OXFORD

Journal of Deaf Studies and Deaf Education, 2023, 1–15

https://doi.org/10.1093/deafed/enad046

Empirical Manuscript

The quality of teaching behaviors in learning environments of DHH students

M. Christina Rivera, PhD[1,*], Jennifer A. Catalano, PhD[2], Lee Branum-Martin, PhD[3], Amy R. Lederberg, PhD[3], Shirin D. Antia, PhD[1]

[1]Department of Disability and Psychoeducational Studies, University of Arizona, Tucson, AZ, United States
[2]Department of Education, Flagler College, St. Augustine, FL, United States
[3]Department of Learning Sciences, Georgia State University, Atlanta, GA, United States
Corresponding author: M. Christina Rivera College of Education, Department of Disability and Psychoeducational Studies, University of Arizona, 1430 E. 2nd St. Tucson, AZ 85721-0069, United States. Email: mcrnuna@arizona.edu

Abstract

Classrooms are complex learning environments, with instruction, climate, and teacher–student interactions playing important roles in students' academic progress. To investigate the learning environments of deaf and hard-of-hearing (DHH) students, we developed a new observational tool called the Quality of the Learning Environment-DHH rating scale (QLE-DHH) and rated 98 teachers of DHH students being educated in a range of classroom environments. The present study sought to (1) determine if the items on the QLE-DHH are good indicators of theoretically meaningful dimensions of classroom quality; (2) determine to what extent these dimensions predicted language and reading outcomes of DHH students; and (3) examine how teachers of DHH students were rated on the indicators of classroom quality. The findings suggested that the QLE-DHH has excellent structural validity. Ratings predicted student reading outcomes. Finally, the QLE-DHH was able to capture teachers' strengths and skills in need of improvement. The QLE-DHH appears to hold promise for use in both research and teacher preparation programs.

Introduction

Classrooms are complex learning environments, with instruction, climate, and teacher–student interactions playing important and complementary roles in students' academic progress. Teachers must be able to deliver instructions in a way that effectively maintains student engagement, challenges and responds to student learning, and encourages students to be independent learners. Teacher behavior within the learning environment has been shown to be a predictor of student outcomes in general education (Connor et al., 2014; Rivkin et al., 2005; Sanders & Rivers, 1996; Wright, 1997), but it is less studied in special education classrooms or in classrooms of deaf and hard-of-hearing (DHH) students. Researchers (Jones & Brownell, 2014; Knoors & Hermans, 2010) suggest that special education teachers, including teachers of DHH students, should use a combination of good general instructional strategies and teaching procedures adapted to the educational needs of their students. The purpose of this study is to examine the reliability and predictive value of a rating scale, the Quality of the Learning Environment-Deaf/Hard of Hearing (QLE-DHH), which was designed to capture aspects of general instructional strategies as they apply to the needs of DHH students. Additionally, this study describes the teaching behaviors of a large sample of teachers of DHH students (TODHH) whose teaching was rated using the scale.

Effective teaching

Effective teaching is an essential component of the classroom learning environment. Good teaching requires a teacher to manage effectively several systems within the learning environment that, when coupled with evidence-based practices (EBPs), can lead to better student outcomes. In their seminal work, Brophy & Good (1986) reviewed the literature on teaching behaviors that best support student achievement. They reported that the amount of time spent in instruction had the greatest effect on student outcomes. Pacing of instruction, opportunities to learn, teacher expectations, and classroom management were also factors that positively influenced student outcomes. These factors were particularly impactful in the early grades (K-3). Students in the early grades learned best when teachers used explicit instruction delivered systematically and clearly and managed the learning environment in such a way that students understood the expectations and had frequent opportunities to participate. Teacher questions, both higher order and lower level (i.e., closed-ended), and teacher reactions to student responses (e.g., acknowledgment of correct responses, rephrasing and explanations of incorrect responses) also had an influence on student learning (Brophy, 1988; Brophy & Good, 1986; Connor et al., 2014).

Because the learning environments of special education students differ from those of typical learners, special education teachers must additionally consider individual student characteristics and the context of the classroom when planning and providing instruction. Research on special education has established that explicit instruction, systematic feedback, metacognitive instruction, and small, interactive groups have the greatest impact on students' reading, math, and writing abilities (Brophy, 1988). Jones & Brownell (2014) define effective special education teachers as "…those who have a deep and sophisticated knowledge of instructional practices and can integrate these in ways that promote student engagement and understanding" (p. 115). Using their knowledge of specific disability areas, special education teachers should be able to provide instruction that is

Received: May 16, 2023. **Revised:** September 5, 2023. **Accepted:** September 14, 2023

explicit, cohesive, intensive, engaging, responsive, and content- and skills-focused.

Researchers in the field of deafness have examined the effect of student demographic variables on academic outcomes of DHH students, such as hearing level (e.g., Antia et al., 2009), cochlear implants (e.g., Geers & Hayes, 2011), early identification and intervention (e.g., Moeller, 2000), and educational placement (e.g., Stinson & Kluwin, 2011). However, there has been little systematic investigation of the quality of teacher instruction and teaching behaviors in classrooms with DHH students and the effect of teaching behaviors on language, literacy, or other academic outcomes. Knoors & Hermans (2010) suggest that the quality of classroom instruction might explain a large amount of unexplained variance in the academic achievement of DHH students and therefore should be carefully examined. Knoors (2007) suggests that teaching behaviors that are effective for typically hearing (TH) or special education students should also be effective for DHH students. At the same time, Knoors & Hermans (2010) acknowledge that there are no data as to whether, and to what extent, application and adaptation of effective teaching behaviors occur in the learning environments of DHH students. TODHH need to consider a broad spectrum of student needs (e.g., mode of communication, primary language, hearing assistive technology).

In short, the field needs to obtain data on classroom teaching behaviors and examine whether and how these behaviors connect to student learning. To this end, we examined the ratings of the teaching behaviors of a large number of TODHH using the QLE-DHH in three ways. First, we examined if the teacher behaviors included in the scale were good indicators of theoretically meaningful dimensions of classroom quality. Second, we determined whether information about the teaching and learning environment, captured by the QLE-DHH, was related to progress in language and literacy learning of young DHH children. Finally, we report how teachers rate on the QLE-DHH indicators while teaching language arts.

Impactful teaching behaviors

Based on a review of the literature, five dimensions of teaching behaviors emerged as having the most impact on student outcomes: instructional delivery, instructional clarity, organization and planning, behavior management, and warmth and responsiveness. Although we describe these as though they are "stand-alone" dimensions, there is considerable overlap among them, and the cohesive, integrative combination of these dimensions is what contributes to high-quality learning environments. In the following sections, we define each dimension, describe briefly selected research that links the dimension with outcomes of TH students, and describe the research (if it exists) on DHH students.

Instructional delivery

Instructional delivery can be defined as "...the myriad teacher responsibilities that provide the connection between the curriculum and the student" (Stronge et al., 2011, p. 340). Effective instructional delivery is purposeful and focused on student outcomes. It requires that teachers spend a majority of classroom time engaged in meaningful instruction (Brophy, 1988; Taylor et al., 1999). To this end, teachers must apply a variety of strategies to teach, communicate, and interact with students around academic content and to support student engagement. Teachers who demonstrate effective instructional delivery individualize instruction by teaching in small groups (Zahorik et al., 2003), ask purposeful questions, provide appropriate wait time to give all students an opportunity to respond (Brophy, 1988), and systematically deliver instructional content with each step in the instructional sequence supporting the lesson objective (Brophy & Good, 1986).

Effective instructional delivery leads to better student outcomes. Connor et al. (2014) found that at-risk TH students showed greater gains in vocabulary and reading comprehension when their teachers spent more time in instruction. In an observational study of 44 third-grade teachers, Carlisle et al. (2011) reported that teacher-directed instruction that included modeling, asking questions, and providing practice and review activities significantly improved reading comprehension outcomes for TH students. In a study of the quality of instruction received by DHH students in various settings, Knoors & Hermans (2010) found that teachers at a school for the deaf provided fewer instructional activities than those provided by general education teachers in a coenrollment program. However, the authors did not explicitly link differences in instruction to differences in DHH students' learning outcomes.

Instructional clarity

Instructional clarity is defined as the teacher's ability to explain clearly the content and purpose of the lesson (Civikly, 1992; Stronge, 2018). While researchers have not always agreed on teaching behaviors that constitute instructional clarity, it typically includes behaviors such as describing instructional objectives, intentionally explaining instructional content, and responding to students in ways that maximize learning. In their observational study, Carlisle et al. (2011) found that teacher-directed instruction (e.g., presenting content to support learning) and support for student learning (e.g., promoting active student involvement) improved reading comprehension for students in third-grade classrooms. In classrooms with DHH students, the dimension of instructional clarity should also include student access to communication and comprehensibility of instruction. Thus, the teacher's fluency in the language being used (e.g., American Sign Language [ASL] or English) and mode of communication (spoken or signed) may be a component of clarity along with physical environments that promote visual and auditory access to communication. Unfortunately, there is no current research that links DHH students' academic outcomes or engagement with instructional clarity.

