Hammill Institute on Disabilities

Communication Disorders Quarterly
2020, Vol. 41(2) 86–99

DOI: 10.1177/1525740118795158
cdq.sagepub.com
SAGE

Collaborating With an Expert Panel to Establish the Content Validity of an Intervention for Preschoolers With Language Impairment

Megan Schneider Dinnesen, EdD[1], Arnold Olszewski, PhD[2], Allison Breit-Smith, PhD[1], and Ying Guo, PhD[1]

Abstract

Expanding on the extant body of research on content validity, this study applied the tenets of content validity research to the development of an expository book reading intervention focused on science for preschool students with language impairment. This example case explains how, guided by content validity research in healthcare interventions, collaboration with an expert panel was systematically approached in the development of an educational intervention. Specifics as to how feedback was gathered and analyzed are discussed. Of practical importance for practitioners and researchers, this study provides a detailed description of how the knowledge and experience of a panel of experts was captured, analyzed, and incorporated into the development process of an educational intervention to validate its application.

Keywords

intervention strategies, language learning disorders, exceptionalities, 3 to 5 years age, evidence-based practices

Content validity is generally defined as the extent to which components of a given assessment or program represent the construct they are intended to address (Haynes, Richard, & Kubany, 1995). Evidence of content validity can be captured using qualitative and/or quantitative methods and is typically based on the judgment of individuals deemed to be experts in the field of study being targeted. Content validity studies are most often undertaken to evaluate assessment instruments prior to their implementation with their intended audience. However, content validity may also be applied to evaluate programs, such as educational interventions, prior to their implementation in the field. Estimations of content validity seek to determine whether the items or constructs contained in the assessment or intervention being evaluated match their definition, are relevant to its purpose, and are representative of the construct being measured (Grant & Davis, 1997; Lynn, 1986; Polit & Beck, 2006). Too often, the specifics of how experts actually helped to shape the constructs addressed are left out; the reader is left with the assurance that content validity was addressed and established, without explicit details on how this was accomplished. Therefore, this study sought to employ strategies for gathering expertise from a panel of researchers and practitioners (e.g., teachers) to inform the development of a book reading intervention designed to teach the expository language of science to preschool children with language impairment.

Context of the Book Reading Intervention

Nearly half of the approximately 700,000 children who receive preschool special education services in the United States are identified with a primary disability of language impairment (U.S. Department of Education, National Center for Education Statistics, 2016). Reading disabilities in school age children have been related to identification of language impairment in young children (Catts, Fey, Tomblin, & Zhang, 2002). Language and literacy interventions to address language impairments and decrease the likelihood of future reading difficulties are necessary.

Dialogic book reading is one approach used to build language skills in preschool children (Wasik & Bond, 2001). Teachers, therapists, and parents using this technique read interactively, including the child in discussion of the text as it is being read (Whitehurst et al., 1988). Although this practice has become more common, it is used mostly with

[1]University of Cincinnati, OH, USA
[2]Miami University, Oxford, OH, USA

Corresponding Author:
Megan Schneider Dinnesen, School of Education, University of Cincinnati, 2600 McMicken Cir., TDC 260E, Cincinnati, OH 45221, USA.
Email: megan.dinnesen@uc.edu

narrative texts. Research has shown that young children have limited exposure to informational or expository text (Duke & Kays, 1998; Pentimonti, Zucker, Justice, & Kaderavek, 2010). Instruction in the early grades often focuses on learning to read, while later the focus shifts to reading to learn. Without explicit instruction in the text structure and signal words associated with this genre, children often struggle; this is particularly true for children with language impairment (Williams et al., 2005).

Text Organization for Preschoolers in Special Education (TOPS) was a 4-year project funded by the Institute of Education Sciences, U.S. Department of Education. TOPS focused on the development of an interactive expository book reading intervention for preschool children with language impairments. The intervention was comprised of three instructional modules to be implemented in teams that included one early childhood special educator and one speech-language pathologist each. These practitioners would incorporate the intervention with a small group of three students as a supplement to their instruction and therapy. Years 1, 2, and 3 of the project were spent developing and testing the intervention. The intervention was implemented in its entirety (Modules 1, 2, and 3) in the final year. The study described in this article was completed during the first phase of development and focused only on Module 1, which addressed the sequence text structure and life science (specifically, plant life cycles).

Content Validity and Intervention Development

Establishing content validity, particularly when developing educational interventions, is critically important to student outcomes. Nevertheless, content validity has not been investigated or reported in the research literature on educational interventions. Educational interventions should be developed to address the construct one is trying to teach; examining content validity provides a means for assuring that the intervention components are a reflection of that construct. Interventions that do not establish content validity risk teaching outside of the intended constructs or diffusion of constructs they intended to teach, compromising potential positive effects of the intervention on intended student outcomes. Also, establishing content validity requires a clear set of definitions and criteria by which consumers can determine the legitimacy of the intervention for addressing the construct it claims to teach. Determining the content validity of effective interventions informs future intervention development as it allows researchers to pinpoint specific constructs that are responsible for observable behavior change. Therefore, as with assessment development, the design and implementation of content validity processes when developing an intervention are significant to the credibility, future acceptability, generalization, and application of program findings (Davis, 1992; McKenzie, Wood, Kotecki, Clark, & Brey, 1999).

There is an ongoing effort to bridge the researcher–practitioner gap by utilizing models of implementation science during intervention development to produce interventions that are both evidence-based and feasibly sustainable in real-world settings. The principle of content validity fits neatly within many implementation frameworks. For example, the Consolidated Framework for Implementation Research (CFIR; Damschroder et al., 2009) includes five domains to consider during intervention development: the intervention itself, the inner setting, the outer setting, the individuals involved, and the process for accomplishing the intervention. Including an expert panel of researchers and practitioners allows for careful consideration of these five domains in a way that can improve the overall implementation of interventions.

