Deafness & Education International

ISSN: 1464-3154 (Print) 1557-069X (Online) Journal homepage: https://www.tandfonline.com/loi/ydei20

Storybook selection criteria used by teachers of d/Deaf and hard-of-hearing prereaders communicating in English

Amy Louise Schwarz, Meagan Jurica, Charlsa Matson, Rachel Stiller, Taylor Webb-Culver & Hervé Abdi

To cite this article: Amy Louise Schwarz, Meagan Jurica, Charlsa Matson, Rachel Stiller, Taylor Webb-Culver & Hervé Abdi (2019): Storybook selection criteria used by teachers of d/Deaf and hard-of-hearing prereaders communicating in English, Deafness & Education International, DOI: 10.1080/14643154.2019.1572284

To link to this article: https://doi.org/10.1080/14643154.2019.1572284

Published online: 06 Feb 2019.

DEAFNESS & EDUCATION INTERNATIONAL
https://doi.org/10.1080/14643154.2019.1572284

Storybook selection criteria used by teachers of d/Deaf and hard-of-hearing prereaders communicating in English

Amy Louise Schwarz [a], Meagan Jurica[a], Charlsa Matson[a], Rachel Stiller[a], Taylor Webb-Culver[a] and Hervé Abdi[b]

[a]Department of Communication Disorders, Texas State University, Round Rock, TX, USA; [b]School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX, USA

ABSTRACT

For d/Deaf and hard-of-hearing prereaders who communicate predominately in spoken and/or signed English (DHH-English), Teachers of the d/Deaf (TODs) read books aloud to increase English skills, auditory-verbal comprehension, sequencing skills, verbal reasoning, background knowledge, and sight word recognition. Teachers struggle to select appropriate books for read alouds. Unfortunately, the only available book selection system for TODs serving DHH-English was developed for a different purpose than identifying storybooks to increase English skills and was based on a methodology difficult to evaluate. To help teachers select books, we created an empirically-derived difficulty-level system based on the judgments of 69 experienced TODs serving DHH-English preschoolers and kindergartener across the United States. We used 14 storybooks and asked the TODs to group these books and order these groups based on how difficult the storybooks would be for DHH-English prereaders to understand when the books were read aloud. Additionally, TODs described the groups they created and why they assigned the books in their groups. We used these descriptions in a content analysis to derive a glossary of book characteristics. Finally, we analysed these results with a multivariate technique that simultaneously analysed the glossary and ranking data to create a difficulty-level scale and to identify exemplar books. The book selection system includes a two-tiered glossary, a six-point difficulty-level scale and exemplar books for each scale level. We discuss how this empirically-derived book selection system for read alouds with DHH-English prereaders can be paired with the three existing evidence-based read-aloud interventions designed to increase oral language skills.

ARTICLE HISTORY
Received 23 August 2018
Revised 3 January 2019
Accepted 14 January 2019

KEYWORDS
Preliteracy; spoken English; signed English; book sharing; read alouds; DiSTATIS; sorting task; hearing impaired

Introduction

Reading aloud books to children that are too advanced for them to read independently is a critical activity because it helps children acquire the knowledge

CONTACT Amy Louise Schwarz amy.louise.schwarz@txstate.edu Department of Communication Disorders, Texas State University, 200 Bobcat Way, Round Rock, TX 78665, USA

they need to succeed as independent readers (Anderson, Hiebert, Scott, & Wilkinson, 1985). For the 67% of d/Deaf and hard-of-hearing (DHH) prereaders who communicate predominately in English (DHH-English, Gallaudet Research Institute, August, 2013) – regardless of the modality (spoken or signed) used – teachers of the d/Deaf (TODs) focus on increasing basic spoken English language skills, auditory-verbal comprehension, sequencing skills, verbal reasoning, background knowledge, and sight word recognition (Schwarz, Guajardo, & Hart, 2017). Little attention, however, has been paid to how storybooks for read alouds are selected by TODs serving DHH children communicating in spoken English only or in both spoken and some form of signed English (DHH-English): Only one study identified book characteristics, a difficulty-level scale, and exemplar books that TODs serving DHH-English can use to guide read-aloud book selection (Stewart, Bennett, & Bonkowski, 1992). Unfortunately, this selection system cannot help TODs select storybooks for read alouds meant to increase English language skills because it was developed for a different purpose, uses a methodology difficult to evaluate, and includes books with an insufficient number of multicultural themes and characters. Our study addresses these concerns by presenting an empirically-derived book selection system based on the judgments of experienced TODs serving DHH-English preschoolers and kindergarteners.

To frame our study within the existing literature, we first discuss the oral language domain of emergent literacy, how the read-aloud goals that TODs have for DHH-English align with these domains, and the three evidence-based read-aloud interventions that focus on this domain. Then we discuss Stewart and colleagues' Signability Index, and why TODs serving DHH-English need a different approach for selecting storybooks for read-alouds meant to increase English language skills. Finally, we explain how identifying similarities in educators' book sorting decisions can be used to create a book selection system.

Emergent literacy skills, DHH read-aloud goals, and evidence-based read-aloud interventions

The current study focuses on oral language, which includes vocabulary, sentence structures, grasp of story structure (Storch & Whitehurst, 2002), and the ability to make inferences (Lepola, Lynch, Kiuru, Laakkonen, & Niemi, 2016) because these skills are linked to reading comprehension in third-grade (Language and Reading Research Consortium, 2018; Lepola et al., 2016). In prior work, Schwarz et al. (2017) conducted a content analysis of read-aloud goal statements that included TODs serving DHH prereaders communicating in only spoken English and TODs serving DHH prereaders communicating in both spoken English and some form of signed English. Schwarz et al. (2017) found that these two groups of TODs used read alouds to increase DHH-English prereaders':

- English vocabulary,
- English syntax,
- auditory-verbal comprehension,
- background knowledge,
- ability to answer questions about stories, and
- ability to sequence stories.

Note that the same TODs who participated in the Schwarz et al. (2017, 2018) studies also participated in the current study. With Schwarz et al. (2017) finding that these two groups of TODs reported similar read-aloud goals, we did not expect to find differences based on how these two groups of TODs selected storybooks for read alouds. However, answering a survey and sorting storybooks based on difficulty are very different tasks. So, in the current study, we verified whether these two groups of TODs differed in the criteria they used to select storybooks.

There are three empirically tested read-aloud interventions that focus on building oral language skills in young children. These read-aloud interventions, which we define in the following sections, include: a dialogic reading style (e.g. Zevenbergen & Whitehurst, 2003), a literal-inferential reading style (e.g. van Kleeck, Vander Woude, & Hammett, 2006), and a TextTalk reading style (e.g. Beck & McKeown, 2007).

Dialogic reading promotes conversational vocabulary and grammar in order to help young children become storytellers (for a review, see Zevenbergen & Whitehurst, 2003). An adult who uses this read-aloud style, asks the child open-ended and wh-questions and also repeats, expands on, and recasts what the child says. The adult follows the child's lead, which in dialogic reading, means that the adult does not read all the words in the text but instead talks about the illustrations that interest the child (for a review, see Zevenbergen & Whitehurst, 2003). Dialogic reading has been linked to significant increases in conversational vocabulary and grammar abilities in hearing toddlers (see Huebner, 2006; Mol, Bus, De Jong, & Smeets, 2008; Zevenbergen & Whitehurst, 2003 for reviews) and DHH children (Fung, Chow, & McBride-Chang, 2005; Lederberg, Miller, Eastabrooks, & Mcdonald Connor, 2014; Trussell & Easterbrooks, 2014). Also, caregiver training in dialogic reading has led to significant changes in how caregivers read books to DHH toddlers (Dirks & Wauters, 2015, 2018). When selecting books for dialogic reading intervention, Whitehurst and colleagues advise adults to select books with illustrations that can introduce new vocabulary and can tell the complete story without relying on information from the text (Whitehurst et al., 1994).

The literal-inferential style of read aloud attempts to develop literal and inferential skills in young children. In the context of read alouds, the difference between literal and inferential language hinges on whether the storybook text or illustrations provides all the information the child needs to understand the story (van Kleeck et al., 2006). It is recommended that adults direct

approximately 60% to 70% of the discussion to promote literal-level language using the strategies described above in the description of dialogic reading, such as labelling and asking questions about content present in the text or in the illustrations (van Kleeck et al., 2006). It is recommended that adults direct approximately 30% to 40% of the discussion about the book to promoting inferential-level language, which requires the adult to ask the child questions that go beyond information explicitly stated in the text or illustrations (van Kleeck, 2014; van Kleeck et al., 2006). For example, the adult may ask the child about the different feelings, perspectives, and motives of the characters or the adult may ask the child to make connections between the child's own background and events in the storybook. The adult may also ask the child to define words in the text and make connections across other storybooks that the child knows, and finally, the adult may ask the child to comment on possible cause-and-effect relationships in the story (van Kleeck et al., 2006). This literal-inferential procedure seems beneficial because hearing preschoolers with language impairment who were exposed to a literal-inferential read-aloud intervention improved their literal and inferential language skills (van Kleeck et al., 2006).

The *TextTalk* read aloud style promotes the child's acquisition of high-level vocabulary that is often found in children's books (Beck & McKeown, 2001, 2007). In a TextTalk read aloud, the adult does not show the illustrations while reading the text. Instead, the adult reads a section of the story aloud, asks the child to construct an understanding of the text, and then shows the illustrations. The adult encourages the child to explain ideas and scaffolds the child's thinking by asking follow-up questions that require the child to elaborate on those ideas. The adult also selects three high-level vocabulary words per book and explicitly teaches the words to the child after reading the book (Beck & McKeown, 2001). When compared to business as usual read alouds, the *TextTalk* style of read aloud significantly increased hearing children's acquisition of sophisticated vocabulary (Beck & McKeown, 2007). Storybooks appropriate for TextTalk read alouds are intellectually challenging and do not rely on illustrations to relay the story content (Beck & McKeown, 2001, 2007).

These three evidence-based read-aloud interventions require storybooks across a range of difficulty levels. How successful are teachers at selecting books for read alouds? It seems that professionals often struggle to select books that are appropriate for read alouds (Beck & McKeown, 2001; Damber, 2015; McGee & Schickendanz, 2007). Although several scholars have suggested factors that educators need to consider when selecting books for read alouds (e.g. Anderson, Anderson, Shapiro, & Lynch, 2001; Beck & McKeown, 2001; Elster, 1998; Griffin, 1970; Martinez & Roser, 1985; McGee & Schickendanz, 2007; Schwarz et al., 2015) only one system – which we discuss next – exists for TODs serving d/Deaf and hard-of-hearing (DHH) prereaders and only for those prereaders using some form of signed English with ASL features.

