Journal of MultiDisciplinary Evaluation
Volume 16, Issue 34, 2020

JMDE
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180
http://www.jmde.com

Mixed Methods Research in Designing an Instrument for Consumer-Oriented Evaluation

Dorinda J. Gallant
The Ohio State University

Nicole Luthy
The Ohio State University

Background: The educational product market has been gradually shifting from primarily print to primarily digital content. Educators must make quick decisions when selecting materials that will assist students in their learning.

Purpose: Purposes of this study were to describe the application of a two-stage sequential mixed-method, mixed-model design in designing an instrument for consumer-oriented evaluation and to describe implications of using mixed methods research in developing a rubric to evaluate prekindergarten through Grade 12 digital content.

Setting: The Ohio State University, Columbus, OH.

Intervention: N/A.

Research design: A two-stage sequential mixed-method, mixed-model design.

Data collection & analysis: In Stage 1, a modified electronic Delphi survey technique was implemented with US geographically dispersed subject matter experts. In Stage 2, cross-sectional focus group interviews were conducted with local teachers, administrators, and textbook publishers.

Findings: Inclusion of multiple perspectives and viewpoints from teachers, administrators, textbook publishers, and experts on importance, clarity, and appropriateness of criteria to evaluate digital content resulted in a final version of the rubric that can be used by teachers and administrators to evaluate digital content that supports students' learning in prekindergarten through Grade 12.

Keywords: *mixed methods research; modified e-Delphi survey; digital content; consumer evaluation; instrument development.*

Introduction

In recent years, the educational product market has been gradually shifting from primarily print to primarily digital content. The shift has created an incursion of digital products, from which educators can select, to remediate, supplement, or enrich students' learning. Furthermore, the shift has created challenges for educators who must select from among the plethora of digital content options available, generally without having the knowledge or skills, to make informed selections. For many educators, they must make quick decisions when selecting materials that will assist students in their learning. In prekindergarten through Grade 12 (preK-12) education, there are sources that provide reviews of educational products or instructional materials (e.g., SEEN [Southeast Education Network] Magazine and EdReports.org). The availability of information on validity evidence to support the reviews may or may not be publicly available or appear in the research literature. The lack of availability of validity evidence has implications for the credibility and trustworthiness of reviews as well as for the interpretability and use of the reviews.

In consumer-oriented evaluation, *valuing* is the key component (Fitzpatrick, Sanders, & Worthen, 2011). A set of criteria (including the standards for the criteria) presumed to be of value (i.e., important, desirable, or useful) to the consumer are used by an evaluator or evaluators to determine the merit or worth of the product. One of the shortcomings of the consumer-oriented approach is the lack of consumer (or stakeholder) input into criteria used to evaluate products (Stufflebeam & Coryn, 2014). Hence, the extent to which consumers (or stakeholders) provided input regarding criteria of importance (or of value) to them in the evaluation process, which ultimately may lead to them purchasing a product, is not widely incorporated. This study aims to address this shortcoming and show the value of incorporating stakeholders' input.

This article has two purposes. First, we describe the application of a two-stage sequential mixed-method, mixed-model design in designing an instrument for consumer-oriented evaluation. Specifically, we describe the development and initial content validation of the Digital Content Evaluation Rubric (DCER), an instrument to be used by teachers in several state of Ohio school districts to evaluate preK-12 digital products. We intentionally incorporate the perspectives of multiple stakeholders through quantitative and qualitative data collection to understand what they *value* in order to provide evidence of the appropriateness of the instrument to evaluate digital content. The DCER was developed for use by teachers in the Evaluating Digital Content for Instructional and Teaching Excellence (EDCITE) project to evaluate English language arts, mathematics, science, and social studies digital content (e.g., online courses, digital curriculum, electronic textbooks) in order to support learning in traditional face-to-face and blended elementary, middle, and high schools. It is important to note that while the delivery mode of instruction in EDCITE schools was primarily traditional face-to-face instruction, teachers at the middle and high school levels also implemented blended learning instruction. The question guiding the development and initial content validation of the DCER was, "What are the criteria essential to evaluate digital content that supports learning in preK-12 educational settings?" Second, we briefly describe implications of using mixed methods research in developing the DCER.

Mixed Methods Research and Significance of this Mixed Methods Research Study

Mixed methods research is one of three research approaches (i.e., qualitative research, quantitative research, and mixed methods research) implemented in many disciplines or fields to address research problems or phenomenon. Based on an analysis of definitions provided by leaders in the field of mixed methods research, Johnson, Onwuegbuzie, and Turner (2007) provided the following general definition:

> Mixed methods research is the type of research in which a researcher or team of researchers combines elements of

> qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration (p. 123).

The definition suggests that—as a type of research—mixed methods research follows a systematic inquiry process to collect and integrate multiple viewpoints through the collection of quantitative and qualitative data within a single study to describe, explain, or understand phenomenon.

Purposes (or rationales) for researchers to engage in mixed methods research have evolved over time. These purposes include establishing credibility of evidence in evaluation (Mertens & Hesse-Biber, 2013) as well as instrument development (Creswell & Creswell, 2018). This article focuses on maximizing the appropriateness and usefulness of an instrument, within the context of consumer evaluation for preK-12 educational products, by including the views of consumers (or stakeholders). We believe that *illustrating* the successful application of a two-stage sequential mixed-method, mixed-model design in designing an instrument for consumer-oriented evaluation and addressing challenges made in the process may contribute to researchers' understanding of mixed methods research, in general, and the sequential mixed-method, mixed-model design, specifically. We also believe that describing how consumers' value is captured through using mixed methods research to develop and content validate an instrument in which multiple stakeholders are included in the development and validation process can contribute to the preK-12 consumer-oriented evaluation literature.

Specifically, we aimed to maximize the appropriateness and usefulness of the DCER by including multiple viewpoints through the inclusion of nationally recognized subject matter experts, teachers, administrators, and textbook publishers and through the collection of quantitative and qualitative data. We also aimed to ensure that the criteria, descriptions of the criteria, and accompanying rating scale for the DCER represented what stakeholders deemed appropriate for digital content that supports student learning so that reviews reflect stakeholders' views and values. We believed that by including stakeholders in the process of instrument development, we would establish credibility for the reviews and stakeholders would use the reviews to select digital content that supports student learning.

Rubric and Instrument Development in Mixed Methods Research

A rubric is "[a] coherent set of rules you use to evaluate the quality of a …performance: They guide your judgments and ensure that you apply the rules consistently…" (Nitko & Brookhart, 2011, p. 512). Rubrics include two main defining aspects: criteria and performance-level descriptions (Brookhart, 2013) that need to be included in the process. Criteria that are appropriate, definable, observable, distinct from one another, complete, and able to support descriptions along a continuum of quality are desired characteristics of criteria for rubrics. Performance levels that are descriptive, clear, cover the whole range of performance, distinguish among levels, center the target performance at the appropriate level, and feature parallel descriptions from level to level are desired characteristics of descriptions of levels of performance (Brookhart, 2013). General steps in a top-down approach to developing a rubric are to create (or adapt from an existing source) a conceptual framework of criteria to be assessed, write a general scoring rubric using the dimensions and performance levels, use the rubric to assess, and adapt the rubric as needed for final use (Nitko & Brookhart, 2011).