Organization and planning

Organization and planning include the systems, strategies, and methods that a teacher implements to manage the daily operation of the learning environment. Learning environments are well organized when advanced planning and routines are evident. Effective teaching and learning cannot occur in poorly managed classrooms (Korpershoek et al., 2016; Marzano et al., 2003; Oliver et al., 2011). High-quality classroom learning environments include procedures and systems for organizing transitions and materials to ensure that students can spend the maximum amount of time in meaningful instruction. In a 3-year longitudinal study of 26 first-grade teachers and 108 TH students, Cameron et al. (2008) found that the amount of time teachers spent on classroom organization was associated with gains in word reading over the school year. Interestingly, the authors found that the classrooms where teachers spent more time supporting classroom organization in the fall and less time in the spring were the most effective in increasing student reading outcomes,

suggesting that once students understand the organizational systems in the classroom, teachers can spend less time on classroom organization and more time in instructional activities. Stronge et al. (2011) examined the practices of 32 fifth-grade teachers, comparing those whose students achieved in the top and bottom quartile in math and reading. They found that teachers whose students scored in the top quartile were rated significantly higher on observations of classroom management and organization than those whose students scored in the bottom quartile. Guardino & Antia (2012) found that DHH students' disruptive behaviors decreased while engagement increased when their teachers organized the learning environments by decreasing visual distractions and providing appropriately labeled storage space. Catalano et al. (2022) found that when three TODHH working in elementary self-contained classrooms were coached on how to develop instructional arrangements that support individualized instruction, their implementation of classroom rules, procedures, and organizational systems fostered learners' independence and increased the students' active engagement.

Behavior management

A consistent and effective behavior management plan contributes to a high-quality learning environment and is evidenced by the proactive measures a teacher takes to ensure students are ready to learn. Specifically, in classrooms with effective behavior management, the teacher is in charge; discipline is consistently proactive, positive, and corrective; and classroom disruptions are handled quickly and effectively. In a meta-analysis of more than 100 studies from 1967 to 1996, Marzano et al. (2003) reported that effective classroom disciplinary interventions resulted in a 32 percentile decrease in student behavioral disruptions. Day et al. (2015) reported that in high-quality learning environments in which teachers implemented classroom and behavior management procedures that limited unproductive non-instruction time (e.g., off-task time, waiting, disruptions), first-grade students demonstrated improvement in literacy skills from the beginning to the end of the year.

Although there is little research on the behavior management strategies in classrooms of DHH children, the research that is available indicates that TODHH may not be fully prepared to handle many of the behaviors presented in their classrooms. For example, according to Garberoglio et al. (2012), TODHH with fewer than 5 years of experience felt less effective with their implementation of classroom management strategies than their implementation of instructional strategies. In a single-subject study investigating the use and effectiveness of function-based interventions with DHH students, Gann et al. (2015) found that classroom antecedent conditions for behavior management did not represent best practices for two out of the three student participants. The teachers of these two students did not have established routines and rules, nor did they provide behavior-specific praise. Student disruptions during instruction were common. When teachers implemented behavior intervention plans with fidelity, which included behavior management best practices, student on-task behaviors increased by 50%–60% (Gann et al., 2015).

Warmth and responsiveness

Warmth and responsiveness refer to the affect and demeanor of the teacher that creates a learning environment in which students are comfortable taking risks, participating in classroom activities, and approaching the teacher. When teachers are warm and responsive, students gain confidence in their abilities and become independent learners (Reeve, 2006). Students are less likely to comply with rules and procedures, and will be more resistant to disciplinary actions, if they think their teacher does not care about them (Marzano et al., 2003). Students with teachers who are responsive to their questions and interests as well as supportive of the emotional climate of the learning environment have better academic outcomes. In a large national study of 787 first-grade TH students, Connor et al. (2005) reported that those students whose teachers scored high on warmth and responsiveness (based on systematic classroom observations) made greater gains in reading and vocabulary by the end of first grade. Stronge et al. (2011) compared teachers of students in the top and bottom quartile in math and reading achievement and reported that teachers of students in the top quartile had significantly higher observer ratings on positive relationships with their students. Woolsey et al. (2004) found that TODHH working with middle school students in public, residential, and treatment centers had neutral affects for most of the classroom time as compared to approval or disapproval. However, no research was identified that examined overall teacher responsiveness in classrooms of DHH students.

Classroom observation systems

Several observation systems have been developed and used to describe instruction in general education (Connor et al., 2014; Reddy et al., 2013*a*; Smolkowski & Gunn, 2012) and special education classrooms (Johnson et al., 2020), but none have been created specifically for deaf education (Knoors & Hermans, 2010). While it is necessary that TODHH understand the unique needs of their students, it is equally important that they have the pedagogical knowledge and skills they need to effectively deliver instruction and manage the learning environment (Garberoglio et al., 2012). Effective teachers are able to use a range of instructional methods to meet the individual needs of their students (Jones et al., 2022). While some aspects outside of the teacher's control may impact student learning (e.g., class size, school climate/culture, students' SES), scores on teacher observation systems are an effective way to predict student outcomes (Gill et al., 2016).

Results from assessments that measure student outcomes or progress as a direct measure of teacher behaviors (i.e., value-added testing) indicate that teachers and the learning environments they create matter (Pianta & Hamre, 2009). However, while value-added testing has been used in general education, it may not be suitable for special education populations (Jones & Brownell, 2014) for a variety of reasons including small class sizes and lack of reliability and validity of value-added tests for special education students. Additionally, value-added measures provide limited guidance for improving pre-service teacher education and in-service teacher professional development (Pianta & Hamre, 2009). Value-added assessments "....do not reveal what is going on in the classroom and the characteristics of the environment that explain the variability in teachers' value-added scores" (Connor et al., 2014, p. 763). Alternatively, researchers have linked effective teaching behaviors captured by observation systems to TH student outcomes. Observation systems used within general education classrooms hold promise for use in special education classrooms if they are adapted to match the needs of a given special education population, which may eliminate some issues associated with value-added measures, such as reliably

measuring growth in students with disabilities (Jones & Brownell, 2014).

Special education instruments

The Recognizing Effective Special Education Teachers (RESET; Johnson et al., 2020) observation system was created for teachers of students with high-incidence disabilities (i.e., those students with mild emotional/behavioral disorders, learning disabilities, or language impairment). The creators sought to align the instrument with EBPs for this population of students to determine the extent to which teachers are implementing EBPs, provide feedback so teachers can improve their practices, and improve student outcomes. RESET includes three subscales that measure instructional methods, content organization and delivery, and individualization. Each subscale includes EBPs (called "items") that special education teachers should implement (e.g., explicit instruction, reading for meaning, and assistive technology). In a study that included 10 special education teachers from three states, the researchers found that the items within the instructional methods subscale explained a large portion (36%) of the variance for student outcomes.

The Framework for Teaching (FFT; Danielson, 2013) is a teacher evaluation instrument that is used widely in the United States for general education teacher evaluation. Jones et al. (2022) investigated if the FFT adequately captured the range of skills necessary to effectively teach students receiving special education services. Their findings suggest that the FFT is not sensitive to effective special education practices, particularly the need for teacher-directed (i.e., explicit instruction) rather than student-centered instruction.

General education instruments

The Classroom Observation of Student–Teacher Interactions (COSTI; Smolkowski & Gunn, 2012) documents four student–teacher interactions considered important during beginning reading instruction: teacher demonstrations, independent student practice, student errors, and teacher corrective feedback. The researchers found that the variance in kindergarteners' word reading was predicted by opportunities for independent practice and teacher corrections of student errors followed by opportunities for additional practice. Opportunities for independent practice had the most significant impact on overall literacy measures except for receptive vocabulary.

The Classroom Strategies Scale-Observer Form (CSS; Reddy et al., 2013*a*) is a "...multidimensional formative assessment designed to evaluate teachers' use of evidence-based instructional and behavioral management practices..." (p. 311). The two-part classroom observation scale counts frequency of use of specific instructional strategies and rates teacher use of instructional and behavior management strategies. Instructional strategies captured in the CSS include providing concept summaries, creating opportunities for students to respond, providing specific instructional or behavioral requests, giving vague commands, and providing verbal and nonverbal praise for learning. When rating instructional strategies, observers record how teachers use teacher- or student-directed instruction, engage students in learning, deliver clear instruction, monitor student understanding, promote student thinking, and provide performance feedback. When rating behavior management strategies, observers record how teachers respond to students' appropriate and inappropriate behaviors, provide verbal and non-verbal praise for behavior, provide corrective feedback, use proactive and preventative management methods, and provide effective directives. In a study that included 23 general education teachers, Reddy et al. (2013*b*) found that teachers who were rated higher on the instructional strategies scale had students who achieved better math and English language arts results on state-level standardized assessments. They did not find a similar connection between ratings of teachers' behavior management and student academic outcomes.