Stewart and colleagues' Signability Index

Stewart et al. (1992) created a Signability Index to help TODs organize books across literary genres. The Signability Index targets preschool through second grade total communication classrooms and was developed for a variety of objectives based on how difficult books were for TODs to sign in English with ASL features. This system has a broader purpose and applies to a larger grade-level range than the current study, which focuses on how TODs select storybooks for read alouds meant to increase the English language skills in DHH preschoolers and kindergarteners that are necessary for later text comprehension. Although the Signability Index includes ASL features in signing English, the Signability Index's focus is fundamentally on transmitting the English language. The Signability Index does not account for the visual and spatial language characteristics fundamental to a natural sign language, such as American Sign Language (Hayes & Shaw, 1994). In fact, Hayes and Shaw (1994) were so opposed to the use of the Signability Index to select storybooks for ASL storytelling sessions that Hayes and Shaw (1994) created a storybook selection system specifically designed for ASL storytelling sessions.

To develop the Signability Index, Stewart and colleagues worked with TODs to derive 13 characteristics, a six-level difficulty scale, and several exemplar books for each level of the scale. The 13 characteristics are: (1) reading rate needed to maintain DHH children's attention, (2) sentence length, (3) language complexity (4) repetition, (5) text density, (6) word/verbal imagery, (7) passage complexity, (8) concreteness, (9) book length, (10) plot complexity, (11) level of ASL skills needed to translate the book, (12) ease of articulating the story in sign, and (13) whether sign practice was recommended for TODs before attempting to read the books aloud.

Stewart and colleagues only selected books that TODs personally enjoyed reading aloud and that translated well into ASL. Stewart *et al*. do not mention how the books were selected or how they determined whether the books would translate well into ASL. Based on Stewart and colleagues' description, it is difficult to properly evaluate the scientific rigour underlying the Signability Index and it is unclear whether the characteristics identified are mutually exclusive. For example, how does passage complexity differ from language complexity, word/verbal imagery, plot complexity, and concreteness? Furthermore, the Signability Index includes books across literary genres, which is problematic, because adults use different book selection criteria when selecting read alouds that include sound play (Ukrainetz, Cooney, Dyer, Kysar, & Harris, 2000) and print-referencing (Zucker, Ward, & Justice, 2009). Finally, only approximately 12% of the books in the Stewart and colleagues' Signability Index include multicultural characters and themes. This percentage falls short of the recommendation from the International Literacy Association (July, 2018) that requires educators to cultivate students' awareness of, and respect for, diversity

through the curriculum by 'using instructional materials that are linked to students' backgrounds' (p. 6). In the United States, however, 35% of DHH children have a home language other than English or ASL, with 19.4% of the children being raised in households in which Spanish is regularly used (Gallaudet Research Institute, August, 2013). We could not find data indicating how many of these DHH children will acquire their home language and/or English in a spoken or signed form, but we expect this percentage of culturally and linguistically diverse DHH children to increase because the United States is projected to lack one racial or ethnic majority by 2055, with most of the growth created by immigrants from Latin America and Asia (Cohn & Caumont, 2016).

Using similarity decisions in sorting tasks to create a book selection system

Our goal was to create a book selection system for TODs serving DHH-English using a rigorous statistical methodology. To do so, we adapted a sorting task methodology that is popular in taste sensation and perception research because it is efficient and has strong validity (Chollet, Lelièvre, Abdi, & Valentin, 2011). Sorting tasks are based on our natural ability to categorize items. In taste sensation and perception research, people group items – such as beer or wine – into as many groups as they want based on sensory and perceptual similarities (Chollet et al., 2011). In many sorting studies in the taste sensation and perception literature, participants describe the groups they created once they have completed the sorting task (e.g. Blancher *et al.,* 2007; Chollet *et al.,* 2011; Patris, Gufoni, Chollet, and Valentin, 2007; Santosa *et al.,* 2010; Tang and Heymann, 2002). The statistical analysis of these sorting tasks proceeds in two steps: First, the sorting task results are represented by maps of the stimuli that describe the major sources of variance in the data and can be used to identify groups of similar stimuli. Second, the participants' verbal descriptions are superimposed onto the data map and provide insight into the participants' reasons for creating the groups of products. Nonparametric statistical inferences, obtained by cross-validation techniques, can be used to evaluate the robustness and significance of the results.

In prior work, Schwarz et al. (2015) adapted the sorting task to create a book selection system for speech-language pathologists (SLPs) serving children at the preschool language level. In that study, Schwarz et al. (2015) asked SLPs to sort storybooks into similar groups based on how difficult SLPs thought the storybooks would be for prereaders to understand when the storybooks were read aloud to them. SLPs then explained the characteristics defining each group of similar storybooks that they had created. Schwarz et al. (2015) conducted a content analysis on the SLP interview data and extracted a glossary of key book characteristics. Schwarz et al. (2015) used the same methodology to analyse the data as Chollet et al. (2011). Schwarz et al. (2015) used the number of ranked groups of storybooks the SLPs collectively created to identify

the number of gradients in a difficulty level scale for the storybooks. Schwarz et al. (2015) ordered the scale levels from easy to difficult based on the SLPs' descriptions of the stacks of storybooks captured in the glossary of book characteristics. Schwarz et al. (2015) created a Likert-type scale in the form of a table by using the levels of difficulty identified by storybooks as the column headers, with the easiest level of difficulty on the far left of the table and the hardest level of difficulty on the far right of the table. Then Schwarz et al. (2015) placed the book characteristics as the row headers. The qualifiers for each book characteristic populated the remaining cells of the table and explained the attributes of each book characteristic for each scale level. Schwarz et al. (2015) then identified exemplar books for each scale level. In the current study, we extend this methodology to a new group of educators and a new group of storybooks to discover the criteria TODs use to select storybooks for read alouds.

Summary and research questions

The English language goals of TODs serving DHH-English are similar to those mentioned in the oral language domain of emergent literacy models (Lepola et al., 2016; Scarborough, 2001; Storch & Whitehurst, 2002), which are linked to reading comprehension in third grade hearing children (Language and Reading Research Consortium, 2018; Lepola et al., 2016). Evidence-based read-aloud interventions focused on building oral language skills require storybooks based on different difficulty levels. The one existing book selection system operationalized for TODs serving DHH-English (Stewart et al., 1992) cannot be used to help TODs select read alouds meant to increase the English language skills of DHH-English. Our purpose was to identify the criteria actually used by TODs when they select storybooks for read alouds meant to increase the English language skills of DHH-English preschoolers and kindergarteners. In our study, we used storybooks without sound play and without prominent print features because adults use different criteria to select read alouds for these purposes (Ukrainetz et al., 2000; Zucker et al., 2009). To create a book selection system for TODs serving DHH-English using a rigorous statistical methodology, we extended a methodology to TODs that Schwarz et al. (2015) adapted from the taste and perception literature to create a book selection system for SLPs serving children at the preschool language level.

Our specific research questions were:

(1) After sorting storybooks from preschool curricula, what terms do TODs use when describing their reasons for placing a book in a particular stack?
(2) What are the patterns of similarity and dissimilarity in the stacks created by TODs when sorting the storybooks?
(3) When the terms the TODs used to describe why they placed books in particular stacks (Research Question 1) are superimposed onto the patterns of

similarity and dissimilarity in the stacks of sorted storybooks (Research Question 2), what are the terms that best describe each stack of storybooks?

(4) As a group, what are the overall levels of difficulty TODs identified, and which storybooks best represent each level of difficulty?

Methods

This study was approved by the University Institutional Review Board where the first author works on 9 June 2015 (application number: 2015Q3357).

Participants

Data from this study is part of a larger study that explored the preliteracy practices of 85 experienced teachers of DHH prereaders. The current study only includes the 69 TODs serving DHH prereaders who communicate in English and so only their recruitment, geographic locations in relation to the U.S. Census, and demographics are discussed here. To participate, TODs had to have two or more years of experience teaching DHH prereaders at any level and had to have experience teaching DHH prereaders in preschool or kindergarten. Recruiting took place in two phases and resulted in 69 participants, with 27 participants recruited in Phase 1 and 42 participants recruited in Phase 2.

To identify programmes in Phase 1, the programmes had to be listed in one of the following sources: (a) the Guide to Education Programs for Deaf Students in the *American Annals of the Deaf Reference Issue* (2013), (b) the Laurent Clerc Education Center's (2015) website list of Schools and Programs for the DHH in the U.S., and (c) the Educational Service Center Region 11's (2014–2015) list of Regional Day School Program's for the Deaf Directory. The programmes had to contain four or more TODs who taught prereaders at the preschool and/or kindergarten level as either classroom or itinerant teachers. The email address and name of the programme contacts in the aforementioned publications had to be accurate. The schools were then divided into the U.S. census geographic regions. Recruiting information was emailed to the lower school principals, superintendents, directors, and/or d/Deaf education coordinators. Interested school administrators scheduled a conference call with the first author to discuss the study further. In phase two, TODs were recruited to participate while attending one of the following three conventions focused on education of DHH students in 2016: (a) the Conference of American Instructors of the Deaf in Louisville, Kentucky, (b) the Alexander Graham Bell Association Conference in Denver, Colorado, and (c) the Statewide Conference of Education of the DHH in San Marcos, Texas. See Table 1 for a summary of demographic information on participating TODs.

Table 1. Demographic information of TODs.

Categories	Total ($N = 69$)
Education	
BA/BS	21
MA/MS or PhD	48
Years of experience teaching DHH students	$M = 14.45$ ($SD = 7.77$, min = 3.00, max = 25.00)
Females	69
Ethnicity	
African American	1
Asian	3
Caucasian	52
Hispanic	8
Not Reported	5
Hearing status	
Profound loss	1
Hard of hearing	7
Hearing	61
Member of the Deaf Community	
No	62
Yes	7
Current work setting	
Private	6
Public	56
State School	4
Not Reported	3
Type of Teacher	
Classroom	57
Itinerant[a]	12
Communication Modality Used by the DHH Students the TODs	
Serve Speech Only	34
Multimodal Communication	
Signing Exact English	20
Conceptually Accurate English and Pidgin English	9
Signed English	2
Unspecified English sign system	4
Read Alouds	
Number of read alouds each week	$M = 6.93$ ($SD = 4.94$, min = 1.00, max = 21.00)
Group size	$M = 4.81$ ($SD = 2.49$, min = 1.00, max = 9.00)
Regional Boundary of United States Census	
Midwest	5
Northeast	5
South	57
West	2

[a]Itinerant teachers drew on their classroom experience with DHH-English.