Although approaches to instrument and rubric development are present in the literature and may include mixing methods, an intentional effort to incorporate mixed methods research may not be explicitly apparent within traditional instrument development frameworks. Onwuegbuzie, Bustamante, and Nelson (2010) and Luyt (2012) have proposed frameworks that intentionally incorporate mixed methods in the development and validation of quantitative measures. Onwuegbuzie, Bustamante, and Nelson (2010) created the *Instrument Development and Construct Validation* (IDCV)

framework—an interactive and cyclical process—consists of 10 interactive phases. In addition, Luyt (2012) extended and adapted the work of Adcock and Collier [*cf*. Adcock and Collier (2001)] for an integrated and cyclical framework for quantitative measurement development, validation, and revision. Daigneault and Jacob (2014) used the validation frameworks of Onwuegbuzie et al. (2010) and Luyt (2012) to illustrate the process of validating inferences drawn from a measure of stakeholder participation in evaluation (i.e., Participatory Evaluation Measurement Instrument; PEMI). The researchers describe the six steps in the development and validation of the PEMI: conceptualization and operationalization of stakeholder participation, applications of the initial instrument, quantitative validation of the initial instrument, mixed methods validation of the initial instrument, instrument revision, and quantitative validation of the revised instrument.

The use of mixed method research designs in instrument development and validation appears in various fields. The basic mixed method research designs are explanatory sequential design (i.e., collection and analysis of quantitative data followed by collection and analysis of qualitative data), exploratory sequential design (i.e., collection and analysis of qualitative data followed by collection and analysis of quantitative data), and convergent design (i.e., separate collection and analysis of quantitative and qualitative data) (Creswell, 2015). However, mixed methods research designs can be more complex and develop from mixed models (i.e., mix quantitative and qualitative approaches within or across stages of the research process) and mixed methods (i.e., include both quantitative and qualitative stages) (Johnson & Onwuegbuzie, 2004). For example, Curran et al. (2011), using a sequential mixed methods research design, developed and validated a set of interprofessional collaborator competencies and an accompanying assessment rubric, in both English and French, for use in health profession education. In Stage 1, the researchers conducted a comprehensive analysis of the literature on interprofessional collaborator competencies (qualitative data collection). In Stage 2, the researchers collected quantitative and qualitative data using a two-round Delphi survey, and then followed with focus group interviews (qualitative data collection). This mixed methods within-stage mixed-model design illustrates the complexity to which mixed methods research designs can be used in the development and validation of an instrument.

Other examples of the mixed methods research designs in instrument development include studies completed by Ungar and Liebenberg (2011) and Enosh, Tzafrir, and Stolovy (2015). To establish content validity evidence of a culturally sensitive measure of youth resilience (i.e., Child and Youth Resilience Measure; CYRM) across 14 international sites, Ungar and Liebenberg (2011) implemented a sequential and concurrent mixed methods research design. The researchers began with focus group interviews with youth and adults to generate questions for the instrument (qualitative data collection), and followed with the pilot administration of the CYRM (quantitative data collection) and additional individual interviews with youth who had completed the CYRM (qualitative data collection).

To develop, test, and validate a measure of social workers' exposure to client violence (i.e., Client Violence Questionnaire; CVQ), Enosh, Tzafrir, and Stolovy (2015) used a four-stage sequential mixed methods research design. In Stage 1, researchers conducted semi-structured, in-depth interviews to map the forms of client violence experienced by social workers and to develop the instrument (qualitative data collection). In Stage 2, researchers developed the CVQ based on the interviews and evaluated content validity using expert judge interrater reliability (quantitative data collection). In Stage 3, researchers examined internal consistency, content validity, and convergent validity on a second sample of social workers (quantitative and qualitative data collection). In Stage 4, researchers examined the internal reliability, factorial structure, and divergent validity of the CVQ using a second survey (quantitative data collection).

Mixed Methods Research in Consumer Evaluation and Instrument Development

In the context of consumer evaluation, elements of mixed methods research in instrument development is in various fields. For example, Loiacono, Watson, and Goodhue (2007), through conducting a literature review, soliciting criteria from Web suffers, interviewing professional Web designers and users, and testing, developed and refined an instrument for consumer evaluation of Web sites. Chang, Lai, and Chang (2007) interviewed diverse experts and surveyed consumers, administered an initial instrument, and used quantitative analyses (i.e., factor analysis, Cronbach's alpha, and item-total correlation) to reduce the number of items as well as to examine the psychometric properties of an instrument to evaluate attractiveness of passenger car forms targeted at young consumers. In addition to recent works, it is noteworthy to acknowledge the seminal work of Scriven (1974) in the development of the product evaluation checklist that took a methodology approach that extended beyond the evaluator to determine *value* for his product evaluation checklist. A revised version of the product evaluation checklist is on Western Michigan University's The Evaluation Center website (See www.wmich.edu/evaluation/checklists). Similar to Scriven and others, this study takes a stakeholder inclusive approach to establish *value*. *Valuing* was sought to create criteria deemed important to evaluating digital content by including stakeholders in the process of establishing criteria and providing content validation evidence for the rubric.

Methods

Background and Context: Evaluating Digital Content for Instructional and Teaching Excellence (EDCITE)

We conducted this study within the context of larger initiatives that resulted from a grant awarded by the Ohio Department of Education and through established partnerships with local school districts; Ohio Resource Center, The Ohio State University; and a newly formed business entity within the College of Education and Human Ecology at The Ohio State University. Funded through a state innovation initiative, EDCITE was a kindergarten through Grade 12 (K-12), higher education project that addressed the challenges school districts face in transitioning from primarily print to primarily digital curriculum. Many school districts see digital content as a way to provide more customized, personalized learning opportunities for their students. Yet, educators struggle to make informed decisions when reviewing these materials. EDCITE presented a comprehensive approach to the challenge of digital content selection by increasing teachers' capacities to reliably and consistently review digital content using a valid evaluation protocol. For information on the design framework for the EDCITE curriculum see Xui, Kim, Cheng, and Luthy (2017). Our goal was to develop an instrument for the professional development component of the grant that incorporated what stakeholders' valued in the selection of digital content that support student learning that would potentially improve district efficiency in making curriculum decisions by reducing personnel costs and shortening the time to implementation.

Research Design

The question guiding the research study was: "What are the criteria essential to evaluate digital content that supports learning in preK-12 educational settings?" To develop the rubric and content validate it, we used a sequential mixed-methods, mixed-model design (Johnson & Onwuegbuzie, 2004) that began with modified electronic Delphi (e-Delphi) surveys (Stage 1; quantitative and qualitative data collection), and then followed with focus group interviews (Stage 2; qualitative data collection) (See Figure 1). The sequential mixed-methods, mixed-model design was central to this study to provide initial content validity evidence to support the interpretation of ratings regarding the quality of available digital content (i.e., instrument

fidelity). The design also allowed us to increase the number and diversity of participants which allowed for multiple perspectives on what experts and stakeholders valued in digital content that supports student learning (e.g., participant enrichment). Furthermore, it allowed us to use data from e-Delphi surveys and focus group interviews to enhance our understanding of what experts and stakeholders valued in digital content that supports student learning (Collins, Onwuegbuzie, & Sutton, 2006).