In spite of the importance of establishing content validity for interventions, most of the research on content validity related to preschool children with disabilities has focused on its application in assessment development. For instance, content validity has been established for numerous preschool assessments of narrative language such as the *Narrative Language Measures* (NLM: P; Petersen & Spencer, 2012) and the *Narrative Assessment Protocol* (NAP; Justice, Bowles, Pence, & Gosse, 2010) as well as norm-referenced general language measures such as the *Structured Photographic Expressive Language Test–Preschool: Second Edition* (SPELT-P2; Greenslade, Plante, & Vance, 2009). Yet few language interventions apply the tenets of establishing content validity for assessments to the development phase of an intervention. With a few modifications, much can be learned from the development of assessments and applied to the development of interventions.

A Process for Establishing Content Validity

Content validity as it applies to an intervention includes evaluation of all of its components. That is, rather than focusing solely on individual items as in developing an assessment, experts validating the intervention must take into account components such as context, format, and method of instruction. Lynn (1986) described a two-stage process for determining content validity. In the first stage, the developers determine the full content domain by conducting a thorough review of the relevant research literature. In the second stage, a panel of experts systematically judges the content validity of the intervention components as a whole, using qualitative and/or quantitative methods. Because no studies that we are aware of describe the content validity process or the use of expert consultants in the development of an educational intervention, we drew upon healthcare and nursing interventions to inform how

we established content validity of our intervention using qualitative and quantitative methods (Bakas et al., 2009; Davis, 1992).

In healthcare and nursing interventions, qualitative methods for measuring content validity include interviews and feedback with content experts (Brod, Tesler, & Christensen, 2009). For example, Wagner, Smith, Ferguson, van Bakergem, and Hrisko (2011) developed an intervention that provided epilepsy education, primary coping skills, and secondary coping skills to youth with epilepsy and their families. In their development of the intervention, the authors described creating the intervention content from a review of the literature and focus group feedback. They also included two external experts who provided feedback, which the authors subsequently used to revise the intervention.

Although there is general consensus on qualitative methods for establishing content validity, there is a lack of consensus regarding the best way to quantify content validity. For example, the index of content validity (CVI; Polit & Beck, 2006; Waltz & Bausell, 1981) is a commonly used means of quantifying content validity of assessment instruments. Nevertheless, some studies utilize a proportion agreement (e.g., Beck & Gable, 2001; Lindell & Brandt, 1999), and others utilize kappa coefficient analyses (e.g., Wynd, Schmidt, & Schaefer, 2003). Furthermore, in developing scales, some methods require universal agreement on items, whereas other methods average item-level content validity indices (Polit & Beck, 2006). Regardless of the method, it remains important for intervention developers to utilize a systematic and quantifiable method for determining content validity. In one study utilizing quantitative methods for establishing content validity of an intervention, Bakas and colleagues (2009) employed a panel of 10 experts, including researchers, clinicians, and caregivers who had experience working with stroke patients. The experts determined content validity for components of their stroke-patient caregiver intervention across four criteria: accuracy, feasibility, acceptability, and problem relevance. Each member of the panel rated the components of the intervention using a 5-point rating system questionnaire. The panel members also submitted recommendations for improvements. Overall, the use of an expert panel determined initial content validity and informed the development process for future iterations of the intervention. Accordingly, conducting a content validity study in the development phase of an intervention conceivably prevents significant revisions to an intervention after its implementation (Rubio, Berg-Weger, Tebb, Lee, & Rauch, 2003).

Purpose

In this study, we applied the framework of intervention content validity represented in the healthcare field to the development of an expository book reading intervention for preschool children with disabilities. The research design was iterative; therefore, consultation with a panel of expert practitioners and researchers was critical to the development of the intervention. Specifically, this study investigated the following research questions:

Research Question 1: How can content validity of a book reading intervention be systematically measured?
Research Question 2: What themes emerged from data collected from an expert panel of researchers and practitioners related to the development of an book reading intervention?
Research Question 3: How were these themes used to inform further development of the intervention?

Thus, this study investigated approaches to measuring content validity and applied the tenets of this construct in a quantitative and qualitative approach to the development of an intervention. Results from this study attempt to address a gap in educational research and are of practical importance for researchers who have incorporated expert consultants in their research design and want to maximize their ability to harness and apply the knowledge and experience these professionals possess. These findings are also important for practitioners. Understanding the processes by which interventions might undergo content validation creates informed consumers of potentially adopted interventions for classroom use.

Method

Expert Panel Members

Three members of the expert panel evaluated the intervention aims and materials using content validity rubrics (see "Data" section for description of rubrics). Two expert researchers and one expert practitioner participated in this process. Expert 1 has been recognized as a distinguished professor at her academic institution. She was a licensed speech-language pathologist with clinical and classroom experience. She was the recipient of many prestigious fellowships, grants, and awards. She has published books and numerous peer-reviewed articles in leading journals on early language and literacy development and interventions for children with disabilities. Expert 2 was the executive director of a research center at a major research university. She was also a licensed speech-language pathologist with clinical and classroom experience. She was the recipient of numerous grants, fellowships, and awards. She too had many publications including books and peer-reviewed articles in leading journals examining the effects of language and literacy interventions on student outcomes. Expert 3 was an early childhood intervention specialist with more than 25 years of experience in teaching and directing preschool

programs for students with and without disabilities in a local school district. Rubio et al. (2003) pointed out that including practitioners on expert panels can yield beneficial results; they are often the individuals who become research participants and have the greatest insights into the practical implications of a research tool or approach. In addition, collaboration between researchers and practitioners can effectively bridge the gap between research and practice by creating materials and approaches that are more practical, feasible, and relevant.

Procedure

The iterative design of this study incorporated collaboration with the expert panel for multiple reviews of the intervention. Specifically, the structure for the collaboration included three parts: (a) initial expert review of developed materials, (b) a 1-day face-to-face meeting with the expert panel, and (c) subsequent expert review of revised materials to ensure that feedback was captured and incorporated accurately.