Storybook selection

To select authentic children's books used in preschools, we searched (in the United States) the websites of the largest 100 school districts (Nolan, 2015), all state schools for the deaf, many private schools for children with hearing loss that focus on language, speech, and listening skills, and large regional day schools for the deaf (Laurent Clerc National Deaf Education Center, 2015). Our search yielded 25 preschool curricula. The booklists of seven curricula were available online and included 348 books (less duplicates). We excluded 147 books because they were not available in the catalogue of juvenile holdings at the university of the first author. To identify the types of books in the 201 remaining children's books, three undergraduate research assistants working in the first

author's research lab coded the books using Lynch-Brown and Tomlinson's (2008) 11 definitions of children's literature, which are: (1) picture storybooks that tell a fictional narrative through pictures and words, (2) baby books, (3) interactive books that invite recitation and chanting, (4) toy books (e.g. pop-up books), (5) wordless picture books, (6) concept books (i.e. alphabet books, counting books, idea books, colour books), (7) pattern books (i.e. decodable picture books, predictable books, sight word books, books with sound play), (8) picture books with chapters, (9) picture books with mostly talk bubbles or graphic novels, (10) dual language books that tell the whole story in both English and another language, and (11) easy-readers meant for children to read independently and so contain simplified text.

When we began reliability training, only 162 of the 201 storybooks were available in the university library. Three undergraduate research assistants working in the first author's research lab were trained to apply the definitions to a randomly selected 20% (i.e. 32 books) of the 162 books. Once acceptable inter-judge reliability had been achieved in training, the students independently applied the definitions to the remaining 130 children's books. Within a month, we were able to access the remaining 39 books (162 initial books + 39 books = 201 total books) at the university library and to assess reliability in applying Lynch-Brown and Tomlinson's (2008) definitions. The children's books needed for the current study told a fictional narrative through pictures and words only. For the purposes of calculating inter-judge reliability, books coded as telling a fictional narrative through pictures and words were assigned the number 1 and books coded as meeting one of the remaining 10 definitions were assigned the number 0. Table 2 contains the reliability statistics for storybook selection in our study. The few disagreements were resolved by consensus, with most disagreements due to coder fatigue. The undergraduate research assistants agreed that 73 of the 201 books were appropriate for this study.

To ensure that these 73 storybooks had wide appeal, we verified that they were listed in the MetaMetrics database, a privately-owned company that markets lexile scores to match text difficulty to the reading level of independent readers. MetaMetrics calculates lexile scores using the variables of word frequency and sentence length (Nelson, Perfetti, Liben, & Liben, 2012). The scores represent both a given text's difficulty level and an independent

Table 2. Reliability statistics for three undergraduate research assistants identifying storybooks appropriate for this study.

Number of books included in reliability calculation	Fleiss's Kappa	Percentage of Disagreements	Number of Books About Which the Judges Disagreed
130 books (from initial set)	.89	7.70%	10
39 additional books	.93	5.10%	2
Overall: 169 books	.90	7.10%	12

Note: Thirty-two books from the initial set of 162 books were used for reliability training.

reader's reading level (MetaMetrics, 2014). The MetaMetrics database was chosen for two reasons: First, the National Governors Association Center for Best Practices and Council of Chief State School Officers (2012) recommends lexile scores for levelling texts for children learning to read independently (Nelson et al., 2012); Second, in a storybook sorting study with speech-language pathologists, Schwarz et al. (2015) found that lexile scores approximated the extreme ends of the difficulty-level scale. Although all 73 storybooks were available at online booksellers, four storybooks were excluded from consideration because they were not listed in the MetraMetrics database.

According to Chollet et al. (2011), in order to give stable results, the sorting task can be used with between 9 and 20 items, with an optimal number of 12. Here, we included 14 storybooks in our sorting task so that TODs could choose to reject a small number of storybooks during the sorting task without jeopardizing our ability to find a stable pattern. TODs could reject a storybook if they did not think it was appropriate for their students. We needed to select two sets of 14 storybooks, one for the current study and one for a future validity study. Both sets of 14 books were very similar with one set randomly assigned to the current study and the other set assigned to the future validity study. To select the two sets of similar 14 storybooks, we used the following criteria: (a) the distribution of lexile scores present in the 69 storybooks was also present in both sets of the 14 storybooks, (b) a storybook author appeared only one time in the list of 14 storybooks, and (c) the distribution of multicultural characters/themes in both sets of 14 storybooks approximated 36%, which is the percent of DHH children who are from homes in which Spanish or another language besides English and ASL are used (Gallaudet Research Institute, August, 2013). See the Appendix for the list of 14 storybooks used in this study, the titles abbreviations, the books' MetraMetrics lexile scores, the curricula in which the books were found, and the number of schools for the deaf and school districts that use each curriculum.

Procedure

The sorting task methodology used in the current study is an adaptation of the 435 sorting task methodologies used in the taste and perception literature (see, e.g., Chollet et al., 2011). In our application of the sorting task, we asked each TOD to group storybooks based on how difficult they thought the books would be for prereaders to understand when the books were read aloud to them.

Examiners in the current study were trained to administer the procedure through a 1 ½-hour video and one live simulation with a trainer. Examiners read a script during the experiment to ensure that all participants had a similar experience. Participants met individually with an examiner. All meetings were audio-recorded. Each participant received the 14 storybooks in a unique

order that was determined using a random number generator. The procedure took an average of 53 min to complete. After participants consented, they completed two surveys that collected information about their demographics, read-aloud practices, and preferences for and familiarity with the 14 storybooks. Then the examiner gave the 14 storybooks to each TOD at one time and asked him or her to sort the books into as many stacks as necessary according to how difficult the books would be for DHH preschoolers or kindergarteners to understand when the books were read aloud to them. The participants were not provided any guidelines for sorting the storybooks. Therefore, TODs could create as few or as many stacks of storybooks that they believed were necessary in order to capture the different levels of difficulty found in the array of 14 storybooks.

Next the examiner asked each TOD to describe the characteristics of each stack of books (stack-level descriptions) and then to describe why he or she placed each book in each stack (book-level descriptions). TODs were allowed to re-sort the books when giving their descriptions. On the few occasions when TODs re-sorted one or more storybooks, the examiner asked them to provide their stack-level and book-level descriptions again so that their descriptions would reflect how their thinking had changed.

Fidelity

The first author and 13 students administered the experimental script. All were trained using the same 1 ½-hour video and through one live simulation with a trainer. A graduate student calculated fidelity on 100% of the data by comparing the transcription of each examiner's administration of the experimental script to the master script used in training the examiners on a word-for-word basis. The average percentage correct was 93.91% ($SD = 7.18\%$, min = 67.22%, max = 100.00%). Four transcripts had less than 80% fidelity. The first author reviewed these four transcripts and determined that even though the examiners varied on a word-for-word basis from the master script, all steps in the procedure were followed and variations did not materially change the experience these four participants had in comparison to the remaining 65 participants. Therefore, data from all 69 TODs using English were included in this study.

Analysis plan

The 69 TODs who participated in the current study also participated in Schwarz et al.'s (2017) content analysis of read-aloud goal statements. Although the TODs showed no difference in the read-aloud goals based on the modality (spoken English, spoken English and some form of signed English) that their students used to communicate, answering a survey and sorting storybooks are very

different tasks. For this reason, we present the major results from the current study for TODs serving DHH prereaders communicating in only spoken English and TODs serving DHH prereaders communicating in spoken English and some form of signed English to determine whether these groups of TODs differed in their sorting judgements.

Recall that the TODs completed a survey of their preferences and familiarity for the 14 storybooks included in this study. Storybooks preferred by less than 50% of either group of TODs were removed from the analysis. The TODs preferences and familiarity for each storybook are summarized in the results' section.

First question: What terms do TODs use to describe their ranked stacks of books?

We identified the glossary of book characteristics from transcriptions of the TOD's stack-level and book-level descriptions using content analysis (Schreier, 2012a; Weber, 1990). Content analysis reduces preserved forms of communication – transcribed audio recordings in our case – into key content categories. The unit of analysis was the category, which we define as a mutually exclusive collection of words, phrases, sentences, and series of sentences with different meanings that, when taken together, describe one theme that TODs consider important when assessing the difficulty level of storybooks (Schreier, 2012a; Weber, 1990). We developed the coding system using both a concept-driven and data-driven process, as is common in qualitative content analysis (Schreier, 2012a). We took a concept driven approach by considering the book characteristics identified by prior research (Schreier, 2012a). Specifically, we considered the book characteristics identified in Stewart et al. (1992) Signability Index for TODs using an English-based sign system that incorporated ASL features and in Schwarz et al.'s (2015) book selection system based on the judgments of speech language pathologists.

We took a data-driven approach by selecting a small but representative portion of the interview data to identify categories through the three steps of open coding discussed by Schreier (2012b). For open coding, we randomly selected 20% of the interview data, with half of the interview data from TODs serving DHH prereaders who communicate in only spoken English and the other half of the data from TODs serving DHH children who communicate in both spoken English and some form of signed English. In the first step of open coding, the pairs of coders and the first author reread the interviews for 20% of the data several times to identify relevant categories mentioned by the TODs. Through this iterative process, similarities and dissimilarities in the way participants use concepts became apparent (Schreier, 2012b). Second, the research team met to identify similarities in the relevant categories that the pairs of coders and the first author identified in the data, which is the second step in open coding. Third, the research team differentiated between which relevant categories were superordinate categories and which were subordinate

categories in the coding scheme, the third step in open coding. In this third step of open coding, we developed a two-level coding scheme. The first tier of our coding system identified the major book characteristics mentioned by the TODs. These book characteristics were the superordinate categories in our two-tiered coding system. All book characteristics (i.e. superordinate categories) were coded (exhaustiveness requirement); and each main category captured one concept (unidimensional requirement), was mutually exclusive from other main categories (mutual exclusivity requirement), and was coded at least once (Schreier, 2012b).

The eleven categories of book characteristics extracted from the data were: (a) plot/sequence, (b) vocabulary, (c) concept, (d) background knowledge, (e) illustrations, (f) book length, (g) characters, (h) sentence length, (i) text density (amount of text per page), (j) translation (only TODs who used a form of signed English), and (k) repetition (language and/or events). The concept category describes the relative abstractness TODs assigned to the ideas presented in the storybooks. The background knowledge category describes how much the TODs thought that the events described in the storybook were related to the experience of the DHH children they taught. The four decision rules were as follows:

- Participant comments that did not pertain to the stack-level and book-level descriptions were not coded.
- Only qualified categories were coded (e.g. 'it's a little bit too long' was coded 'BL [book length]: too long').
- When a TOD repeated a category while describing a stack of books or an individual book, the category was coded only once.
- If a TOD contradicted herself or refined her thinking, the original idea was not coded, but the final idea expressed was coded.

Schreier (2012c) recommends piloting the coding scheme in a three-phase process, which we followed. First, coders applied the coding scheme independently to a portion of the data. Schreier (2012c) recommends selecting a percentage of the data for piloting the coding scheme that will allow for categories to be coded multiple times and will account for variability in the data. We chose to use 40% of the remaining data for this purpose, with an equal portion of data randomly selected from TODs serving DHH prereaders who communicate in only spoken English and from TODs serving DHH prereaders who communicate in both spoken English and some form of signed English. So, to pilot the coding scheme, the pairs of coders applied the decision rules and 11 categories to 40% of the remaining data. In the second phase of piloting the coding scheme, Schreier (2012c) recommends calculating inter-judge reliability to verify that the coding scheme is being applied consistently across coders. See the top section of Table 3 for category item agreement percentages.