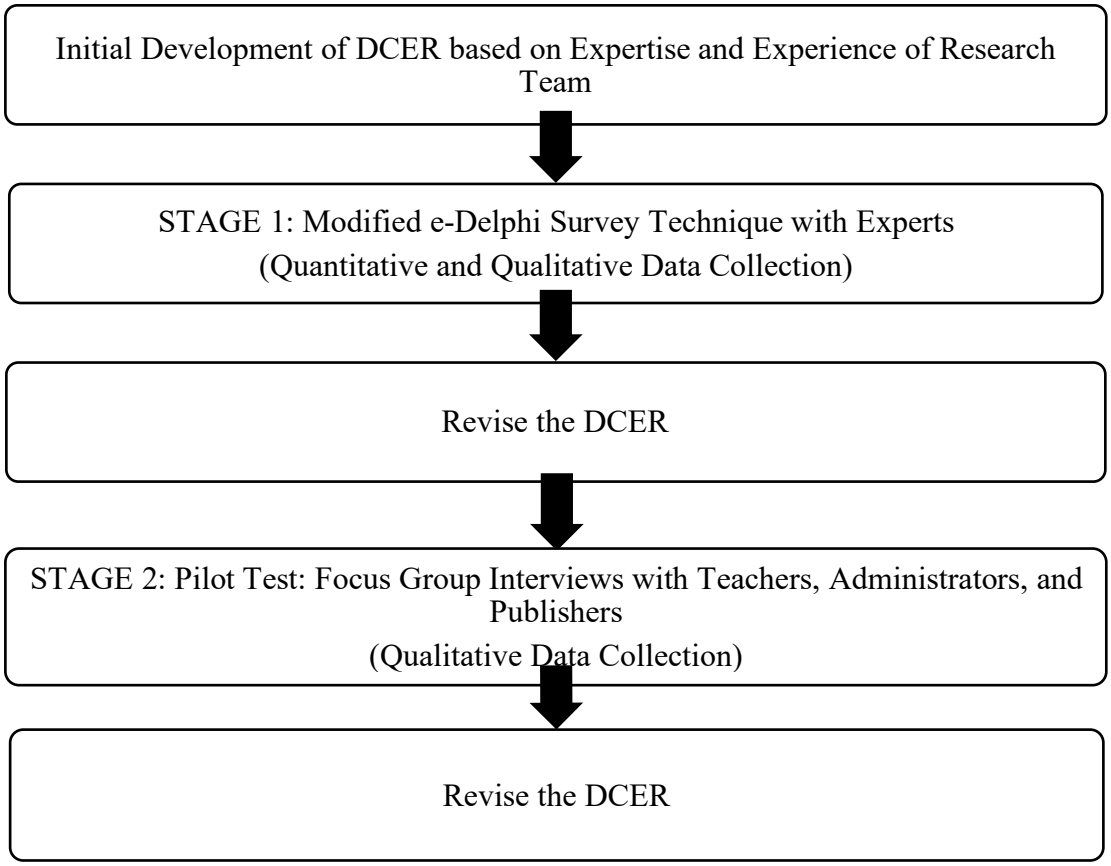

Figure 1. *Rubric development and content validation process for the Digital Content Evaluation Rubric (DCER).*

In Stage 1, we used modified e-Delphi surveys (Donohoe, Stellefson, & Tennant, 2012). A panel of subject matter experts from across the United States engaged in two rounds of modified e-Delphi surveys. The electronic questionnaires included both closed- and open-ended questions related to specific components of the rubric. In Stage 2, face-to-face focus group interviews were held separately with teachers, administrators, and textbook publishers. The use of both modified e-Delphi surveys and focus group interviews allowed us to obtain multiple viewpoints from experts and stakeholders regarding what is important to consider when evaluating digital content in preK-12 settings through the collection and analysis of quantitative and qualitative data.

Stage 1: Modified e-Delphi Survey Technique

The goals of this stage were for experts to reach consensus (i.e., at least 51% of experts in agreement) on importance of criteria, clarity of criteria descriptions, appropriateness of rating scale descriptors, and

recommendations for criteria; gather suggestions and comments; and to use the information to revise the rubric. The initial intent of Stage 1 was to collect quantitative data. Yet, the inclusion of open-ended items regarding suggestions and comments provided us with additional information to assist us with revisions.

In general, the Delphi survey technique, traditionally administered through postal mail, is a group process by which individuals with specific expertise, through an iterative process, reach consensus on some topic (McKenna, 1994). There has not been agreement on what percentage of experts in agreement is considered consensus (Hsu & Sandford, 2007). Hence, for this study, we defined consensus at 51% of experts in agreement because it represented a majority of experts in agreement on each topic. Characteristics of the Delphi technique include: (a) expert respondent anonymity, (b) controlled feedback process, and (c) use of a variety of statistical analysis techniques to interpret the data (Hsu & Sandford, 2007). It is especially beneficial to researchers when experts are geographically dispersed, which was the case for this study. However, limitations include potential of low response rates, time-consuming, potential of molding opinions, and potential of identifying general statements versus specific topic related information (Hsu & Sandford, 2007).

We used a modified version of the Delphi survey technique. We modified the technique in two ways: (a) the survey was delivered electronically and (b) initial questionnaire items were developed by the research team prior to the first round of the Delphi process. First, modifying the process to deliver the survey electronically has several benefits and limitations to both researchers and participants (Donohoe, Stellefson, & Tennant, 2012). Benefits to the research team include convenience, time and cost savings, and data management whereas limitations for participants may include internet access, technology difficulties, and data entry into computer-based data screens. Second, developing initial items for the survey, based on the expertise of the research team, reduced the number of times we asked participants to engage in the process. We believed this could potentially reduce the matriculation rate that often increases with longitudinal designs. A description of each round follows.

e-Delphi Survey, Round 1

Sampling and Participants. Using a purposive sampling approach, the research team identified 43 content area and educational technology experts, located throughout the United States, to invite to participate in the study. Participants in this study were 12 individuals who completed or partially completed items on the questionnaire. In the literature, there does not appear to be agreement on expert sample size for Delphi studies (Akins, Tolson, & Cole, 2005; Hsu & Sandford, 2007). Published Delphi studies have used expert panels ranging from 10 to more than 100 experts (Akins, Tolson, & Cole, 2005). Akins and associates (2005) found that experts with a well-defined knowledge area was stable.

Nine participants provided demographic information. Of the nine participants who provided demographic information, 56% indicated that their area of focus/expertise was mathematics with the remaining 44% indicating English language arts; science; science teacher education, learning with new technologies, and digital library builder; and technologist. The majority of respondents were in higher education (56%), followed by preK-12 (33%) and educational service center (11%). Most respondents (67%) indicated more than 10 years of experience in their current position, followed by 6-10 years (22%) and 1-5 years (11%). The majority of participants had a doctorate degree (56%), followed by master's degree (33%) and bachelor's degree (11%).

Procedure. The study was determined Exempt from the institutional review board (IRB) review process by the Office of Responsible Research Practices at The Ohio State University. In June 2014, we sent to prospective participants an email invitation, along with a link to an electronic questionnaire administered through Qualtrics, an online survey system, to participate in the study. Twenty-nine out of the 43 prospective participants (67%) opened the email invitation. Two emails bounced, one person opted out of the study, 10 participants completed the

questionnaire, two participants partially completed the questionnaire, and one participant did not respond to any items on the questionnaire. We removed the participant's record from the database who did not respond to any items on the questionnaire. In general, the response rate, as the ratio of the total number of questionnaires completed or partially completed to the total number of email invitations sent minus the bounced emails and the person who opted out of the study, was approximately 30% (12/40).