The Quality of the Classroom Learning Environment Rating Scale (Q-CLE; Connor et al., 2014) was developed using a dynamic systems framework, which supports the assumption that different systems within a classroom influence student learning, including both the manner in which the teacher delivers instruction and the general classroom climate. The Q-CLE measures the quality of teacher–student interactions, the developmental appropriateness of instruction, and the general structure of the classroom learning environment (McLean et al., 2016). Several studies have demonstrated that the factors embedded in the Q-CLE predicted literacy in TH students. In an early version of the Q-CLE, Connor et al. (2014) found that TH students who spent their time in high-quality classroom learning environments and received greater amounts of time in meaningful teacher-managed instruction made the greatest gains in vocabulary and reading comprehension. Likewise, using the Q-CLE, McLean et al. (2016) found that better classroom quality predicted greater gains in passage comprehension and vocabulary in TH students. Because it aligned well with teacher effectiveness research, the Q-CLE was used as the foundation for the development of the QLE-DHH.

In summary, achievement outcomes of TH children are related to the overall quality of classroom teaching and learning environments (Connor et al., 2014; Day et al., 2015). Classroom observation systems have been designed to capture the qualities of effective teaching and learning in classrooms. High-quality learning environments can be characterized by five dimensions of effective teaching, namely, teachers' instructional delivery, instructional clarity, organization and planning, behavior management, and warmth and responsiveness. Unfortunately, we know very little about instruction in classrooms of DHH students and even less about how classroom environments influence student outcomes. This research is a first attempt to examine the classroom learning environments of DHH children on a relatively large scale and, specifically, to examine the relationship between the classroom environments that include DHH students and learning outcomes.

The present study

The present study is one component of a larger longitudinal study examining the language and reading progress of DHH students in kindergarten through second grade. For the larger study, trained assessors completed language and reading assessments of 336 DHH students at the beginning and end of the school year (Lederberg et al., 2019). The students attended 103 classrooms in nine states within the United States and one province in Canada. We videorecorded teaching instruction in these classrooms three times a year (fall, winter, and spring). In our initial studies, we examined students' language and reading based on their auditory access to spoken language and/or access to sign language during classroom instruction. In this study, we examined the extent to which the quality of classroom instruction, as determined by QLE-DHH ratings, influences language and reading outcomes of young DHH students.

Our previous analyses of the language and reading data indicated that auditory access (measured through a test of speech perception) did not affect students' expressive vocabulary scores when children were allowed to use either spoken and/or signed words. Specifically, students whose speech perception scores indicated they were able to access language through audition performed similarly on an expressive vocabulary test to those who accessed language only through vision and who used ASL only. However, students who were able to access spoken language auditorily performed better on single word reading and reading comprehension than students who had no speech perception and used ASL (Antia et al., 2020; Lederberg et al., 2019). Therefore, we decided to control for speech perception when examining the influence of the learning environment on language and reading outcomes. The primary research questions are as follows:

1) What is the content validity of the QLE-DHH? That is, how well do the items of the QLE-DHH indicate the five intended dimensions of effective teaching?
2) What is the predictive validity of the QLE-DHH? That is, to what extent do the QLE-DHH dimensions relate to language and reading outcomes of DHH students in grades K-2?
3) How are teachers of DHH students rated on various teaching behaviors on the QLE-DHH? What patterns of strength or weakness are revealed from these ratings?

Methods

Participants

Classes

There were 103 instructional groups in the study. Some teachers taught two groups of students, and each combination of students and teacher were treated as separate instructional groups. Seventy-eight groups (76%) were in self-contained classrooms with only DHH students located either in local elementary programs or in schools serving only DHH students. Class size ranged from 2 to 11. Fifteen additional groups were general education classrooms where DHH students were integrated with hearing peers. These classes had from 18 to 29 students, with one to four DHH students in a class. Teachers used a combination of sign and spoken English in 43 instructional groups; spoken English alone was used in 32 groups; and sign alone was used in 24 groups. There were 25 kindergarten groups, 17 first-grade groups, 15 second-grade classes, and 39 multi-grade classes.

Teachers

Ninety-eight individual teachers were included in the sample. Two observation times each year were included in this study. Eighty-eight teachers were observed in time 1 (T1) fall and 76 in time 2 (T2) spring (70 of the teachers were observed at both times). There were fewer instructional sessions in the spring than in the fall because some teachers withdrew from the study, and other classrooms could not be recorded. The teachers in this sample had, on average, 11 years of experience teaching DHH students ($SD = 8.9$). Eighty-one percent of these teachers held a Master's degree. Eighty-nine percent of the teachers in this sample were White, and 80% were hearing.

Students

Three-hundred twelve DHH students participated in this study; 24 students who were in the original studies (Lederberg et al., 2019) were not included in this study because they were not videorecorded. Eligibility criteria included being DHH (Better Ear-Pure Tone Average, BE-PTA > 25 dB); enrolled in kindergarten, first, or second grade; and between the ages of 5 and 9 years. Students with mild disabilities (identified from teacher reports) were included, but those with severe disabilities (defined as the presence of autism or severe visual or cognitive impairment) were not. A small number of students were in schools or classes that did not allow video recording and thus were not included in the present study. Table 1 lists the student demographic data. We received permission from the university IRBs and from many schools to notify parents rather than require individual written consent. Therefore, we were able to include all students who met our eligibility criteria in most schools.

Table 1. Demographic characteristics of students.

Variable	Count
Gender	
Female	159
Male	141
Grade	
Kindergarten	107
First	108
Second	85
Race/ethnicity	
Black	48
White	125
Latino	90
Asian/Pacific	22
Indigenous	3
Bi/Multiracial	11
Not reported	13
Deaf or hard-of-hearing parent	88
Audiological technology	
Unilateral CI (with or without HA)	57
Bilateral CI	47
Hearing aid(s) only	132
None	38
Early speech perception	
No pattern perception	115
Pattern perception	2
Some word identification	7
Consistent word identification	169

Note. Total $n = 312$. Age in years mean = 6.7, $SD = 1.0$. Not all categories sum to 312, due to a lack of response.

Measures

Vocabulary

We assessed vocabulary with the Expressive One-Word Picture Vocabulary Test-4 (EOWPVT; Martin & Brownell, 2011), which required students to name pictures of increasingly unfamiliar items. We accepted both spoken and signed words (ASL or English). Responses were judged to be correct based on the manual and a list of acceptable signs developed by the researchers (see Antia et al., 2020 for more information). Internal consistency reliability of the EOWPVT for students in the study was .97 for both fall and spring.

Reading

The Woodcock–Johnson Tests of Achievement-III Normative Update (WJ-III; Woodcock et al., 2007) was used to measure students' word reading and reading comprehension. WJ-III Letter-Word Identification (*Letter-Word Id*) required students to identify letters and single words. On the WJ-III Passage Comprehension (*Passage Comp*), initial items required a student to match a rebus

with a picture, the next set of items required a student to match short written phrases to the appropriate picture among three pictures, and the final set required a student to provide a missing word in written sentences and paragraphs (i.e., cloze technique).

For both subtests, standard ceiling and basal rules were used. Spoken and/or signed words were acceptable. The WJ-III Normative Update converts students' total scores to W scores. W scores are scaled total scores that place students on an equal interval metric across levels useful for measuring growth using a common unit. For purposes of this study, to limit the number of analyses, we added the W scores for the two subscales to form one overall score to reflect reading ability. Internal consistency reliability coefficients for students in this study were .95 (both fall and spring) for Letter-Word ID and .88 (fall) and .89 (spring) for Passage Comprehension.

Speech perception

We assessed speech perception with the Early Speech Perception Test (ESP; Moog et al., 1990). Using an acoustic hoop, examiners asked students to select referents from a closed set of pictures for spoken words. Consistent with the test manual, performance was rated on the following ordinal scale: 1 = no pattern perception (*0 correct*), 2 = pattern perception (*discriminated syllabic structure*), 3 = some word identification (*selected correct referents for 33%–65% of words*), and 4 = consistent word identification (*selected correct referents for more than 65% of words*). While the ESP has four categories, more than 95% of children scored either a 1 or 4 (see Table 1 for full distribution). Therefore, children were divided into two categories: having functional hearing (scored 3 or 4, $n = 176$) or not having functional hearing (1 or 2; $n = 117$).

Procedures

Examiners administered the tests individually to students in a quiet room at their schools in the fall and spring of the school year (see Antia et al., 2020 for a fuller description of procedures).

Trained observers videorecorded language arts instructional sessions in each classroom three times during the school year (fall, winter, and spring) using two video cameras, one focused on the teacher and one on the students. We defined language arts instruction to the teachers as instruction focused on reading, writing, and language. Language arts sessions included, but were not limited to, instruction focused on vocabulary, phonics, reading connected text, and spelling. Based on this definition, teachers provided observers the start and stop time of their language arts instructional session.

Observation coding

Classroom observations were coded using an observation rating scale called the QLE-DHH. The QLE-DHH is designed to measure the quality of classroom instruction by assessing the behavior of the teacher for 25 indicators that were organized around five dimensions. After observing the entire instructional session, observers rated each indicator based on detailed descriptions of behaviors with ratings from 1 (*almost never observed*) to 6 (*exemplary*). Table 2 displays the 25 indicators organized by their associated dimension, with a description of the behaviors that would result in an exemplary rating (6). The complete QLE-DHH rating scale can be found in Appendix A (available online).

