi.org/10.1901/jaba.2003.36-147>
- Hanley, G. P. (2012). Functional assessment of problem behavior: Dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis in Practice*, 5(1), 54–72. <https://doi.org/10.1007/BF03391818>
- Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. (1978). *The fourth R: Research in the classroom*. Merrill.
- Haydon, T., Macsuga-Gage, A. S., Simonsen, B., & Hawkins, R. (2012). Opportunities to respond: A key component of effective instruction. *Beyond Behavior*, 22(1), 23–31. <https://doi.org/10.1177/107429561202200105>
- Horner, R. H., & Odom, S. L. (2014). Constructing single-case research designs: Logic and options. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research* (pp. 27–51). American Psychological Association.
- Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). *The ABCs of CBM: A practical guide to curriculum based measurement* (1st ed.). The Guilford Press.
- Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)
- Intervention Central. (n.d.). *Oral Reading Fluency Passage Generator*. Intervention Central. <https://www.interventioncentral.org/teacher-resources/oral-reading-fluency-passages-generator>
- Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis*, 27(2), 197–209. <https://doi.org/10.1901/jaba.1994.27-197>
- Iwata, B. A., DeLeon, I. G., & Roscoe, E. M. (2013). Reliability and validity of the functional analysis screening tool. *Journal of Applied Behavior Analysis*, 46(1), 271–284. <https://doi.org/10.1002/jaba.31>
- Johnson, K., & Street, E. M. (2013). *Response to intervention and precision teaching: Creating synergy in the classroom*. Guilford Press.
- Ledford, J. R., & Gast, D. L. (2018). Combination and other designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (3rd ed., pp. 335–364). Routledge.
- Ledford, J. R., Lane, J. D., & Gast, D. L. (2018). Dependent variables, measurement, and reliability. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (3rd ed., pp. 97–131). Routledge.
- Ledford, J. R., Chazin, K. T., Lane, J. D., Zimmerman, K. N., & Ayres, K. A. (2020). Single case analysis and review framework (SCARF). Retrieved from <http://ebip.vkcsites.org/scarfv2>
- MacSuga-Gage, A. S., & Simonsen, B. (2015). Examining the effects of teacher-directed opportunities to respond on student outcomes: A systematic review of the literature. *Education and Treatment of Children*. <https://doi.org/10.1353/etc.2015.0009>
- Maggin, D. M., Wehby, J. H., & Gilmour, A. F. (2016). Intensive academic interventions for students with emotional and behavioral disorders: An experimental framework. *Journal of Emotional and Behavioral Disorders*, 24(3), 138–147. <https://doi.org/10.1177/10634266166491>
- McIntosh, K., Borgmeier, C., Anderson, C. M., Horner, R. H., Rodriguez, B. J., & Tobin, T. J. (2008). Technical adequacy of the Functional assessment checklist: Teachers and Staff (FACTS) FBA interview measure. *Journal of Positive Behavior Interventions*, 10(1), 33–45. <https://doi.org/10.1177/1098300707311619>
- McKenna, J. W., Flower, A., Kyung Kim, M., Ciullo, S., & Haring, C. (2015). A systematic review of function-based interventions for students with learning disabilities. *Learning Disabilities Research & Practice*, 30(1), 15–28. <https://doi.org/10.1111/ldrp.12049>
- Mellott, J. A., & Ardoin, S. P. (2019). Using brief experimental analysis to identify the right math intervention at the right time. *Journal of Behavioral Education*, 28(4), 435–455. <https://doi.org/10.1007/s10864-019-09324-x>