The questionnaire was available to participants for 13 days. We sent four follow-up reminders to complete the questionnaire to nonrespondents over the 13 days. It is possible that the timing for data collection (i.e., summer), incorrect email address, and delivery of the email invitation to a spam folder may have accounted for 37% of potential participants not opening the email invitation or for participants not participating in the study. Responses to the e-Delphi survey were confidential instead of anonymous; that is, researchers knew who completed each round of the Delphi survey. This information allowed researchers to conduct follow-up communications with non-respondents with the goal of increasing the response rate.

Measures. The questionnaire consisted of a minimum of 97 items within six sections. The first four sections were the main evaluative sections of the rubric. The sections were (a) content quality (5 criteria, 25 items), (b) pedagogy (6 criteria, 30 items), (c) technology use (3 criteria, 15 items), and (d) alignment to standards (4 criteria, 20 items). For each criterion, two additional items appeared based on responses. See Appendix A for sample items in the content quality section of the e-Delphi survey, including rating scale descriptors, and see Appendix C for criterion descriptors for all criteria on the final version of the rubric. The final two sections were suggestions and comments (2 items) and reviewer information (5 items), respectively.

Participants received instructions to respond to the survey by considering criteria they deemed important to assist school personnel, district personnel, and/or state department of education personnel in making decisions regarding the merit and worth of digital content to support learning in preK-12 educational settings. For each criterion, participants answered a series of questions. First, participants indicated the importance of the criterion to evaluate preK-12 digital content by selecting *Unimportant*, *Of little importance*, *Moderately important*, *Important*, or *Very important*. Second, participants indicated the clarity of the criterion description by selecting *Unclear*, *Somewhat unclear*, *Somewhat clear*, *Clear*, or *Very clear*. Third, participants indicated the appropriateness of the rating scale descriptors by selecting either *yes* or *no*. Participants who responded *no* to the appropriateness of a rating scale descriptor received a prompt to describe how to improve the rating scale descriptor or provide an alternative descriptor. Fourth, participants provided suggestion(s) or comment(s), if any, regarding the criterion. Fifth, participants recommended that the criterion be *removed*, *modified*, or *kept*. Participants who selected *modified* received a prompt to describe how to modify the criterion.

Data Analysis. We computed frequencies and percentages to determine consensus for items regarding importance of criteria to evaluate preK-12 digital content quality, clarity of criteria descriptions, and recommendation regarding retaining criteria. For items in which respondents provided constructed responses, we conducted content analysis to identify improvements to rating scale descriptors and modifications to criteria.

e-Delphi Survey, Round 2

Round 2 of the e-Delphi survey technique gave participants in Round 1 an opportunity to review group results and to make additional comments or suggestions regarding each criterion to ensure consensus. In July of 2014, we sent an email invitation to participate in Round 2 of the e-Delphi survey to the 12 participants who participated in Round 1. Eleven out of the 12 participants (91.7%) opened the email invitation. No emails bounced or persons opted out of the study, five participants completed the questionnaire, and two participants partially completed the questionnaire. Hence, seven participants either completed or partially completed the

questionnaire. In general, the response rate, as the ratio of the total number of questionnaires completed or partially completed to the total number of email invitations sent, was approximately 58% (7/12). The questionnaire was available to participants for eight days. We sent three follow-up reminders to complete the questionnaire to nonrespondents over the eight days.

The questionnaire in Round 2 consisted of frequencies for each item on the questionnaire used in Round 1 and all responses provided to the open-ended items (i.e., improvement to ratings scale descriptors, modifications to criteria, suggestions or comments regarding criteria, and additional criteria). Participants received instructions to review each section of the questionnaire (i.e., content quality, pedagogy, technology use, alignment to standards, and suggestions and comments). Then, at the end of each section, indicate whether he/she would like to provide additional comments or suggestions regarding a criterion. Participants who indicated wanting to provide additional comments or suggestions regarding a criterion received a prompt to provide the additional information. We computed frequencies and percentages for closed-ended items and conducted a content analysis for open-ended items.

Results of Stage 1: Modified e-Delphi Survey Technique

Modified e-Delphi Survey, Round 1. Tables 1 and 2 presents the frequencies and percentages of expert reviewers' responses to questions regarding (a) importance of criterion to evaluate preK-12 digital content and (b) clarity of criterion. As reflected in Table 1, at least 77% of experts indicated that each criterion was *Important* or *Very important*. As shown in Table 2, between 40% and 100% of experts indicated that each criterion description was *Clear* or *Very clear*. However, at least 77% of experts indicated that each criterion description was at least *Somewhat clear* (i.e., *Somewhat clear*, *Clear*, or *Very clear*). For three criteria (i.e., *Content Accuracy*, *Depth of Coverage*, and *Balance of Coverage*) in the *Alignment to Standards* section, 50% or less of experts indicated that the criterion was *Clear* or *Very clear*. In regards to appropriateness of rating scale descriptors for each criterion, at least 60% of experts indicated the rating scale descriptors were appropriate for each criterion. Finally, for recommendation on the inclusion of each criterion, 50% to 90% of experts indicated that each criterion be *kept.* At least 80% of experts indicated that each criterion be *modified* or *kept*.

Table 1
Frequencies (and Percentages) of Experts' Ratings of Criterion Importance to Evaluate PreK-12 Digital Content Quality

Section	*n*	Unimportant	Of Little Importance	Moderately Important	Important	Very Important
Content Quality						
Accuracy	12	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	12 (100.0%)
Clarity	12	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (25.0%)	9 (75.0%)
Identifying a sense of purpose	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (30.0%)	7 (70.0%)
Developing content ideas	9	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (22.2%)	7 (77.8%)

Section	*n*	Unimportant	Of Little Importance	Moderately Important	Important	Very Important
Assessing student progress	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (10.0%)	9 (90.0%)
Pedagogy						
Building on student ideas	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	5 (50.0%)	5 (50.0%)
Engaging students	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (30.0%)	7 (70.0%)
Promoting student thinking	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (40.0%)	6 (60.0%)
Developing discipline-based processes and practices	9	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (44.4%)	5 (55.6%)
Enhancing the learning environment	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (60.0%)	4 (40.0%)
Attention to student diversity criterion	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (60.0%)	4 (40.0%)
Technology Use						
Incorporating technology standards	10	0 (0.0%)	1 (10.0%)	0 (0.0%)	5 (50.0%)	4 (40.0%)
Engaging learners	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (30.0%)	7 (70.0%)
Design and navigation	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (30.0%)	7 (70.0%)
Alignment to Standards						
Content accuracy	10	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (30.0%)	7 (70.0%)
Depth of coverage	10	1 (10.0%)	0 (0.0%)	0 (0.0%)	4 (40.0%)	5 (50.0%)
Range of coverage	10	1 (10.0%)	0 (0.0%)	0 (0.0%)	5 (50.0%)	4 (40.0%)
Balance of coverage	9	1 (11.1%)	0 (0.0%)	1 (11.1%)	4 (44.4%)	3 (33.3%)

Table 2
Frequencies (and Percentages) of Experts' Ratings of Criterion Description Clarity

