

Case Study of the Influences on Alternate Assessment Outcomes for Students with Disabilities

Meagan Karvonen
Western Carolina University

Claudia Flowers, Diane M. Browder,
Shawnee Y. Wakeman, and
Bob Algozzine
University of North Carolina at Charlotte

Abstract: The purpose of this year-long, collective case study was to investigate what influences contributed to alternate assessment outcomes for students with significant disabilities. This study of seven students and teachers in two school districts revealed seven main factors that contributed to students' scores on the state's alternate assessment, including resources, curriculum, instructional effectiveness, teacher and student characteristics, data collection and compilation, and features of the state's assessment and accountability system. Implications of this study are discussed in light of current educational reform efforts and related legislation.

In the last decade, states have moved toward standards-based educational reform including the use of large-scale accountability systems to measure student achievement. The Individuals with Disabilities Education Act Amendments (IDEA, 1997) and No Child Left Behind (NCLB, 2002) require the inclusion of all students in state and district assessments. Both federal laws specify the provision of an alternate assessment for students who are unable to participate in state or district assessments with or without accommodations. Alternate assessments provide an option for students with significant cognitive disabilities to demonstrate progress on state standards in academic areas including language arts and mathematics. Since most states did not begin to use alternate assessments until 2000, little research exists on the implementation of high

stakes progress monitoring systems (Browder, Spooner et al., 2003). Given the challenges in developing a statewide procedure for alternate assessment and teachers' potential difficulty in translating this procedure to assess individual students, information is needed on actual practice and alternate assessment's influence on outcome scores.

Variables that Influence Student Scores

One fundamental purpose of an accountability system is to determine whether students are meeting state standards for school achievement. When students do not meet expectations, more information is needed on the reasons for this failure in order for school reform to be effective. Students may perform poorly on outcomes because of program variables. There may be a mismatch between their curriculum and state standards, teacher-made assessments used in many states' portfolio assessments may not tap students' actual progress, or instruction may be ineffective. Prior professional writing, while not addressing alternate assessment per se, has pointed to these potential concerns in programs for students with severe disabilities including curriculum shortcomings (Knowlton, 1998), problems in data collection systems (Grigg, Snell, & Lloyd, 1989), and ineffective instruction prior to learning data-based decision making (Farlow

Support for this research was funded in part by Grant No. H324C010040 of the U.S. Department of Education, Office of Special Education Programs, awarded to the University of North Carolina at Charlotte. The opinions expressed do not necessarily reflect the position or policy of the Department of Education, and no official endorsement should be inferred. Correspondence concerning this article should be addressed to Meagan Karvonen, Western Carolina University, Department of Educational Leadership and Foundations, 250 Killian Building, Cullowhee, NC, 28723. Email: karvonen@email.wcu.edu

& Snell, 1989). Research on educational accountability in general education has identified other variables that may influence student achievement, including student demographics like race and income level and teacher variables like level of training (Darling-Hammond, 2000).

In a statewide survey of teachers using alternate assessment, Kampfer, Horvath, Kleinert, & Kearns (2001) identified a number of variables associated with student scores, including student involvement, portfolio elements being embedded in instruction, time spent on portfolios, perceived benefit to students, and teacher experience. Although teachers reported spending significant hours outside of classroom time on portfolios, this time only minimally related to outcome scores; instead instructional variables like embedding the portfolio elements influenced outcome scores.

In a conceptual model on alternate assessment, Browder, Fallin, Davis, & Karvonen (2003) proposed six factors that may influence alternate assessment scores: (a) technical quality of the assessment, including how it is scored, (b) student characteristics (e.g., behavioral, medical, family), (c) resources the teacher has for doing alternate assessment, (d) access to the general curriculum, (e) use of data collection systems, and (f) instructional effectiveness. The first three factors go beyond the classroom, while the last three are generally under the teacher's control. In the conceptualization of this model, Browder et al. did not propose how these variables may interact to influence outcomes or the relative strength of each variable. These six areas were the starting point for investigating influences on alternate assessment outcomes in this study.

The purpose of this collective case study was to examine factors that contributed to alternate assessment outcomes among seven teachers and their students with significant cognitive disabilities who participate in alternate assessment. This year-long, qualitative study focused on teachers and students who previously had obtained high alternate assessment scores. With this emphasis, we were able to understand how instructional programs were successful for specific individuals, what evidence of student performance was included in

the alternate assessment and how it was presented, and what individual and school factors influenced the experiences of these students and teachers involved in the alternate assessment.

Method

A multiple case study method was used to investigate factors that contributed to student outcomes. A case study is an exploration of a bounded system through detailed, in-depth data collection involving multiple sources of information (Creswell, 1998). In this research, case studies allowed a deeper consideration of the experiences of specific students and teachers. By examining multiple cases, we were able to understand commonalities shared by teacher-student pairs as well as important distinctions in each pair's experiences with alternate assessment.

Context

This study was conducted in two districts with teachers who were assessing their students with the state's Alternate Assessment Portfolio (AAP). The AAP, in its third year of statewide implementation, is a portfolio used to assess students who are unable to participate in other state standardized assessments even with accommodations. Each student's IEP team determines whether the student should be assessed with the AAP. The AAP includes documentation of student knowledge and abilities related to his or her IEP goals. In the year this study was conducted, the portfolio was organized around four domains: community, communication, career/vocational, and personal/home management. These domains were aligned with academic areas of reading, math, and writing. Teachers selected six IEP objectives to document mastery for the alternate assessment. To be considered proficient, a student had to reach the criterion for mastery, maintain this level of performance for most of the last days of the school year, generalize the skill across one person and two settings (or two persons and one other setting), and initiate the skill (e.g., unprompted correct or used skill in response to natural cues). To be distinguished, the student's mastery had to be consistent (nearly all days) with more gener-

alization (across three people and three settings). A portfolio quality score also was assigned based on the completeness of the portfolio, clarity of evidence, age appropriateness of tasks, and IEP links to the task and evidence.