Inter-coder agreement during piloting indicates that the coders were consistent in their application of the coding scheme and decision rules. The third stage of piloting the coding scheme is to adjust the coding scheme based on inter-coder error (Schreier, 2012c). Differences in applying the decision rules and coding scheme were minor and resolved through group discussion. When only minor revisions to the coding scheme are required after piloting, Schreier (2012c) recommends only recoding affected passages and applying the revised coding scheme to the remaining data. We followed this recommendation. After recoding the affected passages, the pairs of coders applied the revised coding scheme to the remaining 40% of the data. Item agreement between coding pairs for the remaining 40% of data showed consistency in application of the coding scheme. Again, differences in applying the decision rules and coding scheme were minor and resolved through group discussion.

We took a data-driven approach to identify the qualifier scales that define each book characteristic, the subordinate tier of our coding scheme. Recall that our second decision rule for coding categories in the data stated that only qualified book characteristics would be coded. We used only the TODs' book-level descriptions to extract qualifiers for each book characteristic because the TODs' book-level descriptions provided more details about each book than their stack-level descriptions. A list of 1,064 qualifiers across categories, storybooks, and participants was compiled. The first author extracted several scales from the list of qualifiers and generated a draft of standardized qualifier scales for each book characteristic, which reduced the list to 73 unique qualifiers. A coding manual was created that included the 73 unique qualifiers and examples of qualifiers used by the TODs in the interview data. For example, when coding the plot/sequence complexity scale, pairs of coders coded the following phrases as *little hard:* little harder, not too complicated, more difficult but still simple, more complex but still simple, able to follow with guidance.

We applied the first step of open coding (Schreier, 2012c) to refine the standardized qualifier scales and to identify the decision rules needed to consistently

Table 3. Item agreement for two pairs of coders.

	Average	Minimum	Maximum
Categories			
Piloting the first tier of the coding scheme			
First 40% subdivision	89.59%	74.19%	100.00%
Main analysis: Applying the first tier of the coding scheme to the remaining data			
Second 40% subdivision	91.89%	83.73%	100.00%
Standardizing the qualifier scales			
Piloting the standardized qualifier scales			
First 40% subdivision	93.39%	57.89%	100.00%
First 40% subdivision less translation category	–	90.91%	100.00%
Main analysis: Applying the standardized qualifier scales			
Second 40% subdivision	96.27%	89.47%	100.00%

apply the scales. For the first step of open coding, we randomly selected 20% of the list of qualifiers for each book characteristic mentioned by half of the TODs serving DHH prereaders who communicate in only Spoken English and the other half from TODs serving DHH children who communicate in both spoken English and some form of signed English. The pairs of coders compared the list of unstandardized qualifiers for 20% of the data several times to the draft of the standardized list of qualifiers in an iterative process to refine the standardized qualifier scales and to develop decision rules for their consistent application. The two coding rules for matching the list of 73 standardized qualifier scales to the list of 1,064 qualifiers were:

- If a TOD mentioned more than one qualifier when discussing a particular book characteristic, only the first qualifier was coded.
- Qualifiers that did not match a level of one of the standardized qualifier scales were not coded.

Following Schreier's (2012c) recommendation, we piloted the application of the list of standardized qualifier scales to a larger portion of the data. As before, we chose to use 40% of the remaining data for this purpose, with an equal portion of data randomly selected from TODs serving DHH prereaders who communicate in only spoken English and from TODs serving DHH prereaders who communicate in both spoken English and some form of signed English. We then calculated inter-judge reliability to verify that the coding scheme is being applied consistently across coders. See the bottom section of Table 3 for item agreement percentages for the standardized qualifier scales.

The disagreements in coding the standardized qualifier scale for the translation category were due to a misapplication of the coding rules and the infrequency in which the translation category occurred in the data of TODs serving DHH prereaders who communicate in both spoken and some form of signed English. The research team met to further clarify how to code the standardized qualifier scale for the translation category and the pairs of coders practiced recoding the standardized qualifier for the translation category in the first 40% of data until consensus was reached. The remaining coding disagreements were minor and were, again, resolved through group discussion. After recoding the affected passages, the pairs of coders applied the revised standardized qualifier scales to the remaining 40% of the data. Item agreement between coding pairs for the remaining 40% of data showed consistency in application of the list of standardized qualifier scales. Again, differences in applying the decision rules and coding scheme were minor and resolved through group discussion.

To identify the most important book characteristics and standardized qualifier scales, we narrowed the list of 11 book characteristics to those mentioned by more than half of the TODs in each group of TODs about more than half of the storybooks. This final reduction of the data resulted in five key book

characteristics and nine standardized qualifier scales that contained a total of 36 individual qualifiers. Table 5 in the results' section shows the two-tiered glossary of book characteristics and standardized qualifier scales created through this process. To determine whether TODs serving DHH prereaders communicating in only spoken English and TODs serving DHH prereaders communicating in both spoken English and some form of signed English significantly differed in their mention of each category in the interviews, we conducted *post hoc* chi square tests with Yates' continuity correction using the *chisq.test* function in the native *stats* package in R version 3.5.1 (R Core Team, 2018).

Second question: How did TODs sort the storybooks?

To answer our second research question, we analysed the stacks each TOD created during the storybook sorting task using DiSTATIS – a generalization of metric multidimensional scaling particularly suited to sorting tasks (Abdi, Valentin, Chollet, & Chrea, 2007, 2012) – designed for the analysis of multiple distance tables collected on the same set of stimuli (Abdi & Valentin, 2007; Abdi & Williams, 2010; Abdi et al., 2012). Each TOD ranked the 14 storybooks into as many stacks as needed to capture the gradient of difficulty of the storybooks when read aloud to DHH children. We treated each TOD's stacks of storybooks as ranked data within a unique difficulty-level scale and numbered the stacks from 1 (for the easiest stack) to the greatest number of the stacks created by each participant (for the hardest stack). With this convention, the (Euclidean) distance between two books is the absolute value of the difference between the numbers assigned to their respective stacks; as a consequence, books in the same stack are at a zero distance from each other and all books in a stack have the same distance to the books in another stack. For each TOD, the distances of all storybooks to each other were stored in a distance matrix and together, the TODs data generated 69 distance tables (i.e. one matrix per TOD).

From these 69 distance tables, DiSTATIS generates two maps. The first map – called the participants' map – describes the pattern of similarity between the participants. In this map, two TODs plotted next to each other ranked the books in a similar way while TODs far from each other ranked the books in a different way. To create the second map – called the stimulus map – DiSTATIS integrates the 69 distance tables into a single data table (i.e. distance matrix) that best describes the pattern of similarity between the books (as perceived by the TODs). To display this pattern as a map, DiSTATIS creates new (orthogonal) variables called components that capture the main sources of variance between storybooks and that give the coordinates of the books on the maps. On these maps, two books near each other were rated in a similar way whereas far away books were rated in a different way; also, in order to easily identify groups of roughly equivalent books, DiSTATIS plots a 95% confidence interval ellipsoid around each book.

Third question: How do the book characteristics and qualifiers align with each stack of storybooks when they are superimposed onto the map of storybooks?

To answer our third research question, we used DiSTATIS to overlay (i.e. project) the book characteristics and qualifier scales (Question 1) onto the stimulus map that plotted the sorted storybooks (Question 2). These new maps show the book characteristics and qualifiers collectively used by the TODs when they explain how they sorted the storybooks.

Fourth question: What are the overall levels of difficulty TODs identified, and which storybooks best represent each level of difficulty?

To identify the overall levels of difficulty created by the TODs and to identify exemplar books for each level of difficulty, we followed the same process used by Schwarz et al. (2015) in the creation of a book selection system for speech-language pathologists (SLPs). We used the number of stacks created collectively by the TODs as the basis for the difficulty scale and ordered the scale levels from easy to difficult based on TODs' descriptions of the stacks of storybooks they created. We superimposed the two-tiered coding scheme onto the map of storybooks created by DiSTATIS. We created a Likert-type scale in the form of a table by using the levels of difficulty identified by storybooks as the column headers, with the easiest level of difficulty on the far left of the table and the hardest level of difficulty on the far right of the table. Then we placed the book characteristics – the first tier of the coding scheme – as the row headers. The qualifiers for each book characteristic populated the remaining cells of the table and explained the attributes of each book characteristic for each scale level. We identified exemplar books for each scale level by applying the following rules. When two books occupied a level of difficulty, either book could serve as the exemplar book. When a level of difficulty included more than two books, the storybook with the intermediate component score was chosen as the exemplar book.

Results

TODs' preferences for the storybooks

Table 4 tabulates data on the TODs storybook preferences and familiarity for each storybook. Approximately half of our participants served DHH prereaders who communicated in only spoken English and the other half served DHH prereaders who communicated in both spoken English and a form of signed English. Storybooks preferred by less than 50% of either group of TODs were removed from the analysis so that both perspectives were represented in the book selection system we extracted from the sorting and interview data. *Abuela's Weave* and *Ugly Duckling* were removed because less than 50% of both groups of TODs preferred them. Although 63.63% of the TODs serving DHH prereaders who communicated in only spoken English preferred *June 29, 1999,* the

Table 4. Storybook preferences and familiarity by book type.

Book Title	Lexile Scores as of February 2016	Preferences for Storybooks	Familiarity with Storybooks
Books with Only Animals as Characters			
Wemberly Worried	AD170L	91.94%	55.88%
Franklin Has a Sleepover	380L	81.82%	70.59%
Three Little Pigs	AD510L	75.00%	79.41%
Friends	AD670L	72.13%	26.47%
Books with People as Characters But No Multicultural Characters or Themes			
Dreams	AD60L	68.25%	19.12%
The Ugly Duckling	AD520L	37.31%	75.00%
Where the Wild Things Are	AD740L	88.06%	98.53%
June 29,1999	AD750L	53.97%	13.24%
Little Red Riding Hood	AD840L	65.15%	77.94%
Books with People as Characters and with Multicultural Characters or Themes			
Dear Juno	AD390L	79.37%	16.16%
Pablo's Tree	410L	71.88%	10.45%
Stone Soup	480L	74.24%	76.47%
Hooray! A Piñata	500L	74.60%	16.18%
Abuela's Weave	AD960L	39.68%	11.76%

Note: AD = A code that stands for 'adult-directed'.

Table 5. Two-tiered glossary of book characteristics and standardized qualifier scales extracted from the content analysis with chi square test results for each category.