Section	*n*	Unclear	Somewhat Unclear	Somewhat Clear	Clear	Very Clear
Content Quality						
Accuracy	12	0 (0.0%)	0 (0.0%)	1 (8.3%)	4 (33.3%)	7 (58.3%)
Clarity	12	1 (8.3%)	0 (0.0%)	0 (0.0%)	8 (66.7%)	3 (25.0%)
Identifying a sense of purpose	10	0 (0%)	0 (0%)	2 (20%)	6 (60%)	2 (20%)
Developing content ideas	9	0 (0%)	2 (22.2%)	1 (11.1%)	3 (33.3%)	3 (33.3%)
Assessing student progress	9	0 (0%)	1 (11.1%)	3 (33.3%)	5 (55.6%)	0 (0%)
Pedagogy						
Building on student ideas	10	0 (0%)	0 (0%)	0 (0%)	10 (100%)	0 (0%)
Engaging students	10	0 (0%)	0 (0%)	1 (10%)	7 (70%)	2 (20%)
Promoting student thinking	10	0 (0%)	1 (10%)	1 (10%)	5 (50%)	3 (30%)
Developing discipline-based processes and practices	10	0 (0%)	0 (0%)	1 (10%)	7 (70%)	2 (20%)
Enhancing the learning environment	10	1 (10%)	0 (0%)	3 (30%)	4 (40%)	2 (20%)
Attention to student diversity	10	0 (0%)	1 (10%)	0 (0%)	7 (70%)	2 (20%)
Technology Use						
Incorporating technology standards	10	0 (0%)	0 (0%)	3 (30%)	5 (50%)	2 (20%)
Engaging learners	10	0 (0%)	1 (10%)	0 (0%)	6 (60%)	3 (30%)
Design and navigation	10	0 (0%)	0 (0%)	2 (20%)	6 (60%)	2 (20%)
Alignment to Standards						
Content accuracy	10	0	1	4	3	2

Section	*n*	Unclear	Somewhat Unclear	Somewhat Clear	Clear	Very Clear
		(0%)	(10%)	(40%)	(30%)	(20%)
Depth of coverage	10	1 (10%)	1 (10%)	4 (40%)	3 (30%)	1 (10%)
Range of coverage	10	1 (10%)	0 (0%)	3 (30%)	4 (40%)	2 (20%)
Balance of coverage	10	1 (10%)	0 (0%)	4 (40%)	3 (30%)	2 (20%)

A content analysis of the open-ended responses to suggestions for improving the rating scale descriptors and modifying the criteria within the sections are as follows. Across the 18 criteria, experts' suggestions primarily fell within the item-writing guidelines. Experts suggested revising the rating scale descriptors so that the descriptors are distinct and dissimilar across the four levels, limiting the number of ideas within descriptors, making mechanics and grammatical corrections, and replacing words for criterion with currently accepted knowledge. In addition, experts suggested that we avoid the use of acronyms in the *Technology Use* section and include rating descriptors in the *Alignment to Standards* section. Criteria recommended to be included on the rubric were:

> Expert A: "Is the reading level of the digital content clearly identified? Is academic language identified and use modeled for students? Is the content supported across platforms?"

> Expert B: "I think a category needs to be included to assess the technical aspect in terms of usability. Does the content work? Is it sluggish, low quality image, all digital learning objects are accessible."

Modified Delphi Survey, Round 2. One respondent provided additional comments or suggestions, after reviewing results from Round 1 of the modified e-Delphi survey for eight of the 18 criterion on the *Content Quality* and *Pedagogy* sections of the rubric. That is, *Accuracy*, *Identifying a Sense of Purpose*, *Developing Content Ideas*, *Assessing Student Progress*, *Building on Student Ideas*, *Engaging Students*, *Promoting Student Thinking*, and *Enhancing the Learning Environment*. The comments or suggestions were primarily utterances of those expressed in Round 1. We made modifications to criteria descriptions and ratings scale descriptors based on results from the modified e-Delphi surveys, prior to the focus group interviews.

Stage 2: Focus Group Interviews

Sampling and Participants. Stage 2 of the study was designed to examine the question of content validity from a different methodological perspective. We expanded the pool of participants to include stakeholders beyond the subject matter experts from Stage 1 of the study. In the selection process for Stage 2, we recruited and screened participants to identify those with experience in such areas as classroom instruction, educational administration, curriculum development, content selection, and digital learning. Three separate focus group interviews were conducted in fall of 2014, one with teachers, one with administrators, and one with publishers.

Participants were a purposive sample of central Ohio teachers and administrators who were identified and recruited by the Ohio Resource Center and compensated for participation. Participants were drawn from a variety of educational settings, including traditional public schools, charter schools, educational service centers, and K-12 publishing companies. Participants in the teacher focus group interview session included

nine middle and high school teachers (four science, two mathematics, one English, and one social studies) and one educational consultant, who had previously been a mathematics teacher. Teachers represented diverse schools, ranging from large urban to small rural districts. Participants in the administrator focus group interview session included five administrators from urban and suburban schools who had content backgrounds in science, social studies, career technology and career readiness, or independent. Participants in the publisher focus group interview session included nine individuals who can be described as independent sales representatives, state/national account manager, editor, or consultant. Across all focus groups, most participants were female (18 of 23 participants) and white (19 of 23 participants). While a more diverse group was desirable, the composition of the focus group reflected the demographics of the school districts participating in the EDCITE project.

Procedure and Data Analysis. An independent and local firm that specializes in market research services was engaged to conduct the focus group interviews using a facilitation guide and script developed by the authors and other members of the research team. Additionally, the research team observed the focus group through a two-way mirror and had real time access to the facilitators through an instant messaging application. This allowed the team to take field notes, respond to participants' questions and provide information to facilitators. See Appendix B for sample questions. Prior to participation in focus group interviews, participants completed and signed a non-disclosure and a consent form. Approximately two hours were allotted for each rubric content validation focus group interview session.

Each focus group session was designed to elicit both individual and group feedback. Individual responses were collected through written feedback, which included coded responses and open-ended comments. At the start of the focus group, each participant was given a print version of the rubric and asked to use a coding system (developed for the focus group) to provide their individual ratings for specific components of the rubric (i.e., rubric criteria, descriptions for each criterion, rating scale, and format and organization). In addition to marking the text with specific codes, participants also indicated their positive and negative responses, noted questions they had, and provided suggestions for revisions and potential application. All copies of the individual feedback documents were collected. These artifacts provided a "document map" of each participant's response to the rubric, as codes and comments were made directly onto the rubric itself. From the individual responses, collective responses were generated, both within and across cohorts, by aggregating codes and comments from each of the individual responses into a single document. This approach highlighted areas of consensus across the three groups and revealed areas where participants' responses diverged. Areas of divergence were then flagged for further examination.

Group interviews were conducted following the individual response portion of the focus group. Group responses were captured through facilitated discussions, which were recorded and later transcribed. Focus group interview questions addressed overall quality and usefulness of the rubric, strengths and limitations of the criteria, clarity of wording, and gaps or missing content. Transcripts from the focus group interviews were completed within one week of the sessions. Through open coding (Strauss & Corbin, 1998), we extracted and integrated participants' comments from the transcripts with other data collected from the sessions.

Results of Stage 2: Focus Group Interviews

We found positive responses across all three groups for the rubric criteria and the overall quality of the rubric, as indicated by individual responses and group comments. Teachers exhibited the most favorable responses, followed by administrators, and then publishers. Across the four sections of the rubric, *Alignment to Standards* received the most positive response with all three groups awarding high marks.