Initial expert review. Two weeks prior to the 1-day face-to face meeting, packets consisting of intervention materials were mailed to the experts. The contents of the packet included an overview of the domains and targets of the intervention and the scope and sequence of the intervention. In addition, tables explaining the selection of books and materials to be used were included as well as detailed plans for each implementation of the intervention. Content validity rubrics, directions for their completion, and a brief rationale were sent along with the materials prepared for the experts. Table 1 provides a list of all documents contained in the packet sent to the expert panel members. Experts were instructed to bring the completed rubrics to the meeting. This first set of completed content validity rubrics constituted Data Package 1.

Expert panel meeting. Members of the expert panel convened for a 1-day meeting to discuss the materials reviewed with the research team. Tojib and Sugianto (2006) pointed out that assembling experts for discussion allows for clarification and expansion of opinions immediately and encourages more in-depth exploration of the construct and concepts being targeted. Two members of the research team created in vivo transcripts of the expert panel meeting. A notepad with open-ended questions was also created prior to the meeting and used by expert panel members during the 1-day meeting to record any comments they did not have a chance to share during the discussion.

Subsequent expert review. Significant changes were made to the intervention components following the initial review and meeting. The subsequent or second review was designed to solicit feedback from the expert panel regarding the substance and form of changes made. This opportunity allowed our team to verify that we incorporated the feedback shared from the initial review and 1-day face-to-face meeting. The materials included in the packet sent for the second review are listed in Table 1.

Data

Four primary data sources were used to collect qualitative and quantitative information from members of the expert panel: (a) rubrics, (b) notes completed by the experts during the face-to-face meeting review of the intervention, (c) transcripts of the face-to-face meeting, and (d) a set of guiding questions designed to solicit more in-depth feedback on changes made following the initial review and expert panel meeting. The content validity rubrics were designed to capture ratings and comments, while the open-ended questions used in the notepad and the guiding questions were targeted to ask for specific suggestions and richer feedback.

Content validity rubrics. Content validity scales or rubrics are often used to capture ratings on materials or assessments being developed (Grant & Davis, 1997). Therefore, we adapted the rubrics from those originally developed by Grant and Davis (1997) to assess caregiver burden within a health care context. Revisions were made to the rating categories to capture information relevant to the intervention materials being evaluated. The categories used for rating the materials on the rubrics were aimed at collecting comments relevant to the purpose of each of the materials. In particular, we were interested in examining the content validity of the intervention lesson plans or, as we called them, Target-Technique Cards. The Target-Technique Cards were a component of the intervention to be used by practitioners while implementing each book reading lesson. Each Target-Technique Card included two parts: the targets or content the practitioners would be teaching and the strategies or techniques they would use to address the targeted content. Therefore, we asked our experts to complete two rubrics during the initial review of the materials: one to evaluate the intervention targets (what to teach) and the other to evaluate the techniques to be used when delivering the intervention (how to teach it).

The content validity rubric that evaluated the targets of the intervention (see rubric in the appendix) included two categories: comprehensiveness and clarity. The content validity rubric that evaluated the techniques of the intervention also included two categories: application and clarity. Both content validity rubrics asked experts to rate the targets and techniques on the corresponding categories (targets-comprehensiveness, clarity; techniques-application, clarity) on a 4-point rating scale: 1—*not clear*, 2—*requires*

Table 1. Contents of Materials Packets Sent to Expert Panel Members.

Packet sections	Data Package 1 December	Data Package 3 June
Section I. Scope and sequence This section included materials that provided an overview of the targets of the intervention as well as an overall plan for its implementation.	Scope and sequence Domains and targets	Scope and sequence
Section II. Selection processes This section included materials that illustrated the processes used to select key components of the program including books to be used and content to be addressed.	Book selection: Coh-Metrix Data Sheet	Instructional targets selection: Selection process of topic knowledge and academic vocabulary
Section III. Instructional materials for practitioners This section included materials designed for practitioner use while implementing the intervention. This included lesson plans (intervention plan, later called target-technique card) as well as scripts, which were later changed to examples.	Techniques chart Intervention plan—Module 1, Week 1 Weekly intervention preorganizer Weekly target card technique card Intervention plan—Week 1, Day 1 Intervention plan—Week 1, Day 2 Intervention plan—Week 1, Day 3	Target-technique card Procedures for interactive book mapping Question bank
Section IV. TOPS alignment with national and state standards charts This section was created in response to expert panel feedback collected in Data Packages 1 and 2. These tables were created to align the targets and techniques of the intervention to applicable state and national standards in both literacy and science.		Targets: Text structure and signal words Targets: Topic knowledge and vocabulary Techniques
Section V. Content validity rubrics These rubrics were created to capture quantitative and qualitative feedback from the expert panel members. Each rubric had directions for completion and was associated with a particular document or set of materials.	Targets Techniques	Target-technique card: Targets Target-technique card: Context for interactive retellings Target-technique card: LEAD strategy Question bank

Note. TOPS = Text Organization for Preschoolers in Special Education. LEAD = L - Lead with a text structure question E - Extend children's responses A - Ask an inferential question D - Develop & Aid

major revisions, 3—*requires minor revisions*, and 4—*is clear*. They also included a space for writing positive comments and/or suggestions for improvement. The same set of rubrics were distributed to the experts with the subsequent expert review materials after significant revisions were made to the intervention materials following the initial review and 1-day face-to-face meeting.

In the subsequent expert review, two additional rubrics were also developed and used. One rubric asked for feedback on a Question Bank developed in response to feedback on the techniques of the intervention during the initial review and meeting. One technique of the intervention included practitioners asking children open-ended questions during the book reading lesson; the Question Bank was developed to provide practitioners with examples of open-ended questions. The other additional rubric sought feedback on extension activities that were part of the book reading intervention. Extension activities utilized graphic organizers or maps to supplement the interactive book reading and were designed to elicit discussion, practice language skills, and provide children with opportunities to demonstrate their understanding of the concepts being taught in the intervention.