Adaptation process

The QLE-DHH is an adaptation of the Q-CLE (Connor et al., 2014). We selected the Q-CLE because it aligned well with teacher effectiveness research and captured important dimensions of classroom instruction. The original Q-CLE contained four dimensions of classroom quality. In studies that used the Q-CLE, raters assigned one rating on a six-point scale for each dimension based on teacher and student behavior. We adapted this scale in several ways. First, we separated and operationalized discrete teaching behaviors (e.g., questioning, management system, and teacher talk) that were collapsed within the Q-CLE. Each teaching behavior, rather than the dimension, received a rating. For example, teachers rated using the Q-CLE dimension of classroom orientation, organization, and planning were given one global rating based on classroom organization, instructional clarity, and the efficiency of transitions. Because one use of observation systems is teacher improvement (Jones et al., 2022), the QLE-DHH authors separated and operationalized these elements into observable teaching behaviors. For example, the classroom orientation, organization, and planning dimension of the QLE-DHH includes four indicators that were separately rated (see Table 2). Second, we expanded the number of behavioral indicators that were considered part of classroom quality based on the research literature and our initial observations of the classroom video recordings. Some of the indicators were specific to DHH students (e.g., visual and auditory access), while others could apply to any student but were not in the original scale (e.g., wait time). Third, we operationally defined some indicators in ways that were specific to the needs of DHH students. For example, the definition of the *explanation of instructional content* indicator within the instruction dimension includes "uses conceptually accurate and correct signs, adjusts mode of communication, and breaks down content into smaller steps." From this process, a detailed coding manual (available from the first author) was created that provided observers with directions for setting up videos and taking notes while observing classrooms, detailed and operationalized descriptions of each indicator as applied to DHH students who use spoken and/or sign language, and a description for how to assign a rating.

Because there is no reason to think that effective PK-12th grade teaching practices identified in the literature for TH students would be any different from effective teaching practices for PK-12th grade DHH students, the QLE-DHH was developed using general best teaching practices as the foundation. The QLE-DHH does not describe or capture strategies or interventions that have been developed specifically for DHH students, nor does it capture cultural aspects that may be unique to the deaf community and classrooms at schools for the deaf. The QLE-DHH is intended to apply to any learning environment that includes a DHH student. The exclusion of these elements should not be taken to mean that the researchers do not recognize their importance; it simply means that these aspects were not included in the original iteration of the QLE-DHH.

Rating

The first two authors served as the primary raters for this study. Each has over 20 years of experience as TODHH in a variety of learning environments, including schools for the deaf and itinerant teaching. Using the QLE-DHH, they rated DHH students' primary teacher while watching the entire videorecorded language arts instructional session. The average instructional session lasted just under 2 hr ($M = 116$ min; $SD = 47$) but ranged from 28 to 265 min. The primary teacher was the teacher with whom the student spent the majority of language arts instruction; for the majority of students, this was a TODHH, but for some, it was a general education teacher. By selecting to code the primary teacher, we were able to capture the manner in which the DHH

Table 2. A description of exemplary ratings of each indicator of the five dimensions of the QLE-DHH.

Dimension and indicators (items)	Exemplary rating (6)
Instructional delivery	
Time in instruction	Teacher devotes entire period to teaching (versus non-instruction).
Checks for understanding	Teacher checks with students to determine their understanding of lesson objectives using a variety of strategies.
Checks inform instruction	Checks for understanding are used to inform instruction .
Individualized instruction	Teacher provides individualized instruction to reinforce learning.
Systematic instruction	Instruction is systematic and consistent with lesson objectives.
Pacing	Pacing of instruction is appropriate.
Questioning	Teacher's questions promote students' learning.
Wait time	Wait time is long enough to allow students to think.
Student participation	Teacher gives all students opportunities to participate.
Instructional clarity	
Objectives explicitly identified	Objectives are explicitly stated at the beginning of a lesson.
Explanation of instructional content	Clear explanation of instructional content according to the objective.
Maximizes learning opportunities	Teacher responds to students in ways that maximize learning.
Visual and auditory access	Teacher ensures that all students have visual and auditory access.
Classroom orientation, organization, and planning	
Organizational system	Classroom runs like a well-oiled machine.
Student disruptions	Activities run smoothly; disruptions are rare.
Materials	Materials for instruction are organized and readily available.
Communicates expectations	Expectations for transitions are clear. Transitions are quick.
Behavior management and control/discipline	
Teacher is in charge	Teacher is clearly in charge.
Management system	Teacher communicates behavioral expectations clearly.
Approach to discipline	Discipline is consistently proactive, positive, and corrective.
Outside disruptions	Nonstudent interruptions are handled quickly.
Warmth and responsiveness	
Teacher talk	Teacher models an encouraging, positive, supportive, and respectful attitude toward others.
Teacher affect	Teacher affect is consistently positive (not neutral or negative).
Student respect	Teacher encourages respect among students.
Positive and specific attention	Teacher provides specific praise and attention to reinforce students.

Note. Each item was rated on a 1–6 scale.

students typically received instruction in the learning environment in which they spent the most time. While instruction was videorecorded three times during the year, only the fall and spring instructional sessions were coded because rating the 103 instructional sessions was time-consuming. When arranging the time for video recording instruction, the primary teacher informed the videographers when they would be teaching language arts. Language arts sessions included, but were not limited to, vocabulary instruction, read alouds, phonics instruction, and spelling lessons.

Ratings for some indicators were adjusted according to educational setting. For example, ID 9 Student Participation was considered differently for a general education classroom versus a self-contained classroom. In a self-contained classroom with fewer students, it was expected that the teacher gave each student an opportunity to participate. However, in a general education classroom with as many as 20 students, it is not possible for the teacher to give each student an opportunity to participate within one lesson. In this example, the general education teacher was not be rated lower for student participation, rather, they were rated according to the variety of students who were called on to participate.

Instructional sessions were randomly assigned to each observer. Videorecorded instructional sessions were viewed using Mangold Interact software. Using this software, observers were able to view instructional sessions from the viewpoint of both the teacher and the DHH students, which allowed them to capture teacher instruction and students' responses. Observers took notes while observing the instructional session capturing the quality of implementation of the teaching behavior, the amount of time spent demonstrating the behavior, and/or the frequency with which the behavior was observed for the given indicator. Observers then selected the rating that best described the teacher's proficiency level for the teaching behavior based on the extensive descriptors in the coding manual.

Interobserver reliability

Percent agreement was used to calculate inter-observer reliability. Observer ratings were determined to be in agreement if they were within one point of each other. To adapt the scale and develop agreement, the two coders observed and rated 11 sessions together. After they obtained 80% reliability on all 25 indicators, they then independently coded the remaining instructional sessions. Because the raters were able to obtain acceptable levels of reliability for an observation instrument (>70%) within a relatively short amount of time, the decision was made to continue rating the remaining instructional sessions using the QLE-DHH as opposed to an interval frequency system (Erlich & Shavelson, 1978). To maintain interobserver reliability, they independently coded some of the same randomly selected instructional sessions each week. Overall, 33% of instructional sessions were coded by both observers and reliability for the indicators was calculated; average reliability was 97% (range = 85%–100%).

Although the first two authors are fluent in ASL, it was important that the instructional sessions during which ASL was the

Table 3. Model fit.

Semester	Model	χ^2	*df*	CFI	TLI	RMSEA	Lower	Higher	SRMR
Fall	Five[a]	440.1	265	0.96	0.96	0.09	0.07	0.10	0.09
	Three	438.7	272	0.97	0.96	0.08	0.07	0.10	0.09
	One	606.7	275	0.93	0.92	0.12	0.11	0.13	0.11
Spring	Five	428.2	265	0.95	0.95	0.09	0.07	0.11	0.10
	Three	434.8	272	0.95	0.95	0.09	0.07	0.10	0.10
	One	602.2	275	0.90	0.89	0.13	0.11	0.14	0.12

Note. "Model" refers to the number of factors. CFI = comparative fit index. TLI = Tucker–Lewis index. RMSEA = root mean square error of approximation. lower/higher = 90% confidence interval for RMSEA. SRMR = standardized root mean square residual. The three-factor model was fit to remediate the estimation warning for the five-factor model. See the text for details. [a]Estimation warning indicating a nonpositive definite latent variable covariance matrix, but with no obvious problems in the resulting estimates.

primary mode of instruction be rated by native ASL users. Two deaf research assistants whose primary language is ASL were trained on the QLE-DHH and observed 30% of the ASL instructional sessions (14 out of 47 total ASL sessions). They did not provide original ratings for these instructional sessions, but, rather, determined if they agreed or disagreed with the original ratings. After viewing the instructional sessions, they reached 97% agreement (range 93%–100%) with the original ratings.

Analysis and results

Factor models for the psychometric properties of the QLE-DHH

To determine the relations of indicators to their intended dimensions (factors), we empirically evaluated the items as a confirmatory factor model (Bollen, 1989; Newby et al., 1983). We fit confirmatory factor models for *ordinal* rating items (Moustaki, 2000; Muthén, 1984; Samejima, 1969) in order to appropriately account for the six-category ratings (fall frequencies are reported in Table 8). Initially, we fit a five-factor model corresponding to the original five dimensions of the QLE-DHH. We also evaluated a single-factor model to determine if all indicators were so highly correlated with each other that classroom quality was unidimensional and that all indicators were consistent across all dimensions. Factor models were fit for each semester separately.