Participants

Teachers and students were purposefully selected based on a combination of past AAP scores and diverse backgrounds. We were especially interested in factors associated with *positive* alternate assessment outcomes (i.e., high scores), so we recruited teachers whose portfolios received high scores in the previous year. Five teachers and students from a large, urban district and two teachers and students from a smaller, more rural district participated in the study. Three teachers taught in one of two separate schools for students with the most significant disabilities, while the other four taught in self-contained classes within regular schools. All participating teachers had received professional development on such topics as collecting student performance data, compiling portfolios, and using portfolio data to improve instruction throughout the year.

Participant characteristics are summarized in Table 1. Students in the study were between the ages of 8 and 13 and included those of African American, European American, and Latino ancestries. Their disability labels were autism or moderate, severe, or profound mental disability. Teachers in the study ranged from those with emergency certification and no special education degree to teachers with graduate degrees in special education and 20 or more years of teaching experience. All participants' portfolios developed during the year of this study were rated "satisfactory" or "superior" for portfolio quality, and all students received achievement level scores of III or IV (proficient or distinguished) in reading and math. Four students had participated in the AAP in the previous year; their prior portfolios were all scored proficient or distinguished.

Data Collection

Data were collected during periodic school visits throughout the year. A data collection

protocol was developed and pilot tested by the researchers prior to the beginning of the study. (A copy of the protocol is available from the first author.) Data sources included documents (IEPs, portfolios contents), quarterly classroom observations, quarterly teacher interviews, and observations of IEP meetings. Data collection methods remained the same across school visits with the exception of teacher interview questions. Interview questions were revised in the second and third marking periods based on analyses of the previous quarters' data. Interviews in the first three marking periods were semi-structured while the fourth quarter interview consisted of two broad questions and a less structured format. Three researchers collected the data. Interview responses and observation notes were handwritten in the first three marking periods. Fourth quarter interviews were audio-taped. Observation and interview notes were transcribed and reviewed by the researchers for accuracy prior to analysis.

Data Analysis

A team of four researchers analyzed data on an ongoing basis and met throughout the school year to discuss findings and refine the focus of future data collection. Transcripts were evaluated using qualitative data analysis methods described by Miles and Huberman (1994). Interview data were chunked according to a coding scheme developed and expanded from Browder, Fallin et al.'s (2003) conceptualization of influences on alternate assessment outcomes. Qualitative data analysis software was used to organize and categorize data and generate reports for each theme for use in subsequent analysis.

Data from each case (within-case analysis) were analyzed using a holistic technique (Yin, 1989) in which a detailed description was developed and themes were identified from the descriptions. After each case was analyzed, a cross-case analysis (Miles & Huberman, 1994) was conducted to examine common themes as well as unique findings within separate cases. The findings reported in this manuscript are based on the cross-case analysis.

Several steps were taken to enhance dependability of the data. A data collection protocol was used to ensure consistency of data

TABLE 1

Characteristics of Participating Students and Teachers

		<i>Student</i>	<i>Teacher</i>	<i>Portfolio Scores</i>				
	<i>Age</i>	<i>Disability Label, Ethnicity and Gender</i>	<i>Years Teaching</i>	<i>Degree, Certification</i>	<i>Reading</i>	<i>Math</i>	<i>Writing*</i>	<i>Portfolio Quality</i>
A	11	Profound Mental Disability, European American Male	3	Bachelor's outside SE MSE	4	4	4	Superior
B	8	Autism, African American Male	15	BSE Behaviorally/Emotionally Disabled, Mentally Disabled, Learning Disabled	4	4	N/A	Superior
C	13	Moderate Mental Disability, Latina Female	4	Bachelor's outside SE Coursework toward certification in Severe/ Profound	4	4	4	Satisfactory
D	8	Autism, European American Male	11	BSE, pursuing doctorate in SE Behavioral/emotional disability, cross categorical, mental disability, orthopedic disability	3	3	N/A	Superior
E	12	Severe Mental Disability, African American Male	20	Bachelor's outside SE, MSE Learning Disabilities, Mental Retardation K-12, Behavior and Emotional Disabilities, Severe/Profound Disabilities	4	4	4	Superior
F	8	Moderate Mental Disability, African American Male	10	Bachelor's outside SE, MSE Elementary education, K-12 mental disabilities, severe/profound disabilities	4	4	N/A	Superior
G	10	Moderate Mental Disability, African American female	25	MSE Mental disabilities, severe/profound disabilities	4	3	4	Superior

Note. BSE = Bachelor's degree in Special Education, MSE = Master's degree in Special Education.
 * Writing scores were based on specific expressive communication goals. Not all students had writing scores.

collected across participants and visits throughout the school year. Consistency of coding was insured by having all interviews coded by two researchers and disagreements discussed until resolutions were reached. A fifth researcher who was familiar with the study but not directly involved in data collection or analysis conducted periodic peer reviews of the analyses. Credibility of the find-

ings was enhanced through triangulation of data sources and researchers. A summary of the findings was shared with participating teachers, who were given an opportunity to respond to the researchers' analysis. Finally, the selection of teachers and students with diverse backgrounds, and the cross-case analysis of their data, may promote transferability of findings to other special educators and

their students who participate in portfolio-based alternate assessment.

Findings

Findings from the cross-case analysis generally supported the conceptual model for score influences proposed by Browder, Fallin et al. (2003), although one factor (teacher characteristics) was added. The seven general influences included (1) resources, (2) teacher characteristics, (3) curriculum, (4) instructional effectiveness, (5) quality of student data, (6) student characteristics, and (7) characteristics of the state's assessment and accountability system. In this section we describe participants' common and unique experiences with each influence. Findings from specific students or teachers are named by letters (A–G) that correspond with the letters in Table 1.