Notepad. A notepad was devised for the face-to-face meeting consisting of four open-ended questions designed to capture feedback the panel members did not have a chance to share during the discussion. The four open-ended questions included the following:

1. What works in this section?
2. What is challenging in this section?
3. Tell us about some suggestions you have.
4. What haven't we discussed?

Members of the expert panel were asked to respond to these questions and record any feedback not discussed during the face-to-face meeting.

Meeting transcripts. In vivo transcripts of the meeting were created. These transcripts included word-by-word transcription of the expert panel meeting. After the face-to-face

Table 2. Contents of Data Packages.

Data Package 1 December Review of materials prior to meeting	Data Package 2 December Data collected at expert panel meeting	Data Package 3 June Review of materials revised to incorporate prior feedback
Content validity rubrics	Meeting transcripts Notepad	Content validity rubrics Guiding Questions

meeting, two members of the research team reviewed and cleaned the transcripts using audio-recordings. The cleaning process included both team members reviewing each transcript and checking for consistency between transcripts and with the recordings. Any discrepancies were discussed until agreement was reached.

Guiding questions. A list of open-ended guiding questions was also provided to the expert panel in the subsequent expert review packet sent after revisions to the intervention were made. The following questions were provided:

1. How do you think the changes we've made reflect the feedback we received?
2. What further changes do you believe we still need to make?
3. What challenges do you think we might encounter during implementation?
4. What suggestions do you have for addressing these challenges?
5. Please provide any other comments you feel are relevant.

Members of the expert panel were asked to respond to these questions after reviewing the revised materials provided in the subsequent expert review packet.

Data Analysis

The data collected was analyzed using content analysis (Johnson & LaMontagne, 1993) and a grounded theory framework (Corbin & Strauss, 2007). Data included in this study were grouped into three time-oriented packages (see Table 2 for data packages). The first package included the rubrics completed during the initial expert review. The second package included notes pages from the 1-day face-to-face meeting and in vivo transcripts created and edited using recordings of the meetings held that day. The third package included data collected during the subsequent expert review after revisions were made to the components as recommended. These data were comprised of the responses to the set of guiding questions and rubrics completed by the panel members. We followed the six steps outlined by Johnson and LaMontagne (1993). Specifically, these steps included: prepare the data for analysis, become familiar with the data, identify units of analysis, define tentative categories for coding the responses, refine categories, and establish category integrity.

Rigor in qualitative research must be formally established using strategies specific to this methodology. To address the trustworthiness of the qualitative data in this study, the criteria and strategies put forth by Guba (1981) and Krefting (1991) were consulted and followed. For example, member checking was completed by consulting with the expert panel prior to submission of this article. Each researcher had the opportunity to review the data sources and "Results" and "Discussion" sections of this article and agreed that their feedback was captured and described accurately. Also, peer examination of coded data was conducted; discussion was held until agreement was reached.

To calculate a quantifiable estimate of content validity, scores on each rubric were tallied and a mean score was calculated. An assessment of the validity of both the structure and content of the intervention was derived by comparing the ratings across experts.

Results

Content Validity Themes

Qualitative analysis of data collected from the expert panel's initial review and face-to-face meeting (Data Packages 1 and 2) revealed several themes relevant to the development process of the intervention. Specifically, three major themes were identified: get the science right; guide more, script less; and get into the field. See Table 3 for the triangulation of data sources and themes.

Get the science right. The primary aim of developing the expository book reading intervention for children with disabilities was to expose them to the text structures and academic language inherent in expository texts. Because science content is expository in nature, science as a subject was used as vehicle for exposing children to various expository text structures and language. During the initial review of the intervention however, the expert panel challenged us to rethink our perspective on science. The panel members stressed that the science must be included not simply as secondary to text structures and academic language, but rather

Table 3. Examples of Data Collected From Expert Panel Members.

Themes	Experts	Data Package 1	Data Package 2	Data Package 3
Get the science right	E1	Look at the NGSS (science framework)—it indicates we should teach science content always in its context—avoid definitions. The three areas of science are earth, life, and physical. You do not have these three areas represented. Read the NGSS—this is a framework for K–12 science education published by the National Research Council in 2012.	Your emphasis needs to be on inquiry; not just explicitly teaching this vocabulary. For science, inquiry is integrated across the curriculum. The science standards say don't teach definitions—just teach it within a framework of inquiry and in a naturalistic approach. This is what teachers are being told right now—it's aligned with the emphasis on science content and instruction right now.	Avoid common science misconceptions (e.g., "Plants need food to eat."). Encourage your teachers to mindfully use language to promote correct understanding. Consider distributed exposure to topics/text structures rather than tying science topics to text structures.
	E2		You have chosen science so it is absolutely critical that you get it right. I think you have some work to do on figuring out what that will look like.	
	E3	The children should have other experiences with a variety of plants during the intervention so they know firsthand "What other plants might grow like a bean plant?"		Come up with a time for children to generate their own questions about the topics.
Guide more, script less	E1			You are still asking for A LOT here. Finding the balance between an intervention that is structured and open-ended will take trial and error.
	E2	This is too comprehensive—and very overwhelming.	If you believe the most important part of this is expressive language—when you have a really heavily scripted book-reading routine, to what extent are you really promoting discussion? There is scripted and there is teacher-guided—they are both effective, but they are very different. Do you want teachers to go about doing the business of delivering the intervention you've designed or learn to new things themselves?	Provide more examples for your teachers—think of making several pages of examples—one for each thing you want them to do.
	E3		I like the scripts as a starting point, but if you want teachers to learn to do these things, eventually you would want those prompts or supports to fade.	
Get into the Field	E1		Consider doing a trial run with a naïve teacher to get a better idea of timing and issues.	
	E2	Can you test this with a few kids?	Taking a lesson and sitting with a focus group of teachers is not useful. The most information you will get will be when you really try it in a classroom. Being in classrooms with teachers is the only thing that will really tell you what works. Be attentive to what teachers identify as barriers—what will keep them from doing what you want them to?	The changes you need will come from your observations and from talking with your participants.
	E3		How will you support teachers who struggle with some of these expectations?	