Category	Standardized Qualifier Scales	% of Books	Number of TODs Mentioning Categories Who Serve DHH Prereaders Communicating Using: Only Spoken English ($n = 34$)	Both Spoken & signed English ($n = 35$)	$\chi^2(1)$[a]	*p*-value
Plot/Sequence (PS)	*Familiarity scale:* Familiar (PS.fam), unfamiliar (PS.unfam) *Complexity scale:* Simple (PS.s), little hard (PS.lh), harder (PS.harder), complex (PS.com) *Predictability scale:* Very predictable (PS.vp), predictable (PS.p), less predictable (PS.lp), unpredictable (PS.up)	100.00%	32	31	0.152	.696
Vocabulary (VO)	*Difficulty-level scale:* low (VO.l), little high (VO.lh), higher (VO.hr), high (VO.h), too high (VO.th)	100.00%	22	25	0.116	.733
Concept (CO)	*Abstractness scale:* concrete (CO.concrete), little abstract (CO.little), abstract (CO.ab), more abstract (CO.m), very abstract (CO.vab), too abstract (CO.tab)	100.00%	28	34	2.675	.102
Background Knowledge (BK)	*Relatability scale:* Very relatable (BK.vr), relatable (BK.r), maybe relatable (BK.maybe), less relatable (BK.lr), unrelatable (BK.ur)	100.00%	29	33	0.702	.402
Illustrations (IL)	*Support scale*: supportive (IL.sup), less supportive (IL.lsup), unsupportive (IL.unsup) *Detailedness scale*: simple (IL.simple), detailed (IL.d), very detailed (IL.vd), too detailed (IL.td) *Realism scale:* realistic (IL.real), less realistic (IL.lessr), unrealistic (IL.unr),	100.00%	29	33	0.702	.402

Note: [a]*Post hoc* chi square tests with Yates' continuity correction using the *chisq.test* function in the native *stats* package in R version 3.5.1 (R Core Team, 2018).

storybook was removed from the analysis because only 41.94% of TODs serving DHH prereaders who communicated in both spoken English and a form of signed English preferred the storybook. TODs were least familiar with the majority of storybooks containing multicultural characters and/or themes.

First question: What terms do TODs use to describe their ranked stacks of books?

Table 5 displays the two-tiered glossary of book characteristics and qualifiers extracted through content analysis. The post hoc *chi square tests* indicate no significant differences in the number of TODs by group who mentioned each category.

Second and third questions: How did TODs sort the storybooks and how do the terms explain the TODs' sorting decisions?

The participants' map (from DiSTATIS, see Methods), shown in Figure 1, indicated that TODs were clustered together regardless of whether they served DHH prereaders communicating in only spoken English or DHH prereaders communicating in both spoken English and some form of signed English. Figure 1 also indicates that none of the TODs was an outlier. Therefore, the picture given by the stimuli map shown in Figure 2 reflects a consensus common to both groups of TODs.

The smallest number of stacks created by a TOD was 2 and the largest number of stacks was 10. DiSTATIS indicated two important patterns of similarity in how the TODs sorted the storybooks. These two patterns were unrelated to each other (i.e. orthogonal) and explained 56% of the total variance in the data. As discussed in the Methods section, these two patterns of similarity are called components and are represented in Figure 2 by each axis. The horizontal axis represents the first component and explains 45% of the variance. The vertical axis represents the second component and explains 11% of the variance. Taken together, these two components explain 56% of the variance of the data.

In Figure 2, dots represent storybooks and the ellipsoids surrounding each dot are 95% confidence intervals (CIs) computed by bootstrap resampling (see Abdi, Williams, Valentin, & Bennani-Dosse, 2012, for details). Storybooks plotted near each other were consistently sorted together by the TODs. When CIs do not overlap, individual storybooks and groups of storybooks (created by overlapping CIs) are considered significantly different from one another at $p < .05$. Recall from Table 4 that the TODs were generally unfamiliar with *Wem, Friends, Dreams, Dear, Pablo, and Piñata*. If the TODs' sorting decisions had been influenced by their lack of familiarity with these storybooks, DiSTATIS would have plotted these storybooks in a similar location in Figure 2. However, Figure 2 shows these storybooks spread across three of the four quadrants of the map. Therefore, the sorting

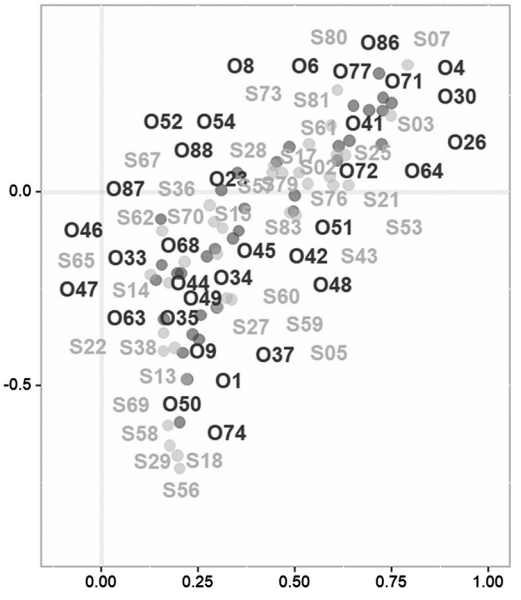

Figure 1. Participants' map showing no group differences based on modality (i.e. spoken language only, both spoken [O#] and signed English [S#]).

decisions of the TODs in our study do not appear to be influenced by their familiarity with the storybooks in the current study.

During the procedure, we asked the TODs to sort the storybooks based on how difficult they thought the books would be for DHH prereaders to

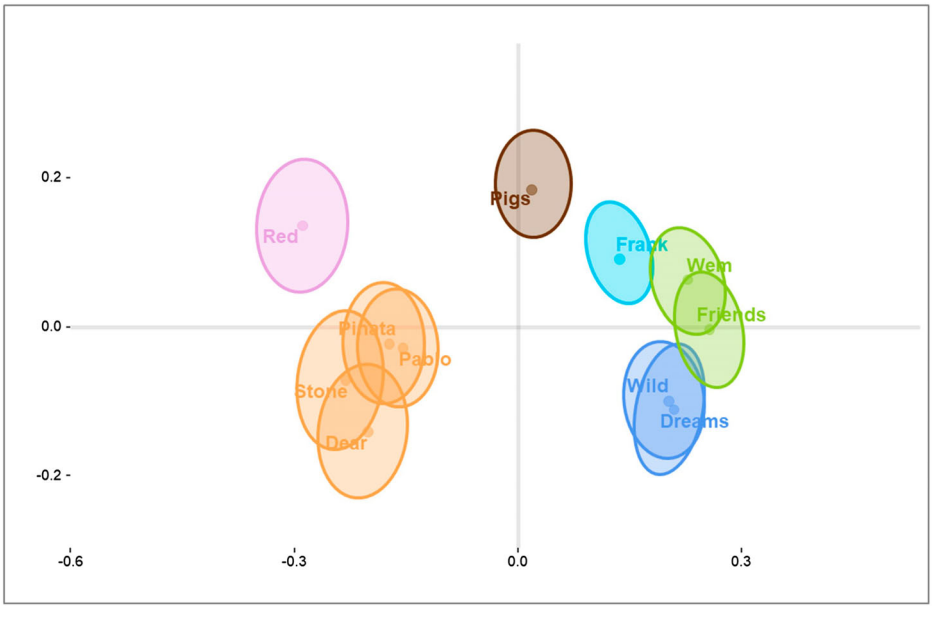

Figure 2. Combined two-dimensional map of how the TODs collectively sorted the storybooks, with storybooks (dots) and 95% confidence intervals.

understand when the books were read aloud to them. So, we treated these individual storybooks and groups of storybooks as separate ranks of difficulty in how TODs evaluated the 14 storybooks. The map's centre (where the two axes cross) corresponds to component scores of 0. Items near (or at) the map's centre do not help explain major patterns of similarity in the data.

TODs using English created six groups of storybooks: (a) *Friends,* and *Wem,* (b) *Frank,* (c) *Dreams* and *Wild*, (d) *Pigs,* (e) *Stone, Dear, Piñata,* and *Pablo,* and (f) *Red*. To understand why TODs collectively sorted the storybooks into these six groups, we incorporated our two-tiered glossary of book characteristics (first tier) and qualifiers (second tier) into our DiSTATIS analysis, which we discuss next.

DiSTATIS computed supplementary component scores for each qualifier associated with each book characteristic (e.g. BL [book length]: too long) and then overlaid (i.e. projected) these scores onto the map of storybooks, as shown in Figure 3. Note that the scores for each qualifier simply indicate which items the TODs frequently associated with particular storybooks and groups of storybooks. For a technical description of this procedure, see Appendix D in Schwarz et al. (2015) or Lahne, Abdi, and Heymann (2018).

Note that in Figure 3, the book characteristic is in capital letters and precedes the qualifier (e.g. CON.concrete = indicates books that include concrete concepts). See Table 5 for the key to the abbreviations.

The book characteristics and qualifiers are colour-coded. The book characteristics and qualifiers in grey in Figure 3 were not associated with one of the six groups of storybooks plotted in Figure 2 so are not important in our analysis. The remaining book characteristics and qualifiers are colour-coded

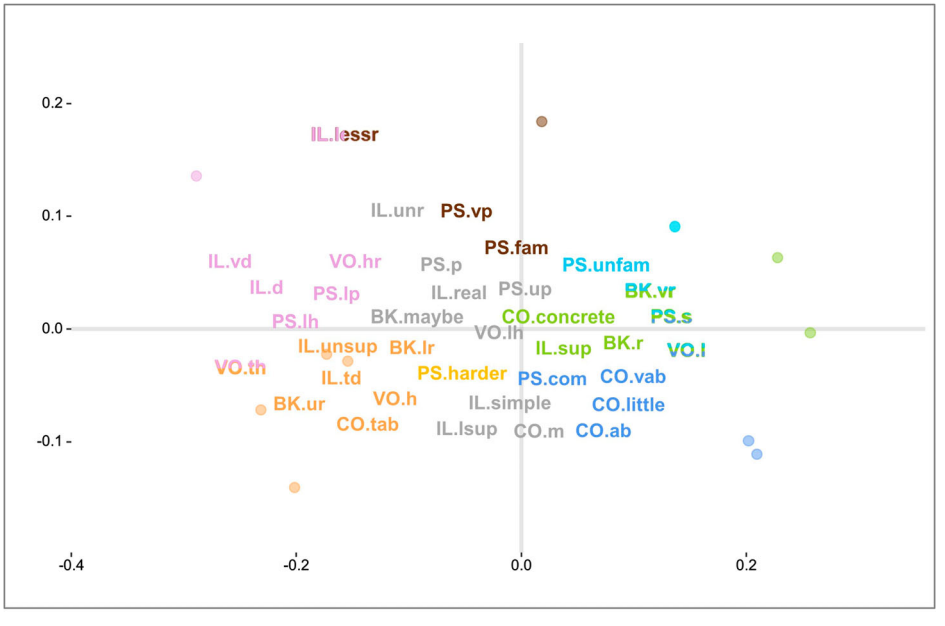

Figure 3. Combined two-dimensional map of storybooks and book characteristics with qualifiers.

based on their proximity on the map to the six groups of storybooks plotted in Figure 2. TODs described books on the far-left side of the horizontal axis shown in Figure 3 as having (a) plot/sequences that were harder or a little hard and less predictable, (b) vocabulary that was high, higher, or too high, and (c) as being either less relatable or unrelatable to the background knowledge of DHH-English prereaders. TODs described books on the far-right side of the map as having low vocabulary, simple plot/sequences, and as being relatable and very relatable to the background knowledge of DHH-English prereaders.