Resources

The category called “resources” was broadly defined to include any form of support that teachers used to create curriculum, provide instruction, or collect or compile portfolio information as well as any means they used to improve their own ability to do any of those activities. We found that teachers used four types of resources: materials, professional development, assistance from other people, and time.

Materials. Participating teachers were able to obtain or create some of the resources they needed to teach students portfolio skills and prepare the portfolios. Teachers used computers at school and at home to create data collection sheets, write summaries, and graph data. Teacher A used assistive technology including an augmentive communication device, a timed switch, and a touch screen computer. Several teachers reported that they received some instructional materials at the beginning of the year, but others purchased additional materials with their own money.

Several teachers described problems with limited resources, ranging from broken computers, to inappropriate materials, to lack of funds for community-based instruction on AAP objectives. Some resource concerns typical in schools were also identified as hindering

instruction, including limited access to a copy machine and small classrooms. Other AAP-specific materials that teachers said they wanted to have were pamphlets explaining the AAP to parents and a manual with guidelines for linking AAP objectives to the state's standard academic curriculum.

Professional development. All seven teachers in this study received training on alternate assessment that went beyond the state's standard orientation session on the AAP. All seven teachers participated in an OSEP-funded model demonstration project to help teachers learn to use behavioral data to improve instruction throughout the year (Browder, Karvonen, Davis, Fallin, & Courtade-Little, 2005). The project consisted of several day-long workshops and technical assistance provided 1:1 in teachers' classrooms. Two of the teachers participated in the project for three years, three participated for two years, and two received more limited assistance through a project replication phase. The timeline for training and technical assistance provided by the project in itself helped teachers monitor their pacing:

The [project] has helped us pace deadlines of the portfolios in our classroom better because otherwise, we make it up ourselves—our pacing of the portfolio, how well have you kept up, when you're checking on them again, changing instruction. But being in a [project], you have deadlines, you know about the next visit. I should have this by the next visit, he should be here. So actually, that has helped.

Participants also reported benefits from the project including access to a wide range of data sheets, assistance in determining *how* to teach students to generalize their skills, and using flexible and creative ideas to meet portfolio demands.

In addition to this project, some teachers received even more training. For example, Teacher F was trained to score portfolios at the state level; she used this training to make sure her own portfolios contained what scorers were looking for, and also to help other teachers at her school with their portfolios. This teacher summarized the importance of training in her success with alternate assessment this way:

... having the administrative support and the right training so you understand the system of the portfolio process, how it works, and the necessary requirements that are in it so that you can carry those out and display them nicely, of course, in a portfolio that's put together well. As long as you understand that, you know what you're doing.

Assistance from others. Teachers in this study received several types of support from others in their school communities. At a minimum, other teachers and staff served as people with whom students could practice generalizing their alternate assessment skills. For example, students practiced using money in the cafeteria and worked on literacy skills with the librarian. Some participants worked with other special educators using more of a team approach to completing portfolios. Teacher B collaborated with one other teacher to check each others' data sheets and review portfolio contents prior to submitting them for scoring. Teachers E and F worked with others at their school to brainstorm instructional strategies, collect generalization data, problem solve, and provide moral support.

Opportunities for teacher collaboration were dependent to some extent on the support of principals. All participants' principals were described as supportive to some degree. One principal allotted teachers a certain number of release days to work on portfolios based on the number of portfolios for which they were responsible. Other principals allowed release time only if teachers could find staff to cover their classes. Release time could not be arranged in classrooms where students had significant medical conditions (e.g., seizure disorders). In other cases, principals were viewed as supportive although they did not provide tangible support. One teacher noted that her principal "heard us belly-aching about all the personal time we were doing on these things and he feels for me, but then there's nothing he can really do to help us."

Regardless of the amount of tangible assistance provided by principals, teachers attributed support of other school staff to the message sent by school leadership. In a school where the principal made it clear to general education teachers that alternate assessment scores counted, teachers and administrative

staff helped with data collection for generalization. In another school, a new principal emphasized academic instruction time and eliminated a policy that previously allowed teachers some time to work on portfolios during the school day.

Most study participants identified paraprofessionals as a critical resource for their success with the alternate assessment. Paraprofessionals provided support by preparing materials, collecting and recording data, and providing instruction on alternate assessment skills.

I think to do a decent portfolio, you have to have a decent coworker or teacher assistant ... I've talked to other teachers who have coworkers who don't want to take data or don't understand or don't want to graph the portfolio way of taking data. It makes it difficult. Without [my assistant] I wouldn't have gotten it done ... We wouldn't have gotten the consistency and that regular data collection.

Two participating teachers' assistants attended training about the alternate assessment portfolio. Other participants reported that after they attended workshops, they came back and shared with their assistants what they learned. Sharing information and techniques for the alternate assessment with paraprofessionals yielded secondary perceived benefits, such as a greater sense of professionalism. Teachers also reported that their assistants were more willing to do work to contribute toward the alternate assessment portfolio when they understood that it was a requirement mandated by the state rather than a requirement imposed by the teacher herself.

Participants' varied experiences with support from colleagues did not appear to depend on whether they were in a self-contained or mainstream school. One self-contained school had very strong principal support for materials, training opportunities, and release time. Teachers believed the principal provided these resources in support of her high expectations for teachers' success. This principal fostered a school culture in which everyone was responsible for students' learning; all staff members were expected to be available to assist with generalization and data collection.

Assistants were expected to do what teachers asked them to do, and all teachers were resources for each other. In contrast, a teacher at another self-contained school was not able to persuade other teachers or therapists to consistently collect data or provide instruction on AAP skills.