Note. NGSS = Next Generation Science Standards.

as equally important in developing children's understanding of science content. Science instruction in early childhood is a timely topic. As one expert pointed out, "You have chosen science so it is absolutely critical that you get it right." The expert researchers recommended that we more closely align the concepts and content of the intervention with the

National Research Council's (NRC) Next Generation Science Standards (NGSS; NGSS Lead States, 2013). They challenged us to make revisions that would demonstrate that science was just as critical to the intervention as the text structures and academic language.

In addition, the academic vocabulary and science content of our intervention came under scrutiny from the expert panel. They recommended targeting fewer words of varied parts of speech (rather than all nouns) and avoiding defining the words. From a science perspective, words should always be taught embedded within their context using an inquiry-based approach. More than one expert cautioned us against defining the words, and the expert practitioner in particular felt strongly that an experiential component was needed. Although the expert researchers brought up each of these topics on the content validity rubrics, the discussion held during our expert panel meeting was in-depth and involved and proved to be crucial to the changes made. The suggestions and insights shared in respect to science content and instruction were invaluable and prompted significant changes to the scope and sequence of the intervention and our study timeline.

Guide more, script less. The intervention materials originally developed for the expository book reading intervention provided scripted information for each week of implementation. The feedback on the scripted nature of the lesson plans indicated another major theme. The expert panel agreed that the overall nature of the intervention, including the lesson plans, was far too scripted. In Data Package 1, one expert researcher commented on the rubric: "This is too comprehensive—and very overwhelming." The panel suggested that the level of prescriptiveness would ultimately impact participant attrition in that practitioners would tire of following a script word-for-word. Instead, the expert panel members emphasized that the practitioners should have room for creativity and that teacher-guided, inquiry-based learning is an essential hallmark of best practices in both literacy and science instruction. The discussion at the expert panel meeting (Data Package 2) included this theme as well; the expert practitioners felt the scripts would be helpful at first, but that the scripts and supports should fade each week. The expert researchers drew a clear distinction between scripted and teacher-guided. Citing their own experience, they suggested a more teacher-guided approach could facilitate practitioner learning and creativity. Thus, the expert panel encouraged us to create a guide for practitioners to use during implementation of an intervention session.

Get into the field. During the 3 months of initial development, our research team conducted extensive reviews of the literature and created all of the intervention materials using the currently available research. In addition, years of experience in the schools as special and early childhood educators as well as speech-language pathologists informed our development. Although a preliminary trial with a small group of three typically developing preschool children had been conducted with the first iteration of the intervention, we had yet to conduct any observations or trials with practitioners in the field with children with disabilities. One of the themes that resonated from the expert panel during the initial review was to collaborate directly with practitioners currently active in the field. Furthermore, it was recommended that we not just talk with these practitioners, but actually observe what they were already doing with expository books in the classroom and enlist them in implementing the intervention. One expert said right away that talking with teachers would not be beneficial; she felt strongly that being in classrooms and observing firsthand was crucial to the development of our intervention. In addition, they encouraged us to get into the field as soon as possible to enlist practitioners in helping us to identify the barriers to successful implementation before executing a full module. The discussion around this theme illustrated the benefit of holding a meeting. As Tojib and Sugianto (2006) pointed out, the ability to talk through each of these themes and ask for more detail proved advantageous. From this feedback, our team sought a team of practitioners who would allow us into their classrooms and could advise us on the feasibility and practicality of the intervention.

Revisions and Further Intervention Development Based on Themes

These three salient themes, once identified, served as catalysts for changes to the intervention itself as well as the study overall. Each theme informed the development process of the intervention and prompted research and revision based directly on recommendations from our expert panel.

Get the science right. Prompted by the expert panel, our team began research on the NRC/NGSS framework and determined we would change the intervention so it would be in direct alignment with three broad science content areas: life science, earth science, and physical science. This decision had ramifications on our study timeline as well as the intervention itself. Rather than the four-module study design that was originally proposed, aligning the modules with the science content areas prompted us to change to three modules.

As a result of changing to three modules aligned with each science content area, we revisited the books that had been selected, as well as the science content and academic vocabulary targets chosen for each one. One of our expert researchers suggested we rethink the number and type of vocabulary words we were targeting. This researcher urged

us to choose words that were not only nouns, but also adjectives and verbs. In addition, she felt strongly that we were focusing on too many words, which would limit children's ability to learn and use these words. She feared children would not demonstrate growth in the area of vocabulary on the intervention-based measure without changes to the scope of our vocabulary instruction. The other expert researcher suggested changes to the method of vocabulary instruction. She gave the vocabulary section of the Target-Technique Card a rating of 3 on the content validity rubric and explained this choice by saying our vocabulary words were not taught within context. She recommended we revisit the NGSS to refine how the practitioners would introduce and embed the targeted vocabulary words and concepts within the intervention lessons. As a result of this feedback, we employed a systematic, research-based approach to the selection of the academic vocabulary words to be targeted. Pulling from the NGSS, we included an inquiry-based approach and changed the techniques used for vocabulary instruction by the practitioner as a guide to discovery, rather than a provider of new facts.

Finally, after conducting an in-depth content analysis of the NRC/NGSS, the Common Core State Standards (CCSS), the National Association for the Education of Young Children (NAEYC), and the State of Ohio's Early Learning and Development Standards, our research team created a series of tables illustrating the alignment of the structure and content of the intervention with these key standards. These tables served as guides as we revised the structure and content of the intervention. In addition, we have shared these tables with practitioners and administrators to demonstrate the standards-based nature of our program.