The second component (vertical axis) contrasts fairy tales – *Pigs and Red* – and storybooks that were conceptually more difficult – *Dear (a multicultural book), Wild,* and *Dreams*. The fairy tales plotted at the top of the map have predictable plot/sequences and have illustrations that are less realistic. The conceptually more difficult books plotted at the bottom of the map have concepts that are very abstract or too abstract.

Fourth question: What are the difficulty levels and exemplar books for each level?

We used the six groups of storybooks identified in Figure 2 as the difficulty levels of our book selection system, shown as column headers in Table 6: (a) *Friends,* and *Wem,* (b) *Frank,* (c) *Dreams* and *Wild*, (d) *Pigs,* (e) *Stone, Dear, Piñata,* and *Pablo,* (f) *Red*. Then we used the colour-coded book characteristics and qualifiers that were associated with each group of storybooks shown in Figure 3 to explain the characteristics of each level of the difficulty scale. The book characteristics are shown as row labels in Table 6 and the qualifiers are shown in the table cells. We ordered the book characteristics in Table 6 based on the level of detail provided for each level of difficulty by the TODs. For example, the plot/sequence book characteristic is shown as the first book characteristic in Table 6 because a separate qualifier is associated with each of the six difficulty levels. The illustrations book characteristic is shown as the last book characteristic in Table 6 because a separate qualifier is associated with only four of the six difficulty levels. For the difficulty levels with only one or two storybooks, those storybooks were used as exemplar books for the associated level of difficulty. However, the fifth difficulty level included the following four storybooks: *Stone, Dear, Piñata,* and *Pablo.* We chose *Pablo* as the exemplar book for difficulty level five because it had the median component score for the group of storybooks.

Note that TODs sorted storybooks with animals as characters at the easier end of the difficulty level scale and storybooks with people only or people and animals in the three most difficult levels. Also, TODs sorted all of the storybooks with multicultural characters and themes in the second most difficult level of the scale. After completing the storybook sorting task, many TODs stated how much they liked and needed books with multicultural characters and themes because they serve DHH children with culturally diverse backgrounds. However, they

Table 6. Book selection system for TODs serving DHH pre-k and kindergarteners who use English to communicate.

Scale & Exemplar Books	1 *Friends, Wemberly Worried*	2 *Franklin Has a Sleepover*	3 *Dreams, Where the Wild Things Are*	4 *Three Little Pigs*	5 *Pablo's Tree*	6 *Little Red Riding Hood*
Plot/Sequence	Simple	Simple & unfamiliar	Complex	Familiar & very predictable	Harder	Little hard & less predictable
Vocabulary	Low	Low	Low		High, too high	Higher, too high
Concepts	Concrete		Little abstract, abstract, very abstract		Too hard	
Background Knowledge	Relatable & Very Relatable	Very relatable			Less relatable, unrelatale	
Illustrations	Supportive			Less realistic	Unsupportive & too detailed	Very detailed, too detailed, & less realistic

wished that the storybooks with multicultural characters and themes included simpler plots/sequences, lower vocabulary, and more supportive illustrations.

Summary

Of the 14 storybooks sorted by the TODs, three storybooks – *Abuela's Weave, June 29,1999, and Ugly Duckling* – were excluded from the analysis because 50% or more of the TODs did not like these storybooks. For the remaining books, the TODs' explanations for sorting the storybooks were used in a content analysis to extract a two-tiered glossary that included five key book characteristics and nine qualifier scales with a total of 36 qualifiers. The DiSTATIS analysis identified six groups of storybooks. To describe these six groups and the rationale for their ranking, we used DiSTATIS to superimpose the qualifier scales onto the six groups of storybooks. We then created a Likert-type scale by making the six difficulty levels the column headers with the easiest level on the far left of the table and the hardest level on the far right of the table. We made the five key book characteristics the row headers and populated the remaining table cells with the qualifiers that the DiSTATIS analysis indicated were associated with each group of storybooks. We then selected exemplar books for each level of the difficulty scale to complete the book selection system.

Discussion

Very little attention has been paid to how TODs select storybooks for read alouds. To implement evidence-based read-aloud interventions focused on building oral language skills, TODs need to be able to select storybooks based on different difficulty levels. Our purpose was to identify the criteria TODs use to select storybooks for read alouds when their goal is to increase the oral language skills of the DHH-English prereaders they serve. By analysing data from a storybook sorting task using content analysis and a multivariate technique, we extracted a book selection system that includes a two-tiered glossary and a six-point overall difficulty scale with exemplar books for each scale level. Next, we discuss how our system compares to the Signability Index created by Stewart et al. (1992). Then we discuss how our system aligns with the levels of difficulty required by the three evidence-based read-aloud interventions that target the increase of oral language skills and present guidelines for using our system to select storybooks for each type of read-aloud intervention. Finally, we discuss the study's limitations and future research directions.

How our results compare to Stewart et al.'s (1992) Signability Index

Until the current study, the only book selection system designed for TODs serving DHH-English was Stewart et al.'s (1992) Signability Index. The Signability

Index organized books across literary genres for a variety of objectives for DHH children in preschool through second grade who were educated in total communication classrooms that incorporated signed English with ASL features. The purpose of our system is to help TODs select storybooks for read alouds meant to increase the English language skills of DHH-English preschoolers and kindergarteners. Our system also is easy to align with the book selection requirements of the three evidence-based read-aloud interventions for increasing oral language skills in prereaders. Our book selection system is based on teachers' judgments and includes mutually exclusive book characteristics. Finally, in order to reflect the increasingly culturally-diverse DHH population, our system includes a larger percentage of storybooks with multicultural characters and themes than the Signability index.

How our results can help TODs select read alouds for evidence-based interventions

The three evidence-based read-aloud interventions include a dialogic reading style (Zevenbergen & Whitehurst, 2003), a literal-inferential reading style (van Kleeck et al., 2006), and a TextTalk reading style (Beck & McKeown, 2001, 2007). Dialogic reading promotes conversational vocabulary and grammar (Dirks & Wauters, 2015, 2018; Fung et al., 2005; Huebner, 2006; Lederberg et al., 2014; Mol et al., 2008; Trussell & Easterbrooks, 2014; Zevenbergen & Whitehurst, 2003) and requires books with illustrations that can both introduce new vocabulary and tell the complete story without relying on information from the text (Whitehurst et al., 1994).

Storybooks that align with the first two levels of our book selection system are particularly appropriate for a dialogic reading intervention. These two levels include books with supportive illustrations, simple and very predictable plot/sequences, mostly concrete concepts, and relate to the background knowledge of DHH-English prereaders. When using the first two gradations of difficulty in our book selection system to select storybooks appropriate for a dialogic reading intervention, we suggest that TODs reread several times the exemplar books *Friends, Wemberly Worried, and Franklin Has a Sleepover*. As TODs compare new storybooks to these exemplar books, we suggest TODs ask themselves the following questions that are based on the key book characteristics and qualifier scales mentioned by the TODs in our study:

- Is the plot relatively simple with a clear sequence of events that focuses on concrete concepts?
- Will my students understand the major plot elements and sequence by relying mainly on the illustrations and their background knowledge?
- Does the storybook text contain mainly vocabulary that the children are exposed to in their everyday lives?

When TODs answer these three questions affirmatively and the storybooks seem similar to the three exemplar books *Friends, Wemberly Worried,* and *Franklin Has a Sleepover,* TODs will have identified storybooks appropriate for a dialogic intervention. When conducting a dialogic reading intervention, TODs should focus on the illustrations because they include major plot elements and often include pictures of concrete concepts that are important to story comprehension. TODs can encourage DHH prereaders to contribute more to the conversation by repeating what the children say and by responding positively to all of their contributions. Based on where the DHH prereaders lead the conversation, TODs can ask the children to label objects, locations, and characters shown in the illustrations and ask children what is happening in the illustrations. TODs should look for opportunities to expand the children's understanding of the story and should encourage the children to retell parts or all of stories during these dialogic reading sessions.

The literal-inferential style of read aloud promotes both the conversational and written language registers (van Kleeck, 2008, 2014; van Kleeck et al., 2006). When transitioning DHH children from a dialogic reading to a literal-inferential read-aloud style, we suggest TODs initially use books at the second and third levels of our book selection system. When using the second and third gradations of difficulty in our book selection system to select storybooks appropriate for a literal-inferential read-aloud intervention, we suggest that TODs reread several times the exemplar books *Franklin Has a Sleepover, Dreams,* and *Where the Wild Things Are*. As TODs compare new storybooks to these exemplar books, we suggest TODs ask themselves the following questions that are based on the book characteristics and qualifier scales mentioned by our participants:

- Will my students have to make one or more inferences to have a complete understanding of the storybooks?
- Do the storybooks contain one or more abstract concepts?
- Will my students understand the plot elements/sequence and many of the concepts by using their background knowledge?
- Does the storybook text contain mainly vocabulary that the children are exposed to in their everyday lives?

When TODs answer these four questions affirmatively and the storybooks seem similar to the three exemplar books *Franklin Has a Sleepover, Dreams,* and *Where the Wild Things Are,* TODs will have identified storybooks appropriate for a literal-inferential read aloud intervention. When preparing for a literal-inferential read aloud, TODs need to determine on which pages the illustrations and text provide all the information that the children need to understand the stories. TODs then need to script literal questions that they can ask DHH prereaders when they find themselves on these pages during the read aloud. These literal questions should comprise between 60% and 70% of the questions TODs

pose to DHH prereaders (van Kleeck, 2014; van Kleeck et al., 2006). The literal questions used in the literal-inferential read aloud style are similar to the questions used in the dialogic reading intervention discussed previously and include asking children to label objects, characters, and locations shown in the illustrations, and to describe scenes in the illustrations (van Kleeck et al., 2006). Then TODs need to determine on which pages the illustrations and text do not provide key information that is important for the children to infer in order to have a deep understanding of the story or to see the story as relevant to their own lives. Next, TODs need to script inferential questions focusing on (a) connections between the storybook and the children's lives, (b) any cause-and-effect concepts in the story, (c) character attitudes, perspectives, and feelings, and (d) connections between the storybook and the children's lives. Only between 30% and 40% of the questions TODs ask during a literal-inferential read aloud session should focus on inferential material (van Kleeck, 2014; van Kleeck et al., 2006). As DHH children gain success in answering inferential questions and engaging in inferential discussions, we suggest TODs gradually increase the difficulty level of storybooks that are used in read alouds. TODs can familiarize themselves with the exemplar storybooks *Pablo's Tree* and *Little Red Ridinghood* in the fifth and sixth levels of difficulty in our book selection system and ask themselves the same four questions above to identify more challenging storybooks for a literal-inferential read intervention.