Time. An underlying theme in all of the teacher interviews was the extensive amount of time required to plan, implement instruction, collect data, and compile portfolios. Some concerns about time were related to the balance between instruction on AAP skills and other content areas. For example, several teachers felt that instruction on other IEP skills was decreasing because they needed to spend so much time on AAP skills. The overall amount of time spent on alternate assessment also was a concern. Teachers reported that they frequently needed planning time to evaluate their students' progress and plan for generalization or other instructional changes. In the months immediately preceding the deadline to submit AAPs, teachers spent many hours recopying data sheets, checking for clarity, and making sure the information in the portfolio was clear to a scorer who would not be familiar with the student. All seven teachers in this study reported that they spent personal time on evenings and weekends working on their portfolios.

Teacher characteristics. Teachers who participated in this study came from different backgrounds and used varied teaching approaches. There was no clear pattern that suggested teachers with higher degrees or more years of experience were more successful with alternate assessment. Instead, the most consistent finding across all of the participants was their strong desire to do everything they could to help their students succeed. In spite of limited resources and other barriers, these teachers persisted. They sacrificed their personal time, sought out additional training, and accessed as many resources as possible in order to support their students' growth. They all maintained very high expectations for what their students could accomplish. As Teacher E noted:

I'll call them my 'academically gifted, severe and profound' students. They would never be where they were if it wasn't for the way

we teach them. I feel that confident in myself as a teacher and my assistants, in the level of expectations we have, and how we push them, that their skills are as high as they are.

In addition to high expectations and motivation to help their students succeed, participants also tended to have strong organizational skills. These skills were evident during classroom visits, as teachers used methods they had devised for organizing data sheets, incorporating data collection into their daily schedules, and periodically reviewing and compiling data. During interviews, teachers spoke of organizational skills when reflecting on the initial portfolio preparation at the beginning of the school year, in their ability to easily locate and summarize data and use them to adjust instruction, and in the context of juggling the instructional and data collection requirements for multiple students.

One other commonality among the teachers in this study was that they accepted the requirements of the AAP and consistently strove to meet those requirements. They adjusted their methods when the state announced changes in the AAP requirements. Rather than ignoring the AAP or making minimal effort to meet its requirements, teachers looked for examples of how to provide what scorers would be looking for. One teacher referred to this as "knowing the rules and knowing how to play the game."

Participants' motivation, experiences, and knowledge about the alternate assessment system made them resources for other teachers. Teacher C provided training to others in her school, and teachers D and E provided technical assistance to teachers from other school districts on identifying portfolio skills and designing data collection methods.

Curriculum

Because this research was conducted during the year when the state made the shift from a functional to academic focus in the alternate assessment to be consistent with the No Child Left Behind requirements for assessing progress in reading and math, we had the unique opportunity to observe how teachers responded to this shift in curricular emphasis

in preparing the portfolios. Several teachers noted that they had written their students' most recent IEPs with the earlier portfolio guidelines in mind so they could easily select skills for the AAP. Because of these guidelines for the AAP, teachers began the year working toward the students' mastery of functional skills. When they were asked in late fall to link those skills to academics, these teachers did not reopen the IEPs to create goals that aligned with specific standards in the general curriculum. Instead, they sought ways to use the skills already identified and documented.

Teachers in this study used varied approaches to adapt their existing goals to meet the requirement for documentation of progress in reading and math. Some teachers, such as Teacher F, had a history of teaching functional academics even before alternate assessments were implemented. Finding ways to link to state academic standards was not as difficult for teachers who were already teaching some reading and math. However, teachers who only had targeted skills from a functional curriculum, found it more challenging to find connections to the state's academic standards. As one teacher put it, "Trying to figure out the [academic] essences for tossing [a] ball is just goofy."

Teachers in this study possessed varying opinions about the value of the new focus on academics in the portfolio. Two veteran teachers who were strongly in favor of functional curricula did not view the general curriculum link as particularly useful:

I think the more relevant the goals are, the more sense it makes to the child and I think sometimes the more progress you make. Sometimes the kids aren't always clear on why they're having to count by 5's when your ultimate goal is to tell time.

Teachers with this viewpoint tended to see functional goals as a means of addressing the student holistically, and they did not want to lose that focus because of the academic emphasis required by the AAP. One teacher said that to abandon functional goals would be unethical.

In contrast, the newest teacher saw value in focusing on academics because it added legitimacy to her instruction; she liked not being

considered a "modified babysitter" because her students were working on math or language arts. She did not see the academic and functional foci as competing. Instead, she discussed the blending of functional and academic curriculum in the following way:

I know that there are some concerns that we're moving away from functional curriculum skills—that some of these kids need just to be independent as much as they possibly can in society, but what I have done is I have just incorporated these goals into the normal daily routine . . . I might tie in numeracy and literacy somewhere as also a functional goal. I kind of make it a two-in-one type [of goal].

For example, this teacher included mathematics in a personal management skill by having her student use a switch to measure the amount of Pediasure poured.

Instructional Effectiveness

In addition to curriculum choices, there were several features of teachers' instructional approaches that seemed to contribute to student performance: (1) direct instruction, (2) data-based decision making, (3) opportunities for student self-determination, and (4) peers without disabilities. Not all practices were observed with all teachers. This section is intended to provide examples of each of these effective instructional practices.

Direct instruction. We use the term direct instruction here to refer to the classroom activity in which the teacher provides information and prompting and gives the student multiple opportunities to respond with teacher feedback. Direct instruction can be contrasted to independent seatwork time or working on projects with peers. During classroom observations, data were collected on the allocation of instructional time among various skills and features of specific teacher-student interactions within instructional sequences. One hundred and eleven separate activities were observed during classroom visits. In most of these activities, the teacher provided direct instruction to the student with opportunities to make the target response and receive feedback. More of these activities were focused on

instruction of the skills targeted for portfolio documentation than non-portfolio skills.