Table 3 presents the fit indices for the factor models in each semester. The five-factor model had excellent fit on all indices in both semesters, but had a technical warning in the fall, likely because the estimated correlations among the factors were too high. In the Fall, the correlation between instructional delivery and instructional clarity and between organization and behavioral management were each .98; the same correlations for the spring were .95 and .94, respectively. Because these factors were highly correlated and there was a technical warning for the five factor model (in the fall), we fit a post-hoc three-factor model in which we collapsed the two pairs of highly correlated factors. The resulting three-factor model had excellent fit on all indices in both semesters (Table 3; CFI ≥ .95; RMSEA < .10). The three factors were instruction (eight instructional delivery and four instructional clarity indicators), management (four classroom organization and four behavior management indicators), and warmth (four indicators).

Finally, we tested the fit of a one-factor model. The fit was not very good in either semester (CFI < .94; RMSEA > .11; see Marsh et al., 2004). Correlations among the three factors in the three-factor model were homogenous and moderate in each semester ($r = .67$–$.77$) and are shown in Table 4. These correlations were not high enough to validate the single-factor model, which had inadequate fit.

Table 4. Correlations for the three-factor model by semester.

Semester	Factor	Inst.	Manag.	Warmth
Fall	Instruction	1		
	Management	.78	1	
	Warmth	.77	.67	1
Spring	Instruction	1		
	Management	.72	1	
	Warmth	.76	.67	1

Note. The model was fit separately in each semester.

Beyond global fit, it is important to examine the detailed parameters for the three-factor model. Standardized loadings, reported in the Appendix (fall in Table A1; spring in Table A2), reflect the correlation of each item with its intended factor (i.e., validity). Validity coefficients ranged from .34 to .88 (median .80) for the instruction factor in the fall. The median of the standardized loadings was .86 for the management factor and .91 for the warmth factor. For the spring, the median standardized loadings were good (.73, .81, and .91 for instruction, management, and warmth, respectively). Tables A1 and A2 (online) show the full results for this three-factor model in each semester.

Relations to student outcomes

Because the results for the three-factor model were reasonable in each semester, to determine the extent to which the QLE-DHH dimensions relate to language and reading outcomes, we pooled the two semesters and refit the model, holding all measurement parameters equal across semesters (i.e., loadings and thresholds for indicators were held equal), and computed factor scores for each teacher on each of the three factors of the QLE-DHH. The resulting factor scores were used to predict student outcomes of reading and vocabulary. In particular, we fit a multilevel model of students nested within classrooms, with the spring score predicted by the fall score, each QLE factor, and by child speech perception score. This model was fit separately for reading and for vocabulary. The conceptual form of the model of students (s) in classrooms (c) is

$$\text{Spring}_{sc} = \text{Fall}_c + \text{Fall}_s + \text{ESP}_s + \text{QLE}_c + e_{sc}$$

where Spring_{sc} is the spring score (reading or vocabulary), Fall_c is the classroom mean score, Fall_s is the student fall score (centered within classroom), ESP_s is the student's speech perception score, QLE_c is the teacher's score on that QLE factor averaged between fall and spring, and e_{sc} is the random error. The QLE factor scores were then each entered into a multilevel model for student reading and vocabulary, separately (fit in SAS PROC MIXED;

Table 5. Descriptive statistics for instructional and student measures.

Variable type	Measure	*n*	Mean	*SD*	Min	Max
QLE factors	Instruction	312	−.02	0.56	−1.50	1.14
	Management	312	−.07	0.71	−1.69	1.58
	Warmth and responsiveness	312	−.08	0.71	−2.29	1.30
Student measures	ESP	300	0.60	0.49	0	1
	Reading, fall	303	407.32	30.34	326.50	505.50
	Reading, spring	296	423.07	29.04	355.50	512.50
	Vocabulary, fall	274	56.47	18.74	13.00	109.00
	Vocabulary, spring	295	66.03	18.79	17.00	12.00

Note. ESP = speech perception: 60% of children had functional hearing; 40% did not. The QLE factors were centered. Represent the average of fall and spring scores. Reading scores are W scores averaged across word reading and comprehension. Vocabulary scores are raw scores (total correct).

Littell et al., 2006). In each model, because students were nested within classrooms, classroom averages were computed and student-level deviations from fall scores were also computed, so that each could be used in the prediction (i.e., the effect of being a high- or low-performing student might not be the same as being in a high- or low-performing classroom).

Table 5 shows descriptive statistics for the three QLE factor scores, as well as for student-level outcome scores. The ESP (speech perception) indicates that 60% of the students had functional hearing. Both reading and vocabulary scores show that children improved from fall to spring.

Table 6 shows the results for the three multilevel regressions for the reading composite W-score. The first set of columns in Table 6 shows the regression estimates for reading predicted by the QLE factor of instruction, the second set shows the estimates for management, and the third set for warmth. We interpret in depth the results for instruction. Each of the other regression estimates can be interpreted in a similar manner. The fixed effects are in the top five rows of estimates of Table 6. The intercept represents the model-predicted W-score for an average student in an average classroom, with no functional hearing (a score of 1 or 2 on the ESP), with average rating on the QLE factor for instruction. The coefficient for class is the number of W-units of spring reading for every unit the classroom was above the grand mean in the fall. The coefficient for student is the number of W-units of spring reading for every unit the child was above their classroom mean for the fall. The coefficient for ESP is the number of W-units of spring reading if the student had functional hearing (a score of 3 or 4 on the ESP). Finally, the row demonstrating the effect of the QLE factor shows the number of W-units of spring reading the child would be predicted to gain, for every unit the teacher was above the QLE mean (see Table 6 for *SD* of the QLE factors). The next two rows of Table 6 show the random effects, with variance and *SD* shown for the classroom intercept, as well as the student-level residuals. The bottom row of Table 6 shows the model-based effect size (*g*; Hedges, 2007), which is the number of *SD*s the reading outcome would change for every unit of the QLE factor.

Table 6 shows that classroom- and student-level fall reading performance significantly predicted spring reading (see asterisks in the last column of each set). Student speech perception was also a statistically significant predictor of reading. QLE-instruction was a statistically significant predictor of reading gains while controlling for speech perception, with an effect size of .37. The effect sizes for QLE-management and QLE-warmth, although not significant, were .19 and .16, respectively.

As an interpretive example, we can consider a student with an average fall reading score in a classroom where QLE-instruction is 1 *SD* above the mean (.54; see Table 5). This student's model-predicted spring reading would be 1.98 (.54 × 3.66; see Table 6) W-score units higher than a similar child in a classroom with average QLE-instruction. This instructional advantage is 46% of a classroom *SD* (4.25; Table 6). A student in a classroom with QLE-instruction 2 *SD* above the mean would be predicted to have double this advantage. Thus, the model suggests that for reading, QLE-instruction can make a difference. Table 6 suggests having a good fall score, being in a good classroom, and having high ratings of QLE-instruction all help to improve spring reading scores.

Table 7 shows the results for expressive vocabulary, one model each for the three QLE factors in the same format as Table 6. While classroom and student fall vocabulary scores were significant predictors of spring vocabulary scores, student speech perception was not. None of the QLE factors were statistically significant predictors of vocabulary. The effect sizes for instruction, management, and warmth were .12, .19, and – .03, respectively.

Descriptive results of QLE-DHH ratings

To answer the third research question as to how teachers of DHH students rated on various instructional behaviors on the QLE-DHH and the patterns of strength or weakness those ratings revealed, we analyzed the frequency of each rating in relation to each indicator. Table 8 shows the response frequencies of teachers who received the given rating for each indicator of the QLE-DHH for fall and spring.

Overall, teachers in this sample were rated in the proficient range (i.e., "satisfactory" and "exemplary") for most QLE-DHH indicators. Indicators remained stable from fall to spring. Within the instructional delivery dimension, the majority of teachers were rated proficient for *time in instruction* and *student participation* indicators. the indicators with the most variability were *checks for understanding*, *checks inform instruction*, *individualized instruction*, and *questioning*, which may indicate that teachers preferred whole class instruction over small groups or 1:1 instruction and that they may need to more frequently ask questions to check for student understanding and engagement.

In the instructional clarity category, most teachers were rated exemplary for the *visual and auditory access* indicator, showing that most teachers ensured their DHH students received the visual and/or auditory access to instruction they require. The majority of teachers were rated as "not observed" and "needs improvements" for the *objectives explicitly identified* indicator in fall and spring. For this indicator, raters considered whether or not the teacher directly referred to the objective at some point during instruction, including displaying the objective on the board. Explicitly identifying the lesson objective helps both the teacher and the students remain focused on the purpose and anticipated outcome of the lesson. For most instructional sessions in this sample, the teacher neither stated the instructional objective nor referred to it posted somewhere in the room.

Table 6. Multilevel predictions for student reading, using fall semester data.

