A Performance Assessment of Teachers' TPACK Using Artifacts from Digital Portfolios

Joshua M. Rosenberg (jrosen@msu.edu)
Spencer P. Greenhalgh (greenha6@msu.edu)
Matthew J. Koehler (mkoehler@msu.edu)
Michigan State University, United States of America

Abstract: Researchers have employed many different methods of measuring teachers' Technological Pedagogical Content Knowledge (TPACK). Existing measures of TPACK have typically focused on teachers' self-report of their understanding, and relatively few approaches directly measure teacher performances. Moreover, to date, no performance assessments of teachers' TPACK have used teachers' digital portfolios or the work samples (or artifacts) included in teachers' portfolios. In this paper, we build on our initial attempt (Koehler, Rosenberg, Greenhalgh, Zellner, & Mishra, 2014) to test the validity and reliability of a performance assessment of the TPACK present in teachers' portfolio artifacts by focusing on two specific types of artifacts: Dream IT (a type of grant proposal) and a sustainable technology initiative. We report the reliability of our attempts to code levels of TPACK evident in these artifacts and discuss the development of more robust performance measures of teachers' TPACK.

Introduction

Technological Pedagogical Content Knowledge (TPACK) is a framework commonly used in both scholarly and practical settings to describe the knowledge needed to effectively teach with technology. As the influence of this framework has expanded, researchers have developed a number of measures for teachers' TPACK; these measures have contributed to a better understanding of how TPACK can develop over time as a result of experience, resources, interventions, or other mechanisms of change. These measures are varied: Interview protocols, open-ended questionnaires, self-report surveys, observations, and performance assessments have all been used to assess teachers' knowledge (Koehler, Shin, & Mishra, 2011).

Self-reports are a particularly common method for assessing teachers' knowledge (Abbitt, 2011; Koehler, Shin, & Mishra, 2011). They are easy to administer, and the data collected from them is easy to analyze, especially when used in pre-post study designs. However, self-reports may exhibit poor construct validity (Creswell, 2013). That is, researchers may, in fact, be measuring something other than teachers' TPACK, stymying efforts to understand and enhance teachers' capacity to integrate technology. For example, self-reports may assess teachers' confidence about teaching with technology rather than their actual knowledge or behavior (Harris, Grandgenett, & Hofer, 2010). Performance assessments of TPACK are attempts to avoid this problem by measuring teachers' competence based on authentic instances of integrating technology with a specific pedagogy and content area. They can complement self-report and other measures of TPACK because they directly measure the action resulting from knowledge rather than taking indirect measures that may inadvertently capture teachers' confidence instead. However, there is not yet a performance assessment for TPACK for use with either teachers' digital portfolios or the specific work samples or artifacts included in portfolios.

The overarching purpose of this study, then, is to test the validity and reliability of a performance assessment of the TPACK evidenced in artifacts included in teachers' portfolios.

Literature Review

We review four areas of prior research in order to establish the need for this study: First, we examine the TPACK framework; then, we review methods of assessing it. Next, we cover specific attempts to develop performance assessments for it. Finally, we address our initial attempt to use digital portfolios as a performance assessment for teachers' TPACK.

Preview version of this paper. Content and pagination may change prior to final publication.

TPACK

The TPACK framework suggests a way of thinking about the knowledge teachers need in order to teach well with technology. It is not enough for teachers to simply know about technology; instead, effectively using technology in teaching and learning requires alignment with the content being taught and the pedagogical practices being used. Mishra and Koehler (2006) suggested that the recent importance of digital technologies in education necessitated changes to Shulman's (1986) framework of the knowledge needed by teachers