Pedagogy, while receiving positive responses overall, showed the most disagreement across groups. In particular, participants disliked the language used in that section, citing obscure references, lack of relevance, and ambiguous language in their comments. In addition to *Pedagogy*, two other areas of the rubric were reviewed negatively for language—*Content Quality* and *Technology*. Focus group participants indicated that the language was not "user-friendly" and seemed too "academic" for the criteria descriptors and rubric rating scale descriptors. Participants expressed a need for greater clarity in the descriptors in order to draw more distinction between criteria and across the rating scale.

In comparing results across groups, the most noticeable differences in responses existed between teachers and publishers. Teachers' provided the most positive responses across the three groups, while publishers expressed more negative responses than any other group. This divergence of responses was most notable in the *Pedagogy* section.

Analysis of data from the focus group interviews revealed broad categories of agreement and disagreement, which were used to guide further revision of the rubric criteria and descriptions. Collective ratings and responses for each criterion, along with the related descriptors, were reviewed to determine what revisions, if any, were needed for each criterion. Final version of the rubric can be found in Appendix C.

Discussion

This article describes the application of a sequential mixed-method, mixed-model design in the development and initial content validation of an instrument for use by teachers to evaluate preK-12 digital content. With a focus on use of mixed methods research for the purpose of instrument fidelity, this study took an approach to rubric development that included multiple perspectives and viewpoints from various stakeholders (i.e., teachers, administrators, and textbook publishers) and experts through quantitative and qualitative data collection. The inclusion of experts as well as multiple stakeholders was intentional, given a focus on use of the rubric by teachers, use of reviews by teachers and administrators, and potential unintended consequences of the reviews to textbook publishers.

Criteria Essential to Evaluate Digital Content That Supports Learning in preK-12

Stage 1: Modified e-Delphi Survey Technique. The use of both closed- and open-ended response options during both rounds of the e-Delphi survey provided information regarding experts' consensus on importance of each criterion, clarity of each criterion description, appropriateness of rating scale descriptors, and recommendations that enabled us to refine the rubric, prior to focus group interviews. Results from the e-Delphi survey allowed us to understand criteria experts—dispersed throughout the United States—valued to evaluate digital content that would support student learning, providing a broader perspective and view than those possibly held locally. The use of closed-ended response options only during the e-Delphi survey rounds would have only provided information regarding the frequencies and percentages of experts' selection of response options for determining consensus. Although determining consensus is at the heart of the e-Delphi survey technique, the inclusion of open-ended responses allowed experts to provide us with information regarding how to improve rating scale descriptors or provide alternative descriptors, how to modify criterion, and suggestions and comments for each criterion.

Findings from the first round of the e-Delphi survey provided evidence that overall, experts reached consensus (i.e., at least 51%) on importance (i.e., *Important* or *Very important*) of each criterion, appropriateness of rating scale descriptors of each criterion, and modifying or keeping each criterion. The finding regarding consensus of importance of each criterion was supported with experts' consensus to modify or keep the criterion. In retrospect, the inclusion of a prompt following experts' selection of an importance rating other than *Important* or *Very important* or selection of a clarity rating other than *Clear* or *Very clear* would have allowed experts an opportunity to explain their rating close to the

item instead of in the comments and suggestions section for each criterion.

Findings from the second round of the e-Delphi survey supported the consensus, and suggestion and comments from the first round of the e-Delphi. We used comments and suggestions provided on the open-ended questions for both rounds of the e-Delphi survey to improve the wording and language used for criterion descriptions and ratings scale descriptors prior to the focus group interviews. The information we obtained from both rounds of the e-Delphi survey provided some evidence to support content validity that lends support to instrument fidelity.

Findings from both rounds of the e-Delphi survey informed the design of the focus group interview protocol. Specifically, we used findings to draft the discussion guide and questions, the format and organization of the rubric, and the mock-up of a product review. In creating focus group interview questions, we gave particular attention to the terminology related to subject-specific content, rubric criteria, and the rubric performance language.

Stage 2: Focus Group Interviews. Focus group interviews, which included stakeholders representing school-based and commercial interests, served multiple purposes and allowed for the inclusion of diverse perspectives in the content validation process. By including multiple stakeholders in the content validation process, we aimed to confirm what experts deemed important and identify discrepant areas where views differed, as well as to ensure that the voices of those most affected by the review process and the outcomes of the reviews were included. Additionally, since focus group interviews were face-to-face, other aspects of the rubric related to and impacting users' ability to interpret content, including organization and layout, its application and ease of use, and the review process (i.e., number of reviewers per product, reviewer training) were addressed. Findings that emerged from the focus group interviews were *likes*, *dislike*s, and *suggestions* for improving the rubric. Findings implied what teachers, administrators, and textbook publishers valued in content and criteria to evaluate digital content.

Implications of Using Mixed Methods Research in DCER Development and Challenges

The collection of quantitative and qualitative data through use of both the modified e-Delphi surveys and focus group interviews enabled us to establish content validity evidence for the DCER, which in turn supports instrument fidelity. Information obtained from subject matter experts provided a foundation for the focus group interview protocol. We integrated data collected from both stages of the sequential mixed-method, mixed-model design regarding what is important to consider when evaluating digital content for preK-12 students to refine the DCER. Mixed methods research allowed us to collect both quantitative and qualitative data to understand what content experts, teachers, administrators, and publishers value when selecting digital content to enhance student learning. It also allowed us to include multiple voices and perspectives in the validation process with an aim of use of the rubric by teachers through professional development and use of reviews created from the rubrics for teachers and administrators.

We encountered two challenges in using the two-stage sequential mixed-method, mixed-model design: time and plethora of data. To fulfill grant requirements, we had a short window of time (i.e., approximately eight weeks) to complete development and content validation of the rubric prior to the first professional development workshop with teachers. The turnaround time between stages of the mixed methods research was a challenge given the amount of data collected from both stages of the design and analysis of data. The collection of quantitative data provided a snapshot into consensus and the collection of qualitative data provided explanation. Although we encountered these challenges, we felt that the use of the two-stage sequential mixed-method, mixed-model design provided an opportunity to include multiple viewpoints and perspectives, which would not only provide content validity evidence to support the use of the instrument, but also demonstrate its potential application as an instrument to be used by teachers.

Limitations and Future Studies

Several limitations exists for this study. First, the timing of the study (i.e., summer) possibly decreased response rates for the e-Delphi surveys because most educators were on break. However, given our time constraint to develop and content validate the instrument for professional development that began in fall of 2014, it was not possible to conduct the e-Delphi study later. Second, there was a lack of diverse demographic representation on the focus group interview panels. For example, although teachers for the focus group interview were screened for participation, there was no representation from elementary school teachers on the focus group interview panel. The lack of representation from elementary school teachers has potential implications for instrument fidelity. That is, what elementary school teachers valued in the content and criteria to evaluate digital content was not included.

Future studies that use a sequential mixed-method, mixed-model design for the purpose of instrument fidelity can build upon the limitations this study. Additional studies can confirm the extent to which the criteria are appropriate for different types of digital content, using a purposive and diverse sample of teachers.

References

Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, *95*, 529-546.

Akins, R. B., Tolson, H., & Cole, B. R. (2005). Stability of response characteristics of a Delphi panel: Application of bootstrap data expansion. *BMC Medical Research Methodology, 5*(37). Available online: http://www.biomedcentral.com/1471-2288/5/37

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria, VA: ASCD.