		Instruction			Management			Warmth		
		Est.	**SE**		**Est.**	**SE**		**Est.**	**SE**	
Fixed effects	Intercept	420.47	1.18	*	420.31	1.19	*	420.45	1.22	*
	Class	0.84	0.03	*	0.85	0.03	*	0.86	0.03	*
	Student	0.87	0.04	*	0.87	0.04	*	0.87	0.04	*
	ESP	4.57	1.45	*	4.98	1.45	*	4.74	1.48	*
	QLE	3.66	1.38	*	1.88	1.10		1.61	1.09	
		Est.	*SD*		Est.	*SD*		Est.	*SD*	
Random effects	Classroom	18.03	4.25		19.51	4.42		20.49	4.53	
	Residual	97.89	9.89		98.53	9.93		98.19	9.91	
Effect size		0.37			0.19			0.16		

Note. ESP = speech perception. SE = standard error. Effect size is the model-estimated standardized effect, *g* (Hedges, 2007) for that QLE factor. *$p < .05$ (for fixed effects).

Table 7. Multilevel predictions for student vocabulary, using fall semester data.

		Instruction			Management			Warmth		
		Est.	**SE**		**Est.**	**SE**		**Est.**	**SE**	
Fixed effects	Intercept	65.33	1.31	*	65.42	1.29	*	65.11	1.32	*
	Class	0.87	0.06	*	0.86	0.06	*	0.88	0.06	*
	Student	0.78	0.05	*	0.78	0.05	*	0.78	0.05	*
	ESP	−.58	1.57		−.56	1.53		−0.31	1.57	
	QLE	1.18	1.49		1.92	1.15		−0.33	1.16	
		Est.	SD		Est.	SD		Est.	SD	
Random effects	Classroom	19.81	4.45		17.84	4.22		2.53	4.53	
	Residual	99.27	9.96		99.65	9.98		99.05	9.95	
Effect size		0.12			0.19			−.03		

Note. ESP = speech perception. SE = standard error. Effect size is the model-estimated standardized effect, *g* (Hedges, 2007) for that QLE factor. *$p < .05$ (for fixed effects).

Overall, the classroom orientation, organization and planning dimension was strong for the teachers in this sample. The *materials* indicator was highest in this category across fall and spring, which shows that the teachers knew the value of having materials prepared for instruction; very little instructional time was lost as teachers looked for or prepared materials. That said, many teachers in this sample did not adequately communicate expectations, which led to instructional time lost to transitions in between instructional topics. The majority of teachers were rated proficient for all indicators in the behavior management and control/discipline dimension. While teachers inconsistently demonstrated use of an established behavior management system or proactive approach to discipline, they were clearly in charge of their learning environments and managed outside disruptions effectively. The warmth and responsiveness dimension was also rated highly across all indicators, though many teachers neglected to provide specific praise to their students.

Discussion

Given that achievement outcomes of TH students are related to the overall quality of their classroom learning environments (Connor et al., 2014; Day et al., 2015), the purpose of this study was to evaluate the validity of a teacher observation rating scale, the QLE-DHH. We found that the QLE-DHH rating scale has excellent structural validity to capture the quality of teaching behaviors of TODHH in a variety of educational settings. We also found that the QLE-DHH was related to student outcomes.

This is the first study that explored the dimensions of classroom quality using a tool that was adapted and designed specifically for learning environments with DHH students. The large and diverse sample size allowed us to measure classroom quality across a variety of learning environments, including self-contained classrooms at schools for the deaf, DHH resource classrooms, and general education classrooms. These environments included instruction in ASL, spoken English, or a combination of the two. Students in these environments had a diverse range of ability and language levels. The high reliability between raters (first by the hearing raters, then affirmed by deaf native ASL users) suggests that the QLE-DHH can capture classroom quality of the varied learning environments of DHH students.

Validity and reliability of the rating scale

The fit of the models suggests that the designed structure of the QLE-DHH indicators of teaching behavior was valid. However, the original conceptualization that classroom quality is best depicted as five factors may be inaccurate because several factors were so highly correlated they appeared to be measuring a single dimension. The resulting three-factor model, which included instruction (which measured both instructional delivery and instructional clarity), management (which measured both classroom management and organization), and warmth (which measured warmth and responsiveness), was more parsimonious and did not have estimation warnings. This three-factor model suggests the indicators have excellent structural validity to measure these three dimensions of classroom quality. In addition, the standardized loadings suggest the indicators had good validity and reliability

Table 8. Number of teachers rated at each level for QLE-DHH indicators.

	Fall semester						Spring semester					
Factor, item/indicator	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
Instructional delivery												
Time in instruction				2	11	**75**				3	10	**63**
Checks for understanding	4	7	13	16	**28**	20	9	7	6	16	**21**	17
Checks inform instruction	3	3	6	18	19	**39**	8	5		11	22	**30**
Individualized instruction	3	11	13	9	22	**30**	1	10	7	9	19	**30**
Systematic delivery		2	8	22	**36**	20		4	8	17	**30**	17
Pacing		5	13	15	26	**29**		4	10	12	**27**	23
Questioning	2	5	6	10	27	**38**	2	4	11	5	26	**28**
Wait time			4	14	**36**	34	1	1	3	12	19	**40**
Student participation			2	5	14	**67**			2	7	13	**54**
Instructional clarity												
Objectives explicitly stated	**26**	7	10	24	19	2	**20**	9	10	21	16	
Explanation of content	2	5	16	14	**28**	23	5	4	9	15	**24**	19
Maximizes learning	1	4	13	15	24	**31**	2	8	2	17	20	**27**
Visual and auditory access		1	3	11	24	**49**		2	4	11	24	**35**
Classroom organization												
Organizational system	1	2	4	13	**39**	29	1	1	8	7	**40**	19
Student disruptions	1	1	6	17	**33**	30		3	5	10	**35**	23
Materials			1	8	23	**56**			1	5	27	**43**
Communicates expectations	1	3	9	9	31	**35**		5	7	11	**29**	24
Behavioral management												
Teacher in charge			2	7	26	**53**			2	3	18	**53**
Management system		2	17	16	**41**	12	1	2	17	14	**35**	7
Approach to discipline		6	15	18	**32**	17		7	14	13	**30**	12
Outside disruptions[a]			4	9	13	**61**			4	10	16	**45**
Warmth & responsiveness												
Teacher talk			1	7	32	**48**		1	8	7	20	**40**
Teacher affect		1	4	17	26	**40**	1	5	4	18	21	**27**
Student respect			7	8	26	**47**		1	4	13	18	**40**
Positive and specific attention	4	7	17	18	14	**28**	2	6	14	17	14	**23**

Note. Each item was rated 1–6 (see the text for descriptive terms). Fall $n = 88$. Spring $n = 76$. The modal response is indicated in bold. [a]One observation was missing.

for these dimensions. Moreover, the structure was highly similar in fall and spring, suggesting that at least the configuration of dimensions is consistent within a school year for the grades examined here. On the other hand, the fit for a one-factor model was poor, suggesting that classroom quality is not unidimensional; while dimensions are moderately correlated, teachers may excel in one dimension over another. For example, a teacher may be strong in the warmth dimension but weaker in the instruction dimension.

Given the nature of this observational rating scale, there was potential for high variability across raters. However, the raters in this study had high reliability, which bodes well for the practical use of this rating scale. Even with minimal training, the operational descriptions of the teaching behavior indicators appear to be detailed enough to support high reliability as indicated by the agreement reached by the two additional deaf raters for the ASL classrooms. While the high reliability could be explained by the fact that the first two authors were the creators of the scale and primary raters for the study, the language used within the operational descriptions of the indicators appear to support the ability to observe these teaching behaviors and determine the quality with which each is implemented.

Quality of teaching behaviors and student outcomes

Previous researchers have used rating scales to measure teaching behaviors and have found a similar predictive value for word reading (Reddy et al., 2013*b*; Smolkowski & Gunn, 2012), vocabulary (Connor et al., 2014; McLean et al., 2016; Smolkowski & Gunn, 2012), and reading comprehension (Connor et al., 2014; McLean et al., 2016; Reddy et al. 2013*b*) in TH students. Results from this study indicate that the teaching behaviors of TODHH have a similar impact on reading outcomes for DHH students. The predictive models suggest that the dimension of instruction is a statistically significant predictor of DHH students' reading performance. The loadings for the instruction dimension for fall (see supplemental Table A1) show that the *time in instruction* (i.e., how much time teachers spend in instruction), *checks for understanding* (i.e., ensuring students understand the content of instruction), and *systematic instructional delivery* (i.e., systematically scaffolding a lesson to support student learning) indicators loaded highest. When implemented with quality, these indicators, along with the other indicators within the instruction dimension, work cohesively to ensure that students are engaged and supported throughout an instructional session. While the management and warmth dimensions were not statistically significant predictors for reading outcomes, their effect sizes (.19 and .16) were promising, suggesting that DHH students who had teachers with higher levels of these teaching behaviors, (e.g., an effective organizational system, limited student disruptions, and positive teacher talk) had better reading outcomes (Connor et al., 2005; Day et al., 2015; Stronge et al., 2011).

The three dimensions of classroom quality did not predict vocabulary outcomes, suggesting vocabulary learning may be less sensitive to the quality of teacher behavior than reading. However,

in a previous analysis of a subset of the current sample, Duncan & Lederberg (2018) found that vocabulary gains were related to the frequency that teachers engaged in vocabulary instruction and expanded children's utterances. Therefore, specific instructional strategies may have a greater impact on vocabulary learning in students who are DHH, rather than general teaching behaviors.