Data-based decision making. During the direct instruction, the teachers collected data on the students' responses. Teachers also used these data on their students' performance to monitor progress and adjust instruction to improve student performance. Teacher D described her approach to using instructional data this way:

I would collect a lot of data and then every couple weeks I'd try to put it together to see how we were doing. I didn't do a lot of graphing but I did calculate their percentages and things so I could kind of see where we were going . . . I'd take one notebook home and just spend a couple of hours on that notebook and that particular child's goals to see where we were and to write down my ideas of what to try next if we weren't succeeding or the next step if they'd reached some success on the first part of it and how to move on.

Other teachers used quarterly reports to summarize progress and plan for the next quarter. These opportunities for reflection helped teachers consider their students' current level of performance in comparison to the goal for the end of the portfolio data collection period. For example, teachers monitored the number of settings and people with which the students had generalized. Since generalization was a critical component in determining the AAP score, this monitoring procedure helped teachers stay on track and address gaps in generalization.

Related to data-based decision making were teachers' decisions to persist until students mastered the skills targeted for portfolio documentation. This approach contrasted with some teachers' approaches prior to alternate assessment. Several teachers noted that before alternate assessment, they would discard goals mid-year when students were not progressing, in favor of goals for which the students were more likely to show growth. With alternate assessment requiring demonstration of progress over time on goals that lasted the entire year, teachers did not have the luxury of changing the content of instruction. Instead, they found creative ways to work

through periods of limited progress to help students improve.

Student self-determination. Elements of student self-determination in the instructional process, including self-evaluation, self-reinforcement, and choice making, were included in several teachers' instructional approaches. A few students were aware of their correct and incorrect responses, and teachers gave them opportunities to self-monitor (e.g., putting stickers on correct work or graphs showing student progress, erasing work to show awareness that it is incorrect). Some students used wall charts to self-schedule throughout the day.

Peers without disabilities. Teacher F had students from other classrooms come to her class and work with students who were assessed with the AAP. These peers without disabilities provided additional instruction, helped with data collection for generalization purposes, and served as sources motivation for student F. Student A also worked with a peer on instructional goals. Student B, who was in a classroom with other students with autism who were following a more academically rigorous curriculum, also benefited from peer influences. His teacher described evidence that he wanted to emulate his classmates to learn academic skills (such as writing his name).

Quality of Student Data

Data collected and included in students' portfolios was perceived by teachers to be important for high scores on the AAP. By the third quarter, several teachers were identifying the quality of the data as the factor they believed would most influence their students' AAP scores. The state designed the AAP to focus on ongoing data collection for skills selected from the IEP to show growth on state standards. Three facets of the portfolio data emerged as important concepts: (1) translating IEP objectives into measurable objectives consistent with the AAP guidelines, (2) the process of data collection itself, and (3) compiling data into the final portfolio.

Creating measurable objectives. Each teacher in the study talked about the importance of selecting and writing IEP objectives for portfolio documentation that would be clearly understandable and measurable. As one teacher

remarked, she wanted her data to be “objective, measurable, and clear.” Teachers also wrote objectives to target mastery during the year of portfolio documentation. The teacher then selected a data system for portfolio documentation based on the wording of the objective (e.g., if the target was mastery of 20 sight words, a data collection system was selected for tracking 20 sight words.) Once objectives were established, teachers changed the objectives only if a new IEP was written or if the state required a change in the information provided.

Collecting data. Every teacher in this study realized the importance of collecting data frequently and making sure the data sheets clearly reflected student performance. As Teacher B noted, “a child could do wonderful [work], but if the teacher doesn’t do a good job of keeping the data, it’s not going to show that the child has made progress.”

The state’s guidelines recommended collected data two times per week on each skill. When translated across skills and students, this meant collecting a large volume of data. Teachers in the case study met or exceeded the state’s guidelines for the amount of data to be kept. Many of the teachers collected data on a daily basis and created a system to make sure data were recorded when the student was practicing the skill. Some used clipboards to organize data for each student, and others used color-coded systems to indicate which days certain types of data were to be collected. One teacher used a wall chart with cards for each student and each skill. Teachers also used computers to design data sheets or generate labels as time-saving methods.

With AAP scores determined in part by evidence of generalization across settings and people, teachers also had to work with colleagues to make sure data were collected by others and in different settings. Others in the school who practiced the skill with the student included general education teachers, other special education teachers, librarians, therapists, receptionists, and nondisabled peers. Teacher D used condensed data sheets to record data for multiple students when the class traveled as a group to other settings in which skills were generalized. The systems teachers created for others to record data varied based on the skill and willingness of the

other collectors, and teachers’ comfort level with having others collect data. Besides the teacher, paraprofessionals were most frequently responsible for data collection and instruction. Some teachers trained their paraprofessionals to record data directly on data sheets, while others asked the paraprofessional to record it on scrap paper or tell the teacher how the student performed. In some cases teachers created simplified data sheets for others to use and then transferred the data to the sheets that were included in the portfolio. For example, two students without disabilities came in to work with student G on sight words. The teacher taught them how to use various materials to work with the student, and created a simplified data sheet with the sight words and a place for the students to write down whether or not the student identified the word correctly. Another teacher sent home similar sheets for parents to record data.

Compiling the data. While behavioral data was the primary source of evidence for the portfolios, teachers also provided anecdotal notes, quarterly summaries of student progress, graphs of the data, and evidence of students’ work. Teachers in this study took seriously the clarity of their students’ data and spent considerable time re-evaluating the contents of the portfolio to make sure it reflected the student’s performance. Participants re-copied data sheets to make handwriting clearer, checked notation about student performance to make sure it was consistent across sheets, and re-computed totals in repeated trials and task analysis goals to make sure they hadn’t made addition errors. One teacher highlighted portions of the portfolio that she wanted the scorer to notice. Another teacher printed graphs in color after she noticed that in a previous year one of the portfolios she submitted with color graphs was rated “superior” in quality while the others (with black and white graphs) were all scored as “satisfactory.”