The TextTalk style of read aloud helps children make meaning based on the storybook text, emphasizing sophisticated vocabulary. Hallmarks of a TextTalk read aloud intervention are that the adult only shows the illustrations after giving the children an opportunity to construct story meaning based solely on the storybook text and that the adult selects three high-level vocabulary words per book and teaches these words after the read aloud. We suggest transitioning DHH-English prereaders into a TextTalk read-aloud intervention using storybooks that align with the third and fourth difficulty levels of our storybook selection system. When selecting storybooks that align with the third and fourth gradations of difficulty, we suggest that TODs reread several times the exemplar books *Dreams, Where the Wild Things Are,* and *Three Little Pigs*. As TODs compare new storybooks to these exemplar books, we suggest TODs ask themselves the following questions that are based on the book characteristics and qualifier scales mentioned by our participants:

- Does the storybook text tell the complete story and the illustrations convey only parts of the story?
- Does the storybook text contain three vocabulary words that are important to the story, are associated with academic/written vocabulary, and will likely appear in other content areas? Beck and McKeown (2001) refer to these words as tier 2 vocabulary words. Some examples of tier 2 vocabulary words from the exemplar books in the third and fourth difficulty levels of

our book selection system are *dream, city, world, cast, gather, shadow, mischief, tame, private,* and *lonely*.

Because the TextTalk read-aloud style requires storybooks that are intellectually challenging and do not rely too much on the illustrations to relay the story content (Beck & McKeown, 2001, 2007), DHH-English prereaders need to have had several successful read-aloud experiences during dialogic and literal-inferential read-aloud interventions before attempting a TextTalk read aloud intervention. Once DHH-English prereaders have been successful at explaining what is happening in sections of text without first seeing the illustrations across several storybooks, we suggest selecting storybooks for TextTalk read alouds that align with the fifth and sixth levels of difficulty shown in Table 6.

Study limitations

Three potential limitations of this study concern validation of the book selection system and the booklists we used. In this study, we created the book selection system but did not validate it. Also, the system is based on storybooks from several published preschool curricula instead of lists of TODs' favourite storybooks for read alouds. We adopted the former approach to ensure that the storybooks we chose represented the range of difficulty in existing preschool curricula for all prereaders, both hearing and DHH. We accounted for the book preferences of TODs in the current study by only including storybooks preferred by more than half of the TODs. This decision ensured that we maintained high expectations for DHH prereaders – a critical factor in the academic success of DHH students (see, e.g. Luckner, 2011).

Future studies

Results from the current study suggest the need for three separate series of studies. First, a series of studies need to determine whether the current system can be validated. We plan to have out-of-field undergraduates, student teachers, and parents use the book selection system in Table 6 to rank the storybooks used in this study and new storybooks to see whether the system helps them approximate the judgment of the highly-skilled TODs who sorted the storybooks in the current study. If the system helps these three groups approximate the judgment of highly-skilled TODs, we will know that the expert judgment captured in our analysis is transferable to people with a wide range of knowledge about DHH children and book selection for read alouds. Second, a series of studies are also needed to determine how choosing books at different levels of difficulty within the three evidence-based read-aloud interventions affect the quality of talk surrounding the read-aloud interaction and the overall success of the read-aloud activity. Finally, it would be interesting to collect

from TODs booklists of their favourite storybooks for read alouds and compare them to the storybooks included in the current study to determine whether and how the range of difficulty between these two sets of storybooks compare.

Conclusion

Teachers struggle to select appropriate books for read alouds, either selecting books at random (Damber, 2015) or books that are too simple when the purpose is to increase children's oral language skills (Beck & McKeown, 2001; McGee & Schickendanz, 2007). The only available book selection system for TODs serving DHH-English prereaders was developed for a different purpose than identifying storybooks to increase English language skills and followed a methodology difficult to evaluate. We addressed the need for a book selection system by creating a system with a two-tiered glossary of terms and a six-point overall difficulty-level scale based on the judgments of 69 highly-skilled TODs from across the United States. Future studies will determine whether our book selection system can be validated and will test its effectiveness as a training tool for TODs-in-training and parents of DHH-English prereaders. Recall that the English language skills often targeted during read alouds are positively linked to reading comprehension in third grade (Language and Reading Research Consortium, 2018; Lepola et al., 2016). Therefore, it is paramount that TODs serving DHH-English prereaders use evidence-based practices to develop these English language skills. The creation of a book selection system and its pairing with evidence-based read-aloud interventions designed to increase the English language skills of prereaders is one small step toward this goal.

Acknowledgements

Meagan Jurica, Charlsa Matson, Rachel Stiller, and Taylor Webb-Culver all hold second authorship position. This research would not have been possible without the teachers of the deaf who took time from their very busy schedules to participate in this study. We also recognize the following members of the Team Schwarz Child Language Lab for helping to create various coding manuals, collect data, and process data: Paid graduate student project managers – Jennifer Guajardo, Rebecca Hart, Lauren Lytle, and Ashley Williams – and the volunteer levelling and undergraduate workforce – Lindsey Adrio, Jordan Bailey, Ashley Bentley, Carol Bittenger, Gracie Cooper, Brittany Dolezal, Dobie Drost, Alyssa Ferrari, Taylor Friend, Edna Echauri Garcia, Rayiah Gassas, Hailey Gilder, Jessica Glaser, Maxine Gonzalez, Megan Green, Vanessa Haros, Jahaira Hernandez, Ashley Keenan, Sarah McCullough, Vanessa Mercado, Clarissa Montes, Olivia Moore, Madelaine O'Neill, Kendall Precup, Julia Rodriquez, Adriana Sanchez, Rhiana Schwartz, Haleigh Straube, Kacey Thomson, Marissa Trevino, Madison Weise, and Kathryn Williamson.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Texas State University [grant number 9000001582]. Hervé Abdi's work benefited from a (2014–2015) EURIAS fellowship at the Institut d'etudes avancées de Paris (France), made possible with the support of the European Union's 7th Framework Program for Research and from funding from the French State managed by the Agence Nationale de la Recherche, programme 'Investissements d'avenir' (ANR-11-LABX-0027-01 Labex RFIEA+).

ORCID

Amy Louise Schwarz http://orcid.org/0000-0001-5883-6327

Notes on contributors

Amy Louise Schwarz, Ph.D., CCC-SLP is an assistant professor at Texas State University. She is both a licensed speech-language pathologist and a certified teacher of the deaf who worked in a total communication classroom for several years. Her primary line of research examines read-aloud practices with prereaders at the preschool language level across sub-specialties of educators in two ways. In one approach, she applies a methodology adapted from the field of sensory and perception research (e.g., wine tasting, olive oil tasting, salt tasting) to capture clinical-decision making and to transform it into evidence-based tools for new clinicians and parents. In the other approach, Dr. Schwarz compares read-aloud goals, read-aloud behaviours, and use of academic language during read alouds across types of educators (e.g., teachers of the deaf, speech-language pathologists, bilingual educators). Her secondary line of research examines how social, cultural, and linguistic biases of professionals and students in Communication Disorders affect their clinical judgments.

Meagan Jurica, B.S. holds second authorship position on this paper. She graduated with a B.S. in Communication Disorders from Texas State University in 2018. She is now pursuing an M.S. in Communication Sciences and Disorders at Baylor University.

Charlsa Matson, B.S. holds second authorship position on this paper. She graduated with a B.S. in Communication Disorders from Texas State University in 2018. She is now pursuing an M.A. in Communication Sciences and Disorders at the University of Texas at Austin.

Rachel Stiller, B.S. holds second authorship position on this paper. She graduated with a B.S. in Communication Disorders from Texas State University in 2018. She is now pursuing an M.S. in Communication Disorders at the University of Texas at Dallas.

Taylor Webb-Culver, B.S., holds second authorship position on this paper. She graduated with a B.S. in Communication Disorders from Texas State University in 2018. She is now pursuing an M.S. in Communication Disorders at Texas State University.

Hervé Abdi, Ph.D. is a full professor in the School of Behavioral and Brain Sciences at The University of Texas at Dallas and an adjunct professor of Radiology at The University of Texas Southwestern Medical Center at Dallas. His research is organized around three areas: psychology of memory, neural networks, and statistics. The psychology of memory research is mainly directed toward the modeling of long-term semantic memory (e.g., scripts, schema, concepts) and memory for faces. The neural networks research is directed at finding a generalization of auto-associative networks and to the analysis of the statistical properties of connectionist models. His work with statistics is oriented towards two domains: analysis of variance

(experimental design), analysis of qualitative data (correspondence analysis, additive tree analysis), and analysis of large datasets as found, for example, in genomics or brain imaging studies.

References

Abdi, H., & Valentin, D. (2007). Multiple correspondence analysis. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 651–657). Thousand Oaks, CA: Sage.

Abdi, H., Valentin, D., Chollet, S., & Chrea, C. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Quality and Preference*, *18*, 627–640.

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*, 433–459.

Abdi, H., Williams, L. J., Valentin, D., & Bennani-Dosse, M. (2012). STATIS and DISTATIS: Optimum multi-table principal component analysis and three way metric multidimensional scaling. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*, 124–167.

Anderson, J., Anderson, A., Shapiro, J., & Lynch, J. (2001). Fathers' and mothers' book selection preferences for their four year old children. *Reading Horizons*, *41*, 189–211.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. (1985). *Becoming a nation of readers: The report of the commission on reading*. Washington, DC: U.S. Department of Education. (pp. 1–155).

Beck, I. L., & McKeown, M. G. (2001). Text talk: Capturing the benefits of read-aloud experiences for young children. *The Reading Teacher*, *55*, 10–20.

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*, 251–271.

Blancher, G., Chollet, S., Kesteloot, R., Nguyen Hoang, D., Cuvelier, G., & Sieffermann, J. M. (2007). French and Vietnamese: How do they describe texture characteristics of the same food? A case study with jellies. *Food Quality and Preference*, *18*, 560–575.

Chollet, S., Lelièvre, M., Abdi, H., & Valentin, D. (2011). Sort and beer: Everything you wanted to know about the sorting task but did not dare to ask. *Food Quality and Preference*, *22*, 507–520.