- Mueller, M. M., Palkovic, C. M., & Maynard, C. S. (2007). Errorless learning: Review and practical application for teaching children with pervasive developmental disorders. *Psychology in the Schools, 44*(7), 691–700. <https://doi.org/10.1002/pits.20258>
- Northup, J., Wacker, D., Sasso, G., Steege, M., Cigrand, K., Cook, J., & DeRaad, A. (1991). A brief functional analysis of aggressive and alternative behavior in an outclinic setting. *Journal of Applied Behavior Analysis, 24*(3), 509–522. <https://doi.org/10.1901/jaba.1991.24-509>
- Oliver, A. C., Pratt, L. A., & Normand, M. P. (2015). A survey of functional behavior assessment methods used by behavior analysts in practice. *Journal of Applied Behavior Analysis, 48*(4), 817–829. <https://doi.org/10.1002/jaba.256>
- Paclawskyj, T. R., Matson, J. L., Rush, K. S., Smalls, Y., & Vollmer, T. R. (2000). Questions about behavioral function (QABF): A behavioral checklist for functional assessment of aberrant behavior. *Research in Developmental Disabilities, 21*(3), 223–229. [https://doi.org/10.1016/s0891-4222\(00\)00036-6](https://doi.org/10.1016/s0891-4222(00)00036-6)
- Parker, D. C., Dickey, B. N., Burns, M. K., & McMaster, K. L. (2012). An application of brief experimental analysis with early writing. *Journal of Behavioral Education, 21*(4), 329–349. <https://doi.org/10.1007/s10864-012-9151-3>
- Rooker, G. W., Iwata, B. A., Harper, J. M., Fahmie, T. A., & Camp, E. M. (2011). False-positive tangible outcomes of functional analyses. *Journal of Applied Behavior Analysis, 44*(4), 737–745. <https://doi.org/10.1901/jaba.2011.44-737>
- Roscoe, E. M., Phillips, K. M., Kelly, M. A., Farber, R., & Dube, W. V. (2015). A statewide survey assessing practitioners' use and perceived utility of functional assessment. *Journal of Applied Behavior Analysis, 48*(4), 830–844. <https://doi.org/10.1002/jaba.259>
- Skinner, B. F. (1968). *The technology of teaching*. Appleton-Century-Crofts.
- Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD: A review. *Remedial and Special Education, 22*(2), 113–121. <https://doi.org/10.1177/074193250102200205>
- Teaching Resource Center. (n.d.). *Dolch List 4th Grade*. Teaching Resource Center. <https://janelldevries.weebly.com/uploads/2/2/1/3/22134306/fourth-grade-sight-word-list.pdf>.
- Vargas, J. S. (1984). What are your exercises teaching? An analysis of stimulus control in instructional materials. In W. L. Heward, T. E. Heron, D. S. Hill, & J. Trap-Porter (Eds.), *Focus on behavior analysis in education* (pp. 126–141). Columbus, OH: Merrill
- Wagner, D., McComas, J. J., Bollman, K., & Holton, E. (2006). The use of functional reading analysis to identify effective reading interventions. *Assessment for Effective Intervention, 32*(1), 40–49.
- Wagner, D. L., Coolong-Chaffin, M., & Deris, A. R. (2017). Comparing brief experimental analysis and teacher judgment for selecting early reading interventions. *Journal of Behavioral Education, 26*(4), 348–370. <https://doi.org/10.2307/44987399>
- Watkins, C. L. & Slocum, T. A. (2004). The components of Direct Instruction. In N.E. Marchand-Martella, T. A. Slocum, & R. C. Martella (Eds.), *Introduction to Direct Instruction* (1st ed., pp. 28–65). Pearson
- Wilder, D. A., Cymbal, D., & Villacorta, J. (2020). The performance diagnostic checklist-human services: A brief review. *Journal of Applied Behavior Analysis, 53*(2), 1170–1176. <https://doi.org/10.1002/jaba.676>
- Wolery, M., Gast, D. L., & Ledford, J. R. (2018). Comparative designs. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (3rd ed., pp. 283–334). Routledge.
- Wolery, M. (2013). A commentary: Single-case design technical document of the What Works Clearinghouse. *Remedial and Special Education, 34*(1), 39–43. <https://doi.org/10.1177/0741932512468038>
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203–214. <https://doi.org/10.1901/jaba.1978.11-203>
- Worksheet Genius. (n.d.). *Design Your Multiplication Worksheet*. WorksheetGenius. <https://worksheetsgenius.com/design/multiplication-worksheet/>.
- Ysseldyke, J. E., & Christenson, S. L. (2002). *The Functional Assessment of Academic Behavior (FAAB)*. Sopris West: Longmont, CO.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Tyler-Curtis C. Elliott<sup>1</sup>  · Andrea M. Zawoyski<sup>2</sup>  · Kevin M. Ayres<sup>1</sup> 

Andrea M. Zawoyski  
andrea.zawoyski@choa.org

Kevin M. Ayres  
kayres@uga.edu

<sup>1</sup> Center for Autism and Behavioral Education Research, University of Georgia, Rivers Crossing, 850 College Station Rd Bldg. 2, Athens, GA 30605, USA

<sup>2</sup> Children's Healthcare of Atlanta, Atlanta, GA, USA