Teachers’ awareness of the scoring system and what scorers were looking for influenced the way they compiled portfolios. Several teachers said they wanted their portfolios to be as clear as possible and to make obvious the points they wanted scorers to notice. Prior to submitting completed portfolios, several

teachers in the study had other teachers review their portfolios to make sure the information would make sense to someone unfamiliar with the student.

Student Characteristics

While the AAP is designed to focus on performance of individualized goals regardless of students' disabilities, we originally thought that students with significant medical or behavioral problems would have a more difficult time demonstrating mastery. In fact, teachers in this study did not view their students' levels of cognitive functioning as influencing their alternate assessment scores. With a state system in which the teachers picked skills to document progress on state standards from the IEP, even students with the most significant disabilities were able to achieve mastery.

On the other hand, teachers did identify several student characteristics that were problematic for data collection purposes. Two teachers noted the difficulty of keeping students with autism focused on the tasks to complete instruction and data collection. Student C exhibited behaviors for the first few months of school that competed with instructional time. Eventually this teacher learned to encourage participation in instruction and to collect data whenever this student was ready to work rather than only being prepared to take data at set times that might be lost to disruptive behavior. Two teachers also mentioned the frustration they felt when medically fragile students were absent and unable to make progress on their goals:

I had a student last year. .he showed growth and [the state] said he showed growth. But he regressed because he was out four months in the hospital. .It really kind of upsets me because it's not like he can make up that time like a regular ed student can make up that time.

Data collection was one area in which students' disabilities were perceived to influence alternate assessment. Some teachers noted that it was more difficult to define correct, independent responses in students with limited expressive communication:

... If you have a student with a learning disability, and it is something really neat and clean, you can put a check or minus, or "yeah, you got that right. Now let's move on to this." But when you're working with students with severe disabilities, you're looking for movement of eyes. Well, I may think he moved his eyes, but you may not think he moved his eyes. How much is he going to move them and where is he going to move them to?

According to participating teachers, their students were aware of and engaged in data collection procedures to varying degrees. Several students made marks or placed stickers on the tops of completed work products. Some students were aware that teachers were writing down information about what the students were doing. One student, who was used to putting stickers on completed data sheets, pointed out to his teacher a data sheet for which she had forgotten to let him add a sticker. The teacher said this student did not understand what the data meant, but that he knew he needed to perform the skill correctly. Another teacher described her student's awareness of the teacher's notation for an incorrect response, and her ability to correct herself when she saw the teacher write a minus. This student was also aware of her writing errors and erased her paper to show she made an error. Teacher F reported that the student wanted to see his data sheets to know how well he performed. On tasks where he did not self-correct mistakes, he was still aware of the mistake. Sometimes teachers reported that their students wanted to do better because they were competitive with other students.

State Assessment and Accountability System

There were three general ways in which the AAP and the accountability system influenced students' outcomes: (1) communication about requirements (2) teachers' awareness of technical quality and scoring of the AAP, and (3) the way that the AAP was included with other students' assessment results for school accountability purposes.

Communication about requirements. This case study was conducted when the state's alternate assessment process was in its early stages and

federal guidelines were still emerging. Teachers were frustrated because they wanted to develop strong portfolios, but had to cope with changing requirements. They credited the model project with helping them keep track of requirements. Most knew teachers not in the model who were confused about the alternate assessment requirements and not prepared to provide the appropriate documentation of student progress. Methods for reporting scores were also still in development. Some of the teachers had not received their students scores (e.g., if the student went to a new school the next year).

Scoring and technical quality of the AAP. During interviews, we asked participants about the accuracy and consistency of AAP scores. A teacher who had been using the AAP since its pilot year of implementation noted that her confidence in the scores was increasing as the state refined the scoring process to include multiple raters and a minimum level of acceptable interrater reliability. However, others viewed the scoring process as subjective. While AAPs were scored separately for portfolio quality and evidence of student performance, teachers indicated that it was difficult to obtain a high student score without also having a well-constructed portfolio. Participants recognized that the need to have nice looking portfolios with clean data sheets to show that their students achieved proficiency.

Another issue teachers noted was how much each student's mastery depended on the expectations they set. One teacher said, "I wouldn't pick something they know already because the student needs to show growth." In contrast, teachers were realistic about targeting mastery by the March date when portfolios were submitted. For example, one of the skills documented for student B's portfolio was writing his first name. His teacher decided that the shortened version of his name (e.g., 'Chris') would be something he would likely master by the time portfolios were due in March. Her own goal was for him to write his entire name ('Christopher'), and she continued to work on the longer name until the end of the school year in late May.

AAP and accountability. This study was conducted prior to the federal guidelines on alternate achievement standards (Title I—Improving the Academic Achievement of the

Disadvantaged, 2003), but in this state alternate assessment scores were counted in school accountability scores and could be counted as proficient. When asked how they would react if AAP scores were automatically considered not proficient for accountability purposes, Teacher E's response reflected the sentiment of other participants:

The experience would have been extremely different. I would not have spent the time and effort on all areas of the portfolios and the quality to make sure scorers could understand my portfolio if the student was only going to score in the 1 or 2 range. Why bother? I would continue to have high expectations for my students and teach them but I would not spend numerous hours, days, after school time, and weekends constructing data sheets which were understandable to [others] . . .

Another teacher noted the changes she saw in other educators once AAP scores started counting for accountability purposes:

I think our expectations are higher. You feel like you're doing something when you have people who are actually interested in what your students are doing and when you're expected to be accountable for what you do in your classroom, it kind of raises the bar a little bit. Not that we didn't do it before, but before it was like, why does this matter? Nobody cared what we did . . . or cared how these kids did.