Data collected in Data Package 3 supported the changes made and prompted further changes to the intervention. The expert panel commented on how thoroughly the science content was now covered and represented. Specifically, the experts shared positive comments on the alignment tables created; they remarked on the significant research that went into their creation. In addition, one expert was impressed with the changes the research team made to the scope and sequence of the intervention as a result of researching science standards and inquiry-based instruction. In addition, the expert researchers whose comments guided changes to the content of the intervention targets rated each area higher on the rubrics than during the first review. Our team included a document explaining the selection of these targets for the second review. Both expert researchers commented that this content was now appropriate.

Suggestions for further changes were also made in Data Package 3. These included comments relevant to the science content of the intervention. One panel member pointed out the importance of representing science content accurately and recommended resources for avoiding common misconceptions often taught in early learning experiences involving science. One of the expert researchers recommended we consider distributed exposure to topics and text structures. Rather than tying each text structure to a science topic, she recommended we consider having each module focus on a text structure, but also introduce the others. For example, rather than having life science taught with only the sequence text structure, we should consider having this module include a week of cause/effect and a week of compare/contrast as well. Furthermore, the expert practitioner complimented the changes made to make the materials more teacher-guided, but felt we still had work to do on including inquiry. She recommended we include a time for students to add their own questions during the lesson; she pointed out that our focus was still on the questions the practitioners would ask, rather than where the students could take the discussion.

Guide more, script less. The insights shared on teacher-guided, inquiry-based learning led our team to significantly revise the intervention plans and materials designed for practitioner use. We took out the scripted discussion and questions from the intervention lesson plans. In their place, we created a new Target-Technique Card that listed the expository structure and content targets practitioners should address during an intervention session and listed the techniques for addressing each target (e.g., ask a question). Instead of scripts, a bank of questions and examples for how one might discuss expository structures and content before, during, and after reading were created for practitioners to reference. Our team researched the hallmarks of inquiry-based instruction (Gerde, Schachter, & Wasik, 2013) and incorporated these tenets into the materials created for practitioners as well as the plans for the accompanying professional development. Emphasis was placed on providing practitioners with materials that were more teacher-guided and incorporated a more dynamic and inquiry-based approach to supporting students in their learning. The new Target-Technique Card could be used as a planning tool; practitioners would learn the techniques and choose how to use them to address the intervention targets.

The changes made to the materials in response to Data Packages 1 and 2 again prompted many favorable comments from the expert panel. For example, one expert researcher complimented the question bank we developed saying, "This is excellent! It is practical and aligns well to what you are asking practitioners to do." Another expert liked the more teacher-guided materials and recommended varying the extension activities included in the lesson (Kamhi, 2014).

Get into the field. Following the recommendation to collaborate directly with active practitioners, we recruited four practitioners working in two teams comprising one early

childhood special educator and one speech-language pathologist each. Over the course of 4 weeks, we conducted classroom observations and interviews with these practitioners. The observations during the first week captured "business as usual" with the aim of discovering what was typical practice with book reading in the classroom. Two members of our research team took detailed field notes during each observation, which were discussed at the weekly team meeting. An interview was conducted with each practitioner after their observation. The questions for Week 1 aimed to ascertain how each practitioner viewed their role in literacy instruction and development. Each practitioner was given one of the books chosen for the intervention at the conclusion of the first week. They were instructed simply to plan and conduct a lesson, which we would observe, during the following week. Again, each practitioner was observed for 1 hr during the second week. Field notes were taken and each practitioner was interviewed. Questions during the second interview focused on planning, instruction, and assessment for the lesson observed. Our research team provided a 1 hr face-to-face professional development session during the third week. Practitioners were introduced to the research behind the intervention, the structure and content of expository text, and the targets and techniques of our intervention. Each practitioner was given materials for 1 week of implementation. During Week 4, each practitioner implemented the intervention in their classroom or therapy session. Again, our team observed these lessons and took detailed field notes. Each practitioner was interviewed about how they planned for, implemented, and assessed the success of their intervention lesson.

The information from the observations and interviews informed further revisions and development of the intervention and intervention materials. For instance, in one classroom, the teacher implemented the intervention within the context of "centers" time. This practitioner expertly rotated small groups of students through the intervention lesson, modifying the content and delivery each time. During the interview, this practitioner discussed his planning and differentiation strategies. Our research team had not considered the possibility of implementation in this way. This practitioner helped us rethink the feasibility of our intervention. In addition, our observations with the practitioners showed that each was consistently able to use the "making connections" technique. Each incorporated text-to-self, text-to-life, and even text-to-text questions (Santoro, Chard, Howard, & Baker, 2008) within their lessons. Ultimately, we dropped this as a technique and focused on asking practitioners to incorporate more inferential and open-ended questions into their lessons.

The insights we gained from these practitioners were invaluable and prompted further changes to the techniques used in teaching the targets of the intervention as well as one of the critical measures of student growth and learning, specifically, the mapping activity, completed at the end of each book reading session. The suggestion to ask for advice from active practitioners proved to be advantageous to the development process.

Expert panel responses in Data Package 3 on the data gathered from working with these four practitioners was positive. Our ability to get into the field and systematically solicit and incorporate feedback from active practitioners was impressive to our panel.

Content Validity Ratings

The assessment of the content validity of this intervention also included quantitative data collected from the rubrics. Data Package 3 included a total of four rubrics designed to elicit expert feedback on the targets and techniques to be used during the intervention as well as the question bank designed to provide examples for practitioners and the extension activities included with the lesson. Panel members rated materials for clarity, comprehensiveness, and ease of application. Each member's score, along with mean scores, are presented in Table 4. These scores enabled the research team to determine each reviewer's impression of the materials; means were calculated to assign an estimated value to the work accomplished and the work yet to be done.