Chang, H.-C., Lai, H.-H., & Chang, Y.-M. (2007). A measurement scale for evaluating the attractiveness of a passenger car form aimed at young consumers. *International Journal of Industrial Ergonomics, 37*, 21-30. doi: 10.1016/j.ergon.2006.09.014

Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed-methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal, 4*, 67-100.

Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Thousand Oaks, CA: Sage.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Thousand Oaks, CA: Sage.

Curran, V., Hollett, A., Casimiro, L. M., Mccarthy, P., Banfield, V., Hall, P., …, Wagner, S. (2011). Development and validation of the interpersonal collaborator assessment rubric (ICAR). *Journal of Interprofessional Care, 25*, 339-344. doi: 10.3109/13561820.2011.589542

Daigneault, P.-M., & Jacob, S. (2014). Unexpected but most welcome: Mixed methods for the validation and revision of the Participatory Evaluation Measurement Instrument. *Journal of Mixed Methods Research, 8*, 6-24. doi: 10.1177/1558689813486190

Donohoe, H., Stellefson, M., & Tennant, B. (2012). Advantages and limitations of the e-Delphi technique: Implications for health education researchers. *American Journal of Health Education, 43*, 38-46. doi: 10.1080/19325037.2012.10599216

Enosh, G., Tzafrir, S. S., & Stolovy, T. (2015). The development of Client Violence Questionnaire (CVQ). *Journal of Mixed Methods Research, 9*, 273-290. doi: 10.1177/1558689814525263

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston, MA: Pearson.

Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation, 12*(10). Available online: http://pareonline.net/getvn.asp?v=12&n=10

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*, 112-133. doi: 10.1177/1558689806298224

Loiacono, E. T., Watson, R. T., & Goodhue, D. L. (2007). WebQual: An instrument for consumer evaluation of web sites. *International Journal of Electronic Commerce, 11*(3), 51-87. doi: 10.2753/JEC1086-4415110302

Luyt, R. (2012). A framework for mixing methods in quantitative measurement development, validation, and revision: A case study. *Journal of Mixed Methods Research, 6*, 294-316. doi: 10.1177/1558689811427912

McKenna, H. P. (1994). The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing, 19*, 1221-1225.

Mertens, D. M., & Hesse-Biber, S. (2013). Mixed methods and credibility of evidence in evaluation. In D. M. Mertens & S. Hesse-Biber (Eds.), *Mixed methods and credibility of evidence in evaluation. New Directions for Evaluation, 138*, 5-13.

Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). Boston, MA: Pearson.

Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research, 4,* 56-78. doi: 10.1177/1558689809355805

Scriven, M. (1974). Prologue: Standards for the evaluation of educational programs and products. In G. D. Borich (Ed.), *Evaluating educational programs and products* (pp. 5-24). Englewood Cliffs, NJ: Educational Technology.

Strauss, A. & Corbin, J. (1998). *Basics of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, and applications* (2nd ed.). San Francisco, CA: Jossey-Bass.

Ungar, M., & Liebenberg, L. (2011). Assessing resilience across cultures using mixed methods: Construction of the Child and Youth Resilience Measure. *Journal of Mixed Research Methods Research, 5*, 126-149. doi: 10.1177/1558689811400607

Xie, K., Kim, M. K., Cheng, S.-L., & Luthy, N. C. (2017). Teacher professional development through digital content evaluation. *Educational Technology Research and Development, 65*, 1067-1103.

Appendix A: Sample Items on the e-Delphi Survey

Directions: The questionnaire consists of six sections (i.e., content quality, pedagogy, technology use, alignment to standards, suggestions and comments, and reviewer information). The question guiding the rubric development and validation process is "What are the criteria essential to evaluate digital content that supports learning in preK-12 educational settings?" So, as you respond to questions within the sections, please consider criteria that are important to assist school personnel, district personnel, and/or state department of education personnel in making decisions regarding the merit and worth of digital content to support learning in preK-12 educational settings.

Section I. Content Quality

1.1. Accuracy criterion: Content is consistently accurate and current. It is free from errors, misleading statements, or statements that may reinforce commonly held student misconceptions.

1.1.1. To what extent is the accuracy criterion important to evaluate preK-12 digital content quality? Unimportant, Of little importance, Moderately important, Important, Very important, Of little importance

1.1.2. To what extent is the accuracy criterion description clearly described? Unclear, Somewhat unclear, Somewhat clear, Clear, Very clear

1.1.3. Accuracy rating scale descriptors

(1) Insufficient: Content contains inaccuracies and is out-of-date. It contains misleading statements that may reinforce commonly held student misconceptions.
(2) Minimal: Content is not inaccurate but may be lacking important details and may not be current. It is free from errors, misleading statements, or statements that may reinforce commonly held student misconceptions.
(3) Adequate: Content is generally accurate and current. It is free from errors, misleading statements, or statements that may reinforce commonly held student misconceptions.
(4) Strong: Content is consistently accurate and current. It is free from errors, misleading statements, or statements that may reinforce commonly held student misconceptions.
Are the rating scale descriptors appropriate for the accuracy criterion? No or Yes
If "No" selected: Please describe how the rating scale descriptors can be improved or provide alternative descriptors for the accuracy criterion.

1.1.4. What suggestion(s) or comment(s), if any, do you have regarding the accuracy criterion?

1.1.5. I recommend that the accuracy criterion be (a) removed, (b) modified, or (c) kept.

If "Modified" selected: Please describe how the accuracy criterion should be modified.

Appendix B. Focus Group Rubric Review

A. First, how useful is the information provided in the rubric for selecting digital content?

B. What do you think about the rubric, overall in terms of evaluating the quality of digital content?

PROBES:

a. Use? Practical? In what ways?
b. Strengths of the rubric. How well does it align with other criteria you think is important in making these decisions?
c. Anything missing?
d. How would you describe the clarity of the wording of the items? (Is there any room for misinterpretation or multiple interpretations? If so, what/where?)

C. Looking at all the different sections / sub-sections, are there any categories that are missing when evaluating content? Which ones?

D. Review the rubric by section:

a. What did you like about this section?
b. What did you dislike about this section?
c. What was confusing or did you have questions about?
d. Anything missing that should be included in this section?

Appendix C. Digital Content Evaluation Rubric [Abridged Version]

Section I: Content Quality

Criterion	Insufficient	Minimal	Adequate	Strong
1.1 Accuracy: Content contains currently accepted knowledge and is free from errors, misleading statements, or statements that may reinforce commonly held student misconceptions.				
1.2 Clarity: Content is expressed in clear concise language that is appropriate for the discipline and the developmental level of the student.				
1.3 Identifying a sense of purpose: Content effectively conveys an overall sense of purpose and direction that is clear to students and involves students in a logical or strategic sequence of learning activities.				
1.4 Developing content ideas: Content builds concept attainment by almost always presenting ideas in a logical sequence, representing ideas accurately and comprehensibly, modeling skills and how to use content knowledge, and providing tasks/questions/problems that allow students to apply knowledge in a variety of situations.				
1.5 Assessing student progress: Content aligns assessment activities to learning objectives; includes assessment tasks that inform instruction and provide opportunities for students to apply ideas and skills as evidence of learning.				