In contrast to studies in which general education teachers view accountability and high-stakes tests as having negative impacts on their students (c.f. Abrams, Pedulla, & Madaus, 2003), the teachers in this study perceived the inclusion of their students in accountability models as enhancing the status and visibility of their students as learners.

Discussion

Experiences of seven teachers and students with histories of positive alternate assessment scores were examined in this study. While the seven contributing influences we identified appeared to be important for all of the teacher-student pairs, a close examination revealed

differences in how those factors manifested themselves. The extent and sources of support teachers received for effective instruction and the AAP varied, as did teachers' perceptions about the roles of academic and functional goals within students' educational programs. However, consistent across the cases were the strength of the instructional programs and the effort teachers put forth to develop high-quality portfolios.

The primary contribution of this study is to provide information on what may influence alternate assessment scores. In earlier research based on teacher interviews, Kleinert, Kearns, and Kennedy (1997) found a relationship between alternate assessment scores and the quality of the students' educational programs in Kentucky. In a survey conducted in that same state, Kampfer et al. (2001) found that scores were related to student involvement, the extent to which portfolio items were embedded in instruction, and teacher perceptions of the benefit of the portfolio to the student. Findings from the current study support earlier research and extend our understanding about influences such as resources, the process of compiling portfolios, and student and teacher characteristics.

The current case study examined the educational programs of students who represented a wide range of disabilities and cognitive levels, all of whom had scored as proficient in the state alternate assessment system. This illustration of positive outcomes for students of varying backgrounds supports Kleinert et al.'s (1997) study in which no relationship was found between students' cognitive level and alternate assessment outcome scores. These studies contrast with research on assessments used in general education, in which student characteristics (e.g., poverty, non-English language status, and minority status) are predictors of outcomes scores (Darling-Hammond, 2000). Tindal et al. (2003) also found a relationship existed between student characteristics and alternate assessment scores in a state that uses a performance-based method of alternate assessment.

Does the influence of teacher characteristics in alternate assessment support or refute its validity as a means to document student achievement of state standards? If teachers, who select assessment items through a portfo-

lio process, influence outcomes merely through constructing the assessment, that assessment's validity in documenting student achievement of state standards is questionable. In contrast, if teachers who select the tasks to document state standards make the assessment accessible to students with the most complex disabilities who might not be able to show achievement in other formats, it may be viewed as a means to make the assessment fair for students with significant cognitive disabilities. What is needed is some step in the assessment scoring process to check that the items selected for the student do indeed reflect both the academic content of the state's standards and the expected achievement level. As states move towards the use of alternate achievement standards, the guidelines that emerge may make these assessments less susceptible to validity concerns related to teacher selection of assessment items.

Another way to view the influence of teacher characteristics is that it is not so much the selection of assessment items as the quality of instruction that produces the outcomes for students. In order to meet the NCLB requirement that all subgroups of students achieve state expectations, the quality of the educational program, teacher characteristics, and school resources will have to exert stronger influence than student characteristics that have traditionally been risk factors for academic failure. The fact that teacher characteristics are strongly influencing alternate assessment scores may be welcome news if it means that the quality of the instructional program is making it possible for students with the most significant disabilities to meet state expectations.

Given that teacher characteristics may have a desirable influence on alternate assessment outcomes, it is interesting to consider the specific characteristics exhibited by teachers in this case study. Teachers strongly influenced alternate assessment outcomes through their selection of skills for documentation, use of data collection and direct instruction, persistence in changing instruction to promote student mastery, and careful documentation of performance for external scoring. Techniques such as the use of frequent opportunities for students to respond and providing immediate feedback are effective for improving students'

engagement and learning (Sutherland, Alder, & Gunter, 2003). These teachers were part of a model project that emphasized the importance of instruction. A quantitative evaluation of all 28 students in the model (Browder et al., 2005) revealed that participants made greater gains than students in the school system that were not in the model.

Taken together, our findings along with Kampfer et al. (2001) and Kleinert et al. (1997) do suggest that instructional variables may influence student outcomes. In contrast, Flowers, Ahlgrim-Delzell, Browder, and Spooner (2004) found in a five state survey that teachers did not agree that instructional changes influenced students' alternate assessment scores. Instead, the survey respondents were more likely to give credit to teachers' knowledge of the alternate assessment system. Teachers participating in the case study also emphasized the importance of understanding how to effectively collect data and compile portfolios that were clear to external scorers. However, they viewed knowledge of the alternate assessment system as a means to ensure their students' learning was accurately understood—not as a method to compensate for limited student growth.

This case study supports Browder, Fallin et al.'s (2003) premise that multiple variables may concurrently influence alternate assessment outcomes. Teachers not only described instructional variables as influencing outcomes, but also noted the need for resources to implement the process and savvy in knowing how to document clearly for scoring.

Also in support of Browder, Fallin et al.'s (2003) conceptual model, this study's findings indicated that not all variables that influence student outcomes are under the teacher's control. For example, teachers in the current case study described how the principal set the tone for school's support for students in the alternate assessment process. Teachers also described limited resources, a lack of time to plan and compile portfolios, and in some cases, minimal support from other school staff members for data collection. This lack of additional resources for the alternate assessment is consistent with Wakeman, Ahlgrim-Delzell, and Browder's (in press) findings from a survey of district-level Exceptional Children's administrators. What was evident in the current

study was that this particular group of teachers found ways to compensate for variables that they could not control. The teacher who had a student with some disruptive behavior compensated by being flexible in scheduling this student's instruction. When the teachers received a mid-year change in state guidelines to document academic skills more directly than they had in the past, they compensated by finding ways to incorporate academics in their ongoing functional skill instruction. When they lacked resources, they compensated by using their own money or personal time to complete the assessments. They also recruited other staff to help with data collection.