For example, after review of the revised materials, Experts 1, 2, and 3 each rated the information presented on one of the targets (signal words) at a 4/4. This rating indicated that the experts recommended little to no further revision to the signal words materials. Conversely, the materials presented on the revised technique, "develop and aid," received a mean score of 2.67/4 (individual expert scores 4, 2, and 2) in both clarity and application. This rating communicated to the research team that the expert consultants believed we had more work to do in refining this concept and its presentation in our materials. The comments the experts provided along with their ratings guided further revision of the materials for the "develop and aid" technique. Because the development of the intervention was an iterative process, the quantitative data collected in Data Package 3 provided evidence of the need for further changes; the comments and answers to guiding questions provided direction for continued improvement. For example, one expert continued to rate the area of topic knowledge as represented on the Target-Technique Card poorly. Even in Data Package 3, E2 rated topic knowledge with a score of 1 for both clarity and comprehensiveness. The comments provided with this rating, however, provided explanation and direction; E2 felt the information presented on the Question Bank relevant to topic knowledge would be useful if included on the Target-Technique Card. In this way, both the quantitative ratings and qualitative feedback were used to incorporate changes to the materials.

Table 4. Quantitative Data Collected With Data Package 3.

Rubric Contents	E1	E2	E3	Mean score
Targets (four items)				
Signal words				
Clarity	4	4	4	4
Comprehensiveness	4	4	4	4
Text structure				
Clarity	4	3	4	3.67
Comprehensiveness	4	4	4	4
Academic vocabulary				
Clarity	4	3	3	3.33
Comprehensiveness	4	3	3	3.33
Topic knowledge				
Clarity	3	1	4	2.67
Comprehensiveness	3	1	4	2.67
Question bank (one item)				
Intervention question bank				
Clarity	3	4	3	3.33
Comprehensiveness	3	4	3	3.33
Retelling activities (three items)				
Dramatization				
Clarity	3	4	3	3.33
Application	3	4	4	3.67
Completed map				
Clarity	3	4	3	3.33
Application	3	4	3	3.33
Map manipulation				
Clarity	3	4	3	3.33
Application	3	4	3	3.33
Techniques (four items)				
Lead with a text structure question				
Clarity	4	3	3	3.33
Application	4	3	1	2.67
Expand and clarify children's response				
Clarity	4	3	2	3
Application	4	3	2	3
Ask an inferential question				
Clarity	3	3	3	3
Application	3	3	3	3
Develop and aid students' ability to identify and utilize evidence to answer questions				
Clarity	4	2	2	2.67
Application	4	2	2	2.67

Note. Rated on a 4-point scale where 1 = *this component is not addressed satisfactorily*, 2 = *the content on this component requires major revisions to be covered satisfactorily*, 3 = *the content on this component requires minor revisions to be covered satisfactorily*, and 4 = *this component is covered satisfactorily*, for all items.

Discussion

Contributing to existing research on content validity, this study identified themes in evidence gathered from expert researchers and practitioners and explained how these insights were incorporated to inform development of an intervention. It adds to the extant body of research on this topic by providing an in-depth look at the benefits of collaboration with a panel of experts and fills a gap by applying the principles of content validity within the context of an educational intervention.

Although an expert panel of researchers was incorporated into the design of this study, we realized a systematic approach would be necessary to gather meaningful data and

fully utilize the expertise of our panel. Drawing from the extant body of literature proved to provide a starting point, however, demonstrated the need for the application of the tenets of content validity to development of an educational intervention. Thus, this article describes the processes we established and followed to collect, analyze, and use data from an expert panel in the development phase of an intervention. Beginning with the tenets put forth by Lynn (1986), a series of content validity rubrics was created. As in health care interventions mentioned in the literature (Bakas et al., 2009; Grant & Davis, 1997), these rubrics provided structured data artifacts that guided experts in providing feedback and provided our team with consistent categories that allowed for comparison between experts and between data collection points (on our development timeline). The inclusion of a meeting with our expert panel provided rich discussion that prompted further research and many changes.

Data collected from consultation with our expert panel was analyzed both qualitatively and quantitatively, and provided evidence of the need for revisions as well as direction for refining the materials, targets, and techniques to be used as part of an educational intervention for preschool children with language impairment. Three salient themes were present and identified throughout the evidence gathered. The insights and suggestions provided by our expert panel served as catalysts; these themes functioned as touchstones in our development process. Using the content validity rubrics we designed, a set of guiding questions and a notepad, as well as coding of in vivo transcripts captured during a face-to-face meeting, enabled our team to revisit our data as we made revisions and continued development. Communicating frequently with our experts helped us to ensure their feedback was being captured and incorporated accurately. Incorporating this research design enabled us to develop a robust intervention prior to implementation. Measures of content validity increase the likelihood that the program contains and measures what it is intended to; it could increase the fidelity to the implementation of the intervention and/or its success as measured by student progress (Davis, 1992; McKenzie et al., 1999; Rubio et al., 2003). Following this process resulted in significant and, we believe, positive changes to our intervention and our development process overall.

The process of determining content validity of our educational intervention corresponds with implementation models, such as CFIR. For example, the expert panel utilized rubrics and discussion to evaluate the comprehensiveness and clarity of the intervention itself. This afforded opportunities to determine whether intervention targets matched current research and theory on children's linguistic and academic skills. Practitioners went beyond the scope of the intervention by also examining the instructional techniques. At the inner setting level, the experts commented on the feasibility of implementing the intervention within classrooms. At the outer setting level, experts examined alignment of instruction with state and national educational standards as well as theoretical frameworks. Inclusion of practitioners and intervention researchers on the expert panel allowed discussion of both practicability and research validity as it related to development of the intervention. Considering each of these domains prior to the intervention process facilitated development and later implementation of the intervention delivered by teachers and speech-language pathologists within early childhood special education classrooms.

Although consultation with a panel of experts is often included in research design, the specifics of how these experts contributed to the development of a program or assessment are often omitted. The reader is left with a mention of experts and an assurance that they were involved. This study took an important next step in describing steps for systematically collecting data when consulting with a group of expert researchers and practitioners. To our knowledge, the application of the construct of content validity to an educational intervention has not been specifically addressed in the literature to date. Of practical importance for researchers and practitioners, this research illustrated the validation of the structure and content of an educational intervention prior to its implementation.