Cohn, D., & Caumont, A. (2016). *10 demographic trends that are shaping the U.S. And the world*. Washington , DC: Pew Research Center. [online] http://www.pewresearch.org/fact-tank/2016/03/31/10-demographic-trends-that-are-shaping-the-u-s-and-the-world/

Damber, U. (2015). Read-alouds in preschool – a matter of discipline. *Journal of Early Childhood Literacy*, *15*, 256–280.

Dirks, E., & Wauters, L. (2015). Enhancing emergent literacy in preschool deaf and hard-of-hearing children through interactive reading. In H. Knoors & M. Marschark (Eds.), *Educating deaf learners: Creating a global evidence base* (pp. 415–441). New York, NY: Oxford University Press.

Dirks, E., & Wauters, L. (2018). It takes two to read: Interactive reading with young deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, *23*, 261–270.

Elster, C. A. (1998). Influences of text and pictures on shared and emergent readings. *Research in the Teaching of English*, *32*, 43–78.

Fung, P., Chow, B., & McBride-Chang, C. (2005). The impact of a dialogic reading program on deaf and hard-of-hearing kindergarten and early primary school-aged students in Hong Kong. *Journal of Deaf Studies and Deaf Education*, *10*, 82–95.

Gallaudet Research Institute. (2013). *Regional and national summary report of data from the 2011–12 annual survey of deaf and hard of hearing children and youth*. Washington, DC: Gallaudet University.

Griffin, L. (1970). *Books in preschool: A guide to selecting, purchasing, and using children's books*. Washington, DC: National Center for Educational Research and Development.

Hayes, P. L., & Shaw, P. C. (1994). Guidelines for selecting read aloud books for deaf children as suggested by storytellers using American sign language. In M. P. Moeller & B. Schick (Eds.), *Deafness and diversity: Sociolinguistic issues* (pp. 105–116). Omah, NE: Boys Town National Research Hospital.

Huebner, C. (2006). Optimizing the effects of shared reading on early language skills. In A. Van Kleeck (Ed.), *Sharing books and stories to promote language and literacy*. (pp. 149–178). San Diego, CA: Plural Publishing.

International Literacy Association. (2018). *Standards 2010: Standard 4*. Newark, DE: International Literacy Association. [online] https://www.literacyworldwide.org/get-resources/standards/standards-for-reading-professionals/standards-2010-standard-4

Lahne, J., Abdi, H., & Heymann, H. (2018). Rapid sensory profiles with DISTATIS and barycentric text projection: An example with Amari, bitter herbal liqueurs. *Food Quality and Preference*, *66*, 36–43.

Language and Reading Research Consortium. (2018). The simple view of reading across development: The prediction of grade 3 reading comprehension by prekindergarten skills. *Remedial and Special Education*, *39*, 289–303.

Laurent Clerc National Deaf Education Center. (2015). *Schools and programs for the deaf and hard of hearing students in the U.S.*. Gallaudet University. [online] http://www3.gallaudet.edu/clerc-center/info-to-go/national-resources-and-directories/schools-and-programs.html

Lederberg, A. R., Miller, E. M., Eastabrooks, S. R., & Mcdonald Connor, C. (2014). Foundations for literacy: An early literacy intervention for deaf and hard-of-hearing children. *Journal of Deaf studies and Deaf Education*, *19*, 438–455.

Lepola, J., Lynch, J., Kiuru, N., Laakkonen, E., & Niemi, P. (2016). Early oral language comprehension, task orientation, and foundational reading skills as predictors of grade 3 reading comprehension. *Reading Research Quarterly*, *51*, 373–390.

Luckner, J. (2011). Promoting resilience: Suggestions for families, professionals, and students. In D. H. Zand & K. J. Pierce (Eds.), *Resilience in deaf children: Adaptation through emerging adulthood* (pp. 207–226). Berlin: Springer Science+Business Media, LLC.

Lynch-Brown, C., & Tomlinson, C. M. (2008). *Essentials of children's literature* (6th ed). Boston, MA: Allyn and Bacon.

Martinez, M., & Roser, N. (1985). A case study of the effects of selected text factors on parent-child story time interactions. In J. A. Niles & R. V. Lalik (Eds.), *Issues in literacy: A research perspective, thirty-fourth yearbook of the national reading conference* (pp. 168–174). Chicago, IL: National Reading Conference.

McGee, L. M., & Schickendanz, J. A. (2007). Repeated interactive read-alouds in preschool and kindergarten. *The reading Teacher*, *60*, 742–751.

Metametrics. (2014). *Text complexity grade bands and lexile bands* [online]. Retrieved from https://lexile.com/using-lexile/lexile-measures-and-the-ccssi/text-complexity-grade-bands-and-lexile-ranges/

Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development*, *19*, 7–26.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2012). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*. Washington, DC.

Nolan, J. (Producer). (2015). The 2014 AS&U 100. *American School and University*. Retrieved from http://www.asumag.com/research/2014-asu-100.

Patris, B., Gufoni, V., Chollet, S., & Valentin, D. (2007). *Impact of training on strategies to realize a beer sorting task: behavioral and verbal assessments*. Paper presented at the SPISE, Ho Chi Minh, Vietnam.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Santosa, M., Abdi, H., & Guinard, J. X. (2010). A modified sorting task to investigate consumer perceptions of extra virgin olive oils. *Food Quality and Preference*, *21*, 881–892.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In D. K. Dickinson (Ed.), *Handbook of early literacy research* (pp. 97–110). New York: Guilford Press.

Schreier, M. (2012a). Building a coding frame. In *Qualitative content analysis in practice* (pp. 80–106). Washington, DC: SAGE Publications.

Schreier, M. (2012b). Strategies for building a data-driven coding frame. In *Qualitative content analysis in practice* (pp. 107–125). Washington, DC: SAGE Publications.

Schreier, M. (2012c). Trying it out: The pilot phase. In *Qualitative content analysis in practice* (pp. 146–165). Washington, DC: SAGE Publications.

Schwarz, A. L., Guajardo, J., & Hart, R. (2017). How do communication modes of deaf and hard-of-hearing prereaders influence teachers' read-aloud goals? *Deafness and Education International*, *19*, 115–125.

Schwarz, A. L., Guajardo, J., & Hart, R. (2018). How do communication modes of d/deaf and hard-of-hearing prereaders influence the amount of storybook text teachers report reading during read alouds? *Deafness and Education International*, 1–30.

Schwarz, A. L., Van Kleeck, A., Beaton, D., Horne, E., McKenzie, H., & Abdi, H. (2015). A read-aloud storybook selection system for pre-reading preschoolers. *Journal of Speech, Language, and Hearing Research*, *58*, 1273–1291.

Stewart, D. A., Bennett, D., & Bonkowski, N. (1992). Books to read, books to sign. *Perspectives in Education and Deafness*, *10*, 4–7.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a logitudinal structural model. *Developmental Psychology*, *38*, 934–947.

Tang, C., & Heymann, H. (2002). Multidimensional sorting, similarity scaling and free-choice profiling of grape jellies. *Journal of Sensory Studies*, *17*, 493–509.

Trussell, J. W., & Easterbrooks, S. R. (2014). The effect of enhanced storybook interaction on signing deaf children's vocabulary. *Journal of Deaf studies and Deaf Education*, *19*, 319–332.

Ukrainetz, T. A., Cooney, M. H., Dyer, S. K., Kysar, A. J., & Harris, T. J. (2000). An investigation into teaching phonemic awareness through shared reading and writing. *Early Childhood Research Quarterly*, *15*, 331–355.

van Kleeck, A. (2008). Providing preschool foundations for later reading comprehension: The importance of an ideas for targeting inferencing in storybook-sharing interventions. *Psychology in the Schools*, *45*, 627–643.

van Kleeck, A. (2014). Distinguishing between casual talk and academic talk beginning in the preschool years: An important consideration for speech-language pathologists. *American Journal of Speech-Language Pathology*, *23*, 724–741.

van Kleeck, A., Vander Woude, J., & Hammett, L. (2006). Fostering literal and inferential language skills in Head Start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology*, *15*, 85–95.

Weber, R. P. (1990). *Basic content analysis* (2nd ed.). London: Sage Publications.

Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in head start. *Journal of Educational Psychology*, *86*, 542–555.

Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading: A shared picture book reading. In A. Van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 177–200). New York, NY: Routledge.

Zucker, T. A., Ward, A. E., & Justice, L. M. (2009). Print referencingg during read alouds: A technique for increasing emergent readers' print knowledge. *The reading Teacher*, *63*, 62–72.

Appendix. References and curricula for the 14 storybooks.

	Abbreviations	Lexile Scores[c]	Preschool Curricula[a,b] The Creative Curriculum	Big Scholastic Day	High Scope	Houghton Mifflin	The Literacy Express	Opening the World of Learning	Read, Play, & Learn
No. of U.S Private & State Schools for the Deaf that Use the Curricula			9	0	3	0	0	0	1
No. of U.S. Public School Districts that Use the Curricula			11	2	5	1	1	5	0
Castaneda, O.S. (1993). *Abuela's Weave.* New York, NY: Lee and Low Books, Inc.	Abuela	AD960L	√						
Pak, S. (1999) *Dear Juno.* New York, NY: Penguin Group.	Dear	AD390L	√						
Keats, E. J. (2000) *Dreams.* New York, NY: Puffin Books.	Dreams	AD60L	√						
Cauley, L. (1979). *The Ugly Duckling.* New York, NY: Harcourt Brace Jovanovich, Publishers.	Duck	AD520L				√			
Bourgeois, P. (2014). *Franklin Has a Sleepover.* Toronto, ON: Kids Can.	Frank	380L							√
Heine, H. (1982). *Friends*	Friends	AD670L							√
Kleven, E. (2000). *Hooray, A Piñata!* New York, NY: Puffin Books	Pinata	500L	√					√	
Wiesner, D. (1992). *June 29,* 1999.New York, NY: Clarion Books.	June29	AD750L					√		
Pinkney, J. (2007). *Little Red Riding Hood.* New York, NY: Little Brown& Co.	Red	AD840L	√						
Mora, P. (1994). *Pablo's Tree.* New York NY: Simon & Schuster Books for Young Readers.	Pablo	410L	√						
Muth, J. J. (2003). *Stone Soup.* New York, NY: Scholastic Press.	Stone	480L				√			
Zemach, M. (1998). *The Three Little Pigs.* New York, NY: Farrar Straus.	Pigs	AD510L	√	√	√	√			
Henkes, K. (2000). *Wemberly Worried.* New York, NY: HarperCollins.	Wem	AD170L	√						
Sendak, M. (1963). *Where the Wild Things Are.* New York, NY: HarperCollins.	Wild	AD740L					√		

[a]Some of the storybooks used in the study were listed by title only on the read aloud lists from some curricula so a different version of the storybook may be used in the curricula than the version used in the current study; [b]For the DLM curriculum, only five of the 30 storybooks were available in the university library of the first author. None of these five books was quasi randomly selected for this study; [c]These lexile scores were reported on the MetaMetrics website as of February 2016.