Although these examples suggest that teachers are of primary importance in creating positive alternate assessment outcomes, caution is warranted against overgeneralizing this finding. The teachers in the current case study were selected because the students in their classrooms performed well in the state's alternate assessment. Our focus was to determine how they helped their students achieve success. The lessons from these teachers may help those who resist alternate assessment understand how it can be implemented effectively, as an integral part of a strong educational program, with the intended benefits of being "counted" in a state's accountability system.

While this study highlighted teachers' effective practices and their successes despite barriers, it did not include a comparison group of teachers with a history of creating low-scoring portfolios. Therefore, this study does not provide evidence for a causal influence of teachers' strong instructional emphasis on the students' scores. A larger scale study including teachers and students with a range of alternate assessment outcomes is necessary to understand how those with lower scores might have contrasting experiences with instruction, resources, data collection and compilation, and other factors. Such a study would need to consider the interrelatedness of the seven contributing influences, as well as the unique contributions of each influence. Some evidence of the complex interactions of the seven influences was evident in this study. However, the multiple case study design does not allow for quantifiable conclusions about the relative importance of certain factors.

It is important to note that the intense investment by teachers in this study occurred in a high stakes accountability state in which the alternate assessment scores counted in school accountability. Teachers stated that they would not have invested as much in the process if the scores could not count as proficient for their school. In the survey by Flowers et al. (2004), teachers who reported that alternate assessment scores influence their school's accountability reported more benefits for students participating in alternate assessment. As states develop and apply alternate achievement standards for students with significant cognitive disabilities, more information needs to be collected on how "counting" these students impacts their educational programs and any resulting change in their outcome scores.

It is clear from this study that resources, teacher characteristics, and instructional effectiveness are important factors that contribute to alternate assessment outcomes. The importance of these features signifies the value of providing ongoing, targeted professional development to enhance teachers' capacities to conduct high-quality alternate assessments. This professional development should emphasize instructional approaches and effective compilation of portfolio contents, but it should also focus on helping teachers persist despite influences they cannot control (e.g., administrator support, changing requirements). States will need to make efforts to equip teachers with new resources to help them be effective in meeting the demands of NCLB to help students with the most significant disabilities meet state expectations.

References

- Abrams, L. M., Pedtulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42*(1), 18–29.
- Browder, D. M., Fallin, K., Davis, S., & Karvonen, M. (2003). Consideration of what may influence student outcomes on alternate assessments. *Education and Training in Developmental Disabilities, 38*, 255–270.
- Browder, D. M., Karvonen, M., Davis, S., Fallin, K., & Courtade-Little, G. (2005). The impact of teacher training on state alternate assessment scores. *Exceptional Children, 71*, 267–282
- Browder, D. M., Spooner, F., Algozzine, R., Ahlgrim-Delzell, L., Flowers, C., & Karvonen, M. (2003). What we know and need to know about alternate assessment. *Exceptional Children, 70*, 45–61
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8*(1). Retrieved May 25, 2004 from <http://epaa.asu.edu/epaa/v8n1/>
- Farlow, L. J., & Snell, M. E. (1989). Teacher use of student performance data to make instructional decisions: Practices in programs for students with moderate to profound disabilities. *Journal of the Association for Persons with Severe Handicaps, 14*, 13–22.
- Flowers, C., Ahlgrim-Delzell, L., Browder, D. M., & Spooner, F. (2004). *Teachers' perceptions of alternate assessments*. Manuscript submitted for publication.
- Grigg, N. C., Snell, M. E., & Loyd, B. H. (1989). Visual analysis of student evaluation data: A qualitative analysis of teacher decision making. *Journal of the Association for Persons with Severe Handicaps, 14*, 23–32.
- Individuals with Disabilities Education Act of 1997*, 120 U.S.C. §1400 et seq.
- Kampfer, S., Horvath, L., Kleinert, H., & Kearns, J. (2001). Teachers' perceptions of one state's alternate assessment portfolio program: Implications for practice and teacher preparation. *Exceptional Children, 67*, 361–374.
- Kleinert, H. L., Kearns, J. F., & Kennedy, S. (1997). Accountability for all students: Kentucky's alternate portfolio assessment for students with moderate and severe cognitive disabilities. *Journal of the Association for Persons with Severe Handicaps, 22*, 88–101.
- Knowlton, E. (1998). Considerations in the design of personalized curricular supports for students with developmental disabilities. *Education and Training in Mental Retardation and Developmental Disabilities, 33*, 95–107.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103) New York: American Council on Education.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- No Child Left Behind Act of 2001*, Pub. L. No. 107–110, 115 Stat.1425 (2002).
- Sutherland, K. S., Alder, N., & Gunter, P. L. (2003). The effect of varying rates of opportunities to respond to academic requests on the classroom behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 11*, 239–248.

Tindal, G., McDonald, M., Tedesco, M., Glasgow, A., Almond, P., Crawford, L. et al. (2003). Alternate assessments in reading and math: Development and validation for students with significant disabilities. *Exceptional Children*, 69, 481–494.

Title I—Improving the Academic Achievement of the Disadvantaged; Final Rule, 68 Fed. Reg. 68698 (Dec. 9, 2003) (to be codified at 34 C.F.R. pt. 200).

Wakeman, S. Y., Ahlgrim-Dezell, L., & Browder,

D. M. (in press). Evaluation of emergent alternate assessment practices by administrators. *Journal of Special Education Leadership*.

Yin, R. K. (1989). *Case study research: Design and method*. Newbury Park, CA: Sage.

Received: 10 March 2005

Initial Acceptance: 13 April 2005

Final Acceptance: 15 June 2005