Appendix

CONTENT VALIDITY RUBRICS: Target-Technique Card

Directions: Please complete the following rubrics during your review of the Target-Technique Card (please see both sides of the card). Please include positive comments and/or suggestions for improvement in each area. Thank you!

Target/Technique Card	COMPREHENSIVENESS	CLARITY
Targets	1 = This component **is not** addressed comprehensively. 2 = The content on this component requires **major revisions** in order to be covered comprehensively. 3 = The content on this component requires **minor revisions** in order to be covered comprehensively. 4 = This component **is** addressed comprehensively.	1 = This component **is not** covered clearly. 2 = The content on this component requires **major revisions** in order to be covered clearly. 3 = The content on this component requires **minor revisions** in order to be covered clearly. 4 = This component **is** covered clearly.
Text Structure	1 2 3 4 Comments	1 2 3 4 Comments
Signal Words	1 2 3 4 Comments	1 2 3 4 Comments
Academic Vocabulary	1 2 3 4 Comments	1 2 3 4 Comments
Topic Knowledge	1 2 3 4 Comments	1 2 3 4 Comments

Source. Adapted from Grant, J.S. & Davis, L. L. (1997). Selection and use of content experts for instrument development. Research in Nursing and Health, 20(3), 269-274.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A130205 to the University of Cincinnati. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

ORCID iD

Megan Schneider Dinnesen https://orcid.org/0000-0002-7321-2595

References

Bakas, T., Farran, C. J., Austin, J. K., Given, B. A., Johnson, E. A., & Williams, L. S. (2009). Content validity and satisfaction with

a stroke caregiver intervention program. *Journal of Nursing Scholarship*, *41*, 368–375.

Beck, C. T., & Gable, R. K. (2001). Ensuring content validity: An illustration of the process. *Journal of Nursing Measurement*, *9*, 201–215.

Brod, M., Tesler, L. E., & Christensen, T. L. (2009). Qualitative research and content validity: Developing best practices based on science and experience. *Quality of Life Research*, *18*, Article 1263.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research*, *45*, 1142–1157.

Corbin, J. M., & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: SAGE.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science*, *4*, Article 50.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*, 194–197.

Duke, N. K., & Kays, J. (1998). "Can I say 'once upon a time'?": Kindergarten children developing knowledge of information book language. *Early Childhood Research Quarterly*, *13*, 295–318.

Gerde, H. K., Schachter, R. E., & Wasik, B. A. (2013). Using the scientific method to guide learning: An integrated approach to early childhood curriculum. *Early Childhood Education Journal*, *41*, 315–323.

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing and Health*, *20*, 269–274.

Greenslade, K. J., Plante, E., & Vance, R. (2009). The diagnostic accuracy and construct validity of the Structured Photographic Expressive Language Test–Preschool. *Language, Speech, and Hearing Services in Schools*, *40*, 150–160.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Technology Research & Development*, *29*(2), 75–91.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*, 238–247.

Johnson, L. J., & LaMontagne, M. J. (1993). Research methods using content analysis to examine the verbal or written communication of stakeholders within early intervention. *Journal of Early Intervention*, *17*, 73–79.

Justice, L. M., Bowles, R., Pence, K., & Gosse, C. (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol). *Early Childhood Research Quarterly*, *25*, 218–234.

Kamhi, A. G. (2014). Improving clinical practices for children with language and learning disorders. *Language, Speech, and Hearing Services in Schools*, *45*, 92–103.

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *The American Journal of Occupational Therapy*, *45*, 214–222.

Lindell, M. K., & Brandt, C. J. (1999). Assessing interrater agreement on the job relevance of a test: A comparison of *CVI, T,* $r_{WG(J)}$, and $r^*_{WG(J)}$ indexes. *Journal of Applied Psychology*, *84*, 640–647.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*, 382–385.

McKenzie, J. F., Wood, M. L., Kotecki, J. E., Clark, J. K., & Brey, R. A. (1999). Establishing content validity: Using qualitative and quantitative steps. *American Journal of Health Behaviors*, *23*, 311–318.

NGSS Lead States (2013). *Next Generation Science Standards: For States, by States*. Washington, DC: The National Academies Press.

Pentimonti, J. M., Zucker, T. A., Justice, L. M., & Kaderavek, J. N. (2010). Informational text use in preschool classroom read-alouds. *The Reading Teacher*, *63*, 656–665.

Petersen, D. B., & Spencer, T. D. (2012). The narrative language measures: Tools for language screening, progress monitoring, and intervention planning. *Perspectives on Language Learning and Education*, *19*, 119–129.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*, 489–497.

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, *27*, 94–104.

Santoro, L. E., Chard, D. J., Howard, L., & Baker, S. K. (2008). Making the very most of classroom read-alouds to promote comprehension and vocabulary. *The Reading Teacher*, *61*, 396–408.

Tojib, D. R., & Sugianto, L.-F. (2006). Content validity of instruments in IS research. *Journal of Information Technology Theory and Application*, *8*(3), 31–56.

U.S. Department of Education, National Center for Education Statistics. (2016). *Digest of education statistics, 2015* (NCES 2016-014, Chap. 2). https://nces.ed.gov/pubs2017/2017094.pdf

Wagner, J. L., Smith, G., Ferguson, P., van Bakergem, K., & Hrisko, S. (2011). Feasibility of a pediatric cognitive-behavioral self-management intervention: Coping openly and personally with epilepsy (COPE). *Seizure*, *20*, 462–467.

Waltz, C. F., & Bausell, R. B. (1981). *Nursing research: Design, statistics and computer analysis*. Philadelphia, PA: F.A. Davis.

Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology*, *93*, 243–250.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*, 552–559.

Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, K. B., DeSisto, L. A., & decani, J. S. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology*, *97*, 538–550.

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, *25*, 508–518.