The QLE-DHH captured the quality of teaching behaviors of TODHH. Using this scale, we can describe the general teaching strengths of TODHH and identify areas that should be targeted for improvement. In this study, most teachers were rated as proficient (e.g., ratings 4–6) on the majority of indicators; however, there were some indicators that could be targeted for improvement. For example, in both the fall and spring, teachers rarely stated their learning objectives, which is an important driver for systematic instruction. When teachers identify the instructional objective in a way that students can understand, they are more likely to remember the purpose of the instruction they are providing (Brophy & Good, 1986). Providing individualized instruction was another indicator that many teachers in this sample rated as not proficient. Given that there is high variability of student ability levels within many classrooms with DHH students, teachers must be skilled at providing their students with individualized instruction, rather than teach to one ability level.

The QLE-DHH indicators and their operationalized descriptions were created according to the needs of DHH students (e.g., visual and auditory access, conceptually accurate signs). While some QLE-DHH indicators loaded higher on their corresponding dimension (i.e., instruction, management, or warmth), it is important for teachers to recognize and develop the discrete teaching behaviors that may have greater impact for students who are DHH. In an effective learning environment, a teacher may demonstrate multiple QLE-DHH teaching behaviors within the same teaching moment (Stronge et al., 2011); however, for novice teachers, it is necessary to deconstruct the elements of quality instruction for the purpose of teacher preparation and assessing and improving performance. Raters using the original Q-CLE rating scale arrived at a single rating for each dimension (Connor et al., 2014); however, the QLE-DHH is more effective in this regard because it breaks down each dimension into multiple descriptive indicators that are individually rated and, therefore, can be used to target areas for improvement. As such, the QLE-DHH can be used to identify those teaching behaviors that should be targeted for improvement through professional development or individualized teacher coaching, with the goal of improved outcomes for DHH students.

The findings of this study showed implementation of instruction that aligns with the QLE-DHH indicators, particularly indicators within the instruction dimension, has a positive impact on DHH students' reading achievement. In the field of deaf education, researchers have long-searched for EBPs that can better support academic outcomes, particularly reading, for DHH learners (Marschark & Knoors, 2018). Given that many of the teaching behaviors within the QLE-DHH align with the Council for Exceptional Children's High-Leverage Practices (McLeskey et al., 2017), and the relationship between these teaching behaviors and student outcomes, suggests that it is a combination of high-quality teaching practices paired with EBPs that will have the greatest impact on student learning.

Implications for teacher preparation

Although not the purpose of this study, we have suggested that the QLE-DHH can be used to target specific teaching behaviors for improvement. Since the completion of the current study, the first author trained four cohorts of graduate students using the QLE-DHH based on the three-factor model in their practicum and internship experiences (Rivera et al., 2020). The teacher candidates showed improvement in their ability to write a well-crafted lesson, manage DHH students during instruction, and adjust instruction to better support DHH students' learning. In addition, the preservice teachers were able to reflect on their own teaching behaviors to identify strengths and areas to target for improvement. Further research is needed to confirm that the QLE-DHH can be used to improve the quality of teacher behaviors and explore the best ways to implement it. Instructional coaching can be used to support teacher candidates and novice teachers as they work to improve their own teaching behaviors (Kretlow & Bartholomew, 2010). Coaching supports, such as bug-in-ear technology, have been used to provide real-time support to teacher candidates' accuracy of discrete trial teaching skills (McKinney & Vasquez, 2014). While this technology will need to be adapted for TODHH who are deaf themselves, it is a promising way to support development of teaching behaviors such as those included in the QLE-DHH.

Limitations and future research

Though the findings of this study are promising for supporting teachers' instructional practices, there were several limitations. When considering the ratings for this sample of teachers, it is important to note that the QLE-DHH only offers a "snapshot" of teaching behaviors on the day of observation, which may or may not be indicative of their teaching in general. Several aspects of data collection, such as having a camera and observer in the room, could have affected teacher and student behaviors on the day of recording. Additional research on another large sample of TODHH is needed to determine if these ratings are consistent across several cohorts of teachers. In addition, it is possible that teachers in different schools adopt consistent approaches to instruction. The current analysis ignores differences across the 40 schools, as well as differences by country or region. Additional research is needed to parse out these differences, particularly those that include culturally responsive practices for DHH students, especially those who attend schools for the deaf.

Some technical/statistical limitations are present in this study. The current models are only preliminary, in that we examined two child outcomes, with each QLE factor separately. Because we know the instructional factors to be highly and homogeneously related, we cannot argue for sharp distinctions between the models here. The models and their results are certainly related, but we present these initial relations to show that instructional quality matters and that the QLE-DHH may be an important and useful way to begin to measure instructional quality for DHH students. Because the three QLE-DHH factor scores were correlated ($r = .67–.76$), we presented separate multilevel regressions as preliminary exploration for the predictive value of the QLE-DHH factors for student outcomes. Testing longitudinal categorical data is complex. In the current study, we used a pragmatic approach of testing the two time points, forcing equivalent scoring, and then subsequently evaluating those equated scores in a second-stage analysis. A larger sample might allow a single, joint analysis of longitudinal invariance, along with student outcomes (multilevel, nested within classrooms). Testing each instructional factor separately upon each of the two student outcomes is not elegant and might have invited Type-I error (we therefore have no basis to make claims of differential effects). However, this is an initial examination of the construct and predictive validity of the QLE-DHH dimensions insofar as there are positive relations with student

gains. The current analysis is suggestive as a preliminary study of these initial items in this understudied population.

Finally, interrater agreement of a subjective rating scale can be problematic. The first and second authors developed the QLE-DHH and then served as the raters in this study. While they were able to achieve high reliability, more research is needed, on a different sample of teachers/learning environments, to determine if others can reliably use the scale.

Conclusions

Continual evaluation and reflection of teaching behaviors is an integral part of being an effective teacher. The QLE-DHH provides preservice and inservice teachers, and their supervisors, with a common language for discussing instructional practices. The systematic use of the QLE-DHH within deaf education teacher preparation programs and as part of inservice professional development through self-reflection and teacher evaluations may be one way to ensure TODHH are prepared to monitor their own teaching to effectively support the language and literacy needs of DHH students. Implementing evidence-based instructional interventions is important, but, without the inclusion of effective teaching behaviors, these interventions have the potential to lose their effectiveness. Using the QLE-DHH to support the implementation of effective teaching behaviors of TODHH may ultimately contribute to the time DHH students spend in meaningful instruction and yield more positive learning outcomes.

Developing a way to capture classroom quality is challenging. This study presents preliminary data that indicates the QLE-DHH is a valid and reliable way to capture the instructional experiences of DHH students. The instrument holds promise for improving TODHH instructional practices through professional development for preservice and in-service teachers and further our understanding of learning environments with DHH students.

Supplementary Materials and Data

Supplementary material is available at *Journal of Deaf Studies and Deaf Education*.

Funding

This research was supported by the Institute of Education Sciences, U.S. Department of Education through Award #R324C120001. The content of this article does not represent views of the Institute or the U.S. Department of Education.

Conflict of interest: The authors have no conflicts of interest fo disclose.

Data availability

The identified data can be obtained from the corresponding author.

References

Antia, S. A., Jones, P. B., Reed, S., & Kreimeyer, K. (2009). Academic status and progress of deaf and hard-of-hearing students in general education classrooms. *Journal of Deaf Studies and Deaf Education*, **14**(3), 293–311. https://doi.org/10.1093/deafed/enp009

Antia, S. D., Lederberg, A. R., Easterbrooks, S., Schick, B., Branum-Martin, L., Connor, C. M., & Webb, M. (2020). Language and reading progress of young deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, **25**(3), 334–350. https://doi.org/10.1093/deafed/enz050

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons, Inc.

Brophy, J. (1988). Research linking teacher behavior to student achievement: potential implications for instruction of chapter 1 students. *Educational Psychologist*, **23**(3), 235–286.

Brophy, J. E., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328–375). Macmillan.

Cameron, C. E., Connor, C. M., Morrison, F. J., & Jewkes, A. M. (2008). Effects of classroom organization on letter-word reading in first grade. *Journal of School Psychology*, **46**, 173–192. https://doi.org/10.1016/j.jsp.2007.03.002

Carlisle, J., Kelcey, B., Berebitsky, D., & Phelps, G. (2011). Embracing the complexity of instruction: a study of the effects of teachers' instruction on students' reading comprehension. *Scientific Studies of Reading*, **15**(5), 409–439. https://doi.org/10.1080/10888438.2010.497521

Catalano, J., Weirick, W., Hasko, J., & Antia, S. (2022). Teacher coaching: Increasing deaf students' active engagement through flexible instructional arrangements. *Journal of Deaf Studies and Deaf Education*, **27**(1), 73–88. https://doi.org/10.1093/deafed/enab031

Civikly, J. M. (1992). Clarity: Teachers and students making sense of instruction. *Communication Education*, **41**(2), 138–152.

Connor, C. M., Son, S., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology*, **43**(4), 343–375. https://doi.org/10.1016/j.jsp.2005.06.001

Connor, C. M., Spencer, M., Day, S. I., Giuliani, S., Ingebrand, S. W., McLean, L., & Morrison, F. (2014). Capturing the complexity: content, type, and amount of instruction and quality of the classroom learning environment synergistically predict third graders' vocabulary and reading comprehension outcomes. *Journal of Educational Psychology*, **106**(3), 762–778.