Section II: Pedagogy

Criterion	Insufficient	Minimal	Adequate	Strong
2.1 Building on student ideas: Resource effectively specifies necessary prerequisite knowledge and skills and provides opportunities to connect to prior knowledge and skills (e.g., alerts teachers to commonly held student ideas and misconceptions, provides strategies for uncovering student ideas prior to introducing new material, addresses commonly held student ideas and misconceptions).				
2.2 Engaging students: Resource provides varied contexts and approaches and firsthand experiences (e.g., inquiry investigations, interviews, real-world problem solving).				
2.3 Promoting student thinking: Resource encourages students to explain, clarify, justify (cite evidence), and represent their ideas; includes tasks/questions/problem sequences to guide student interpretation and reasoning; and encourages students to monitor their progress and think about what they have learned.				
2.4 Developing discipline-based processes and practices: Resource provides multiple and varied opportunities for students to engage in discipline-based processes (e.g. mathematical practices, scientific inquiry and technological design, increasing complexity of text and using evidence from the text to support responses) and promotes behaviors and skills that sustain learning beyond the course/curriculum.				
2.5 Enhancing the learning environment: Resource supports all students by setting high expectations, encouraging curiosity and questioning, and enabling success.				
2.6 Attention to diversity: Resource is culturally accurate, current, and free of bias. Resource supports all students by fostering a sense of inclusion through learning activities that address a variety of learning styles and preferences.				
2.7 Differentiation: Resource differentiates instruction for a diverse population of learners (e.g., English language learners, students with disabilities, etc.)				

Section III: Technology Use

Criterion	Insufficient	Minimal	Adequate	Strong

3.1 Creativity and innovation: Resource utilizes technology in ways that lead to student skills in creativity and innovation. (See ISTE-S Standards 1-4)*

3.2 Communication and collaboration: Resource utilizes technology in ways that lead to student skills in communication and collaboration. (See ISTE-S Standards 1-4)*

3.3 Research and information fluency: Resource utilizes technology in ways that lead to student skills in research and information fluency. (See ISTE-S Standards 1-4)*

3.4 Critical thinking, problem solving, and decision-making: Resource utilizes technology in ways that lead to student skills in critical thinking, problem solving, and decision-making. (See ISTE-S Standards 1-4)*

3.5 Design and navigation: Resource is presented in an easy to use manner and provides guidance for simple navigation and supports the learning objectives.

Section IV: Mathematics Standards Alignment

Criterion	Insufficient	Minimal	Adequate	Strong

4.1 Content Alignment

Content Standards Alignment: Aligns with Ohio's New Learning Standards: K-12 Mathematics. The material is at the correct grade level and aligns with the intent of the Clusters and Standards.

4.2 Depth of Coverage

- Conceptual Understanding of Mathematics Development: Develops mathematical thinking (e.g., embraces developing conceptual understanding, developing and using strategies which then evolve into fluency with skills and procedures).
- Standards for Mathematical Practice Development: Develops essential mathematical habits of mind, specifically these practices:
    - Reason abstractly and quantitatively.
    - Attend to precision.
    - Look for and make use of structure.
    - Look for and express regularity in repeated reasoning.

4.3 Range of Coverage

- Optional Content Standards (+): Includes topics that are reserved for students who plan to take advanced mathematics courses; these topics are denoted by a (+) in the Ohio's New Learning Standards for Mathematics.
- Problem Solving and Rich Problems: Provides rich problems or open-ended questions. Suggestions for student reflection and ideas for follow-up or extension are provided. Guidance is provided to help students formulate and carry out their own investigations including formally communicating and defending their results.
- Contextual Learning: Content is framed in a context that is relevant to students and significant from a global perspective, and students are required to communicate (data/findings/research) to an external audience.
- Standards for Mathematical Practice Development: Develops essential mathematical habits of mind, specifically these practices:
    - Make sense of problems and persevere in solving them.
    - Construct viable arguments and critique the reasoning of others.
    - Model with mathematics.
    - Use appropriate tools strategically.

4.4 Balance of Coverage

- Critical Areas of Focus: Aligns well with the Critical Area(s) of Focus
- Appropriate Amount of Review
- Appropriate Extension(s) of Ideas

Section IV: Science Standards Alignment

Criterion	Insufficient	Minimal	Adequate	Strong

4.1 Content Alignment

Resource is aligned to *Ohio's New Learning Standards: K-12 Science*

4.2 Depth of Coverage

- The resource integrates scientific inquiry and application skills as identified in *Ohio's New Learning Standards: K-12 Science*:
    - Identify questions and concepts that guide scientific investigations;
    - Design and conduct scientific investigations;
    - Use technology and mathematics to improve investigations and communications;
    - Formulate and revise explanations and models using logic and evidence (critical thinking);
    - Recognize and analyze explanations and models; and
    - Communicate and support a scientific argument.
- The resource includes a range of student tasks that reflect the Expectations for Learning: Cognitive Demands identified in *Ohio's New Learning Standards: K-12 Science*:
    - Recalling Accurate Science
    - Interpreting and Communicating Science Concepts
    - Demonstrating Science Knowledge
    - Designing technological/Engineering Solutions Using Science Concepts

4.3 Range of Coverage

The resource provides contextual learning examples and meaningful application as included in *Ohio's New Learning Standards: K-12 Science*.

4.4 Balance of Coverage

- The resource includes a balance of opportunities for students to practice all Cognitive Demands for Science
- Cognitive Demands identified in *Ohio's New Learning Standards: K-12 Science*:
    - Recalling Accurate Science
    - Interpreting and Communicating Science Concepts
    - Demonstrating Science Knowledge
- Designing technological/Engineering Solutions Using Science Concepts
- The resource includes adequate review to link new concepts to previous learning and refrains from extensive re-teaching of concepts addressed in earlier grades.

Section IV: English Language Arts Standards Alignment

Criterion	Insufficient	Minimal	Adequate	Strong
4.1 Content Alignment Resource topics address key content as described in *Ohio's New Learning Standards: K-12 English Language Arts*. 4.2 Depth of Coverage Resource integrates developmentally appropriate levels of text complexity, academic vocabulary and rigorous student tasks as related to the target grade level and standards. 4.3 Range of Coverage The resource readings, lessons, activities and tasks cover the full range of standards as described in *Ohio's New Learning Standards: K-12 English Language Arts;* appropriate to the target grade level. 4.4 Balance of Coverage - The resource shifts the balance of texts and instructional time to include equal measures of literary and informational text. - Questions and tasks cultivate students' abilities to ask and answer questions based on the text are balanced throughout the course.				

Section IV: Social Studies Standards Alignment

Criterion	Insufficient	Minimal	Adequate	Strong
4.1 Content Alignment Digital content fully aligns with the theme, all topics, and all content statements in *Ohio's New Learning Standards in Social Studies* for the targeted grade level or course. 4.2 Depth of Coverage				

Digital content consistently incorporates a rich variety of developmentally appropriate academic vocabulary, text complexity and cognitive demands throughout the resource and challenges students to meet or exceed the level of understanding articulated in the *Expectations for Learning* for the targeted course or grade level.

4.3 Range of Coverage

Digital content (activities, readings, questions, tasks, etc.) consistently promotes building knowledge through the use of content-rich nonfiction or informational texts, and promotes reading and writing grounded in evidence from texts across the range of standards at the targeted grade level or course.

4.4 Balance of Coverage

- Topics and concepts emphasized in the digital content (activities, readings, questions, tasks) are consistent with the emphasis found in Ohio's New Learning Standards in Social Studies and course/grade level.
- Social Studies skills (historical thinking, spatial thinking, civic participation, economic decision making, and financial literacy skills) are integrated, where appropriate, throughout the digital content.