Danielson, C. (2013). *The framework for teaching evaluation instrument* (2013th Ed.). The Danielson Group.

Day, S. L., Connor, C. M., & McLelland, M. M. (2015). Children's behavior regulation and literacy: the impact of the first grade classroom environment. *Journal of School Psychology*, **53**(5), 409–428. https://doi.org/10.1016/j.jsp.2015.07.004

Duncan, M. K., & Lederberg, A. R. (2018). Relations between teacher talk characteristics and child language in spoken-language deaf and hard-of-hearing classrooms. *Journal of Speech, Language, and Hearing Research*, **61**(12), 2977–2995. https://doi.org/10.1044/2018_JSLHR-L-17-0475

Erlich, O., & Shavelson, R. (1978). The search for correlations between measures of teacher behavior and student achievement: Measurement problem, conceptualization problem, or both? *Journal of Educational Measurement*, **15**, 77–89. https://doi.org/10.1111/jedm.1978.15.issue-2

Gann, C. J., Gaines, S. E., Antia, S. A., Umbreit, J., & Liaupsin, C. J. (2015). Evaluating the effects of function-based interventions with deaf or hard-of-hearing students. *Journal of Deaf Studies and Deaf Education*, **20**(3), 252–265. https://doi.org/10.1093/deafed/env011

Garberoglio, C. L., Gobble, M. E., & Cawthon, S. W. (2012). A national perspective on teachers' efficacy beliefs in deaf education. *Journal of Deaf Studies and Deaf Education*, **17**(3), 367–383. https://doi.org/10.1093/deafed/ens014

Geers, A. E., & Hayes, H. (2011). Reading, writing, and phonological processing skills of adolescents with 10 or more years of cochlear implant experience. *Ear and Hearing*, **32**(1), 49S–59S.

Gill, B., Shoji, M., Coen, T., and Place, K. (2016). *The content, predictive power, and potential bias in five widely used teacher observation instruments* (REL 2017–191). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory MidAtlantic. Retrieved from http://ies.ed.gov/ncee/edlabs.

Guardino, C., & Antia, S. D. (2012). Modifying the classroom environment to increase engagement and decrease disruption with students who are deaf or hard of hearing. *Journal of Deaf Studies and Deaf Education*, **17**(4), 518–533.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, **32**(4), 341–370. https://doi.org/10.3102/1076998606298043.

Johnson, E. S., Crawford, A., Moylan, L. A., & Zheng, Y. (2020). Validity of a special education teacher observation system. *Educational Assessment*, **25**(1), 31–46. https://doi.org/10.1080/10627197.2019.1702461

Jones, N. D., Bell, C. A., Brownell, M., Qi, Y., Peyton, D., Pua, D., Fowler, M., & Holtzman, S. (2022). Using classroom observations in the evaluation of special education teachers. *Educational Evaluation and Policy Analysis*, **44**(3), 429–457. https://doi.org/10.3102/01623737211068523

Jones, N. D., & Brownell, M. T. (2014). Examining the use of classroom observations in the evaluation of special education teachers. *Assessment for Effective Instruction*, **39**(2), 112–124. https://doi.org/10.1177/1534508413514103

Knoors, H. (2007). Educational responses to varying objectives of parents of deaf children: a Dutch perspective. *Journal of Deaf Studies and Deaf Education*, **12**(2), 243–253.

Knoors, H., & Hermans, D. (2010). Effective instruction for deaf and hard-of-hearing students: teaching strategies, school settings, and student characteristics. In M. Marschark & P. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (Vol. **2**, pp. 57–71). Oxford University Press.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, **86**(3), 643–680.

Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve the fidelity of evidence-based practices: A review of studies. *Teacher Education and Special Education*, **33**(4), 279–299. https://doi.org/10.1177/0888406410371643

Lederberg, A. R., Branum-Martin, L., Webb, M., Schick, B., Antia, S., Easterbrooks, S., & Connor, C. M. (2019). Modality and interrelations among language, reading, spoken phonological awareness, and fingerspelling. *Journal of Deaf Studies and Deaf Education*, **24**(4), 408–423. https://doi.org/10.1093/deafed/enz011

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for Mixed Models*. SAS Institute Inc.

Marschark, M., & Knoors, H. (2018). Sleuthing the 93% in deaf education. In H. Knoors & M. Marschark (Eds.), *Evidence-based practices in deaf education* (pp. 1–32). Oxford University Press.

Marsh, H. W., Wen, Z., & Hau, K. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psyhological Methods*, **9**(3), 275–300.

Martin, N. A., & Brownell, R. (2011). *Expressive One-Word Picture Vocabulary Test-4 (EOWPVT-4)*. Academic Therapy Publications.

Marzano, R. J., Marzano, J. S., & Pickering, D. (2003). *Classroom management that works: Research-based strategies for every teacher*. ASCD.

McKinney, T., & Vasquez, E. (2014). There's a bug in your ear: using technology to increase accuracy of DTT implementation. *Education and Training in Autism and Developmental Disabilities*, **49**(4), 594–600. https://www.jstor.org/stable/24582354

McLean, L., Sparapani, N., Toste, J. R., & Connor, C. M. (2016). Classroom quality as a predictor of first graders' time in non-instructional activities and literacy achievement. *Journal of School Psychology*, **56**, 45–58. https://doi.org/10.1016/j.jsp.2016.03.004

McLeskey, J., Barringer, M.-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017). *High- leverage practices in special education*. Council for Exceptional Children & CEEDAR Center.

Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics*, **106**(3), 1–9.

Moog, J. S., Popelka, G. R., Geers, A. E., & Russo, M. H. (1990). *Early speech perception test for profoundly hearing-impaired children*. Central Institute for the Deaf.

Moustaki, I. (2000). A latent variable model for ordinal variables. *Applied Psychological Measurement*, **24**(3), 211–223.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**(1), 115–132.

Newby, R. F., Hallenbeck, C. E., & Embretson, S. (1983). Confirmatory factor analysis of four general neuropsychological models with a modified Halstead-Reitan Battery. *Journal of Clinical and Experimental Neuropsychology*, **5**(2), 115–133.

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews*, **7**(1), 1–55.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, **38**(2), 109–119. https://doi.org/10.3102/0013189X09332374

Reddy, L. A., Fabiano, G., Dudek, C. M., & Hsu, L. (2013a). Development and construct validity of the classroom strategies scale-observer form. *School Psychology Quarterly*, **28**(4), 317–341. https://doi.org/10.1037/spq0000043

Reddy, L. A., Fabiano, G., Dudek, C. M., & Hsu, L. (2013b). Predictive validity of the classroom strategies scale-observer form on statewide testing scores: an initial investigation. *School Psychology Quarterly*, **28**(4), 301–316. https://doi.org/10.1037/spq0000041

Reeve, J. (2006). Teacher as facilitator: What autonomy-supportive teachers do and why their students benefit. *The Elementary School Journal*, **106**(3), 225–236.

Rivera, M. C., Hasko, J., Jun, M. J., & Antia, S. D. (2020). The journey to effective pre-service Teacher self-evaluation: Improving the quality of the learning environment. *Paper presented at the Association of College Educators Deaf Hard of Hearing*, Atlanta, GA.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, **73**(2), 417–458.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement.*, **34**, 1–97. https://doi.org/10.1007/BF03372160

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. University of Tennessee Value-Added Research and Assessment Center.

Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the Classroom Observation of Student-Teacher Interactions (COSTI) for kindergarten reading instruction. *Early Childhood*

Research Quarterly, **27**, 316–328. https://doi.org/10.1016/j.ecresq.2011.09.004

Stinson, M. S., & Kluwin, T. N. (2011). Educational consequences of alternative school placement. In M. Marschark & P. Spencer (Eds.), *The Oxford handbook of deaf studies, language, and education* (Vol. **1**, 2nd Ed.). Oxford University Press.

Stronge, J. H. (2018). *Qualities of effective teachers* (3rd Ed.). Association of Supervision and Curriculum Development.

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross- case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, **62**(4), 339–355. https://doi.org/10.1177/0022487111404241

Taylor, B. M., Pearson, P. D., Clark, K. F., & Walpole, S. (1999). Center for the improvement of early reading achievement: Effective schools/accomplished teachers. *Reading Teacher*, **53**(2), 156–159. https://www.jstor.org/stable/20204767

Woodcock, R. W., Shrank, F. A., McGrew, K. S., & Mather, N. (2007). *Woodcock-Johnson III Normative Update (WJ III NU)*. Riverside Publishing.

Woolsey, M. L., Harrison, T. J., & Gardner, R. III (2004). A preliminary examination of instructional arrangements, teaching behaviors, levels of academic responding of deaf middle school students in three different educational settings. *Education and Treatment of Children*, **27**(3), 263–279.

Wright, S. J. (1997). *Primal-dual interior-point methods*. Society for Industrial and Applied Mathematics.

Zahorik, J., Halbach, A., Ehrle, K., & Molnar, A. (2003). Teaching practices for smaller classes. *Educational Leadership*, **61**(1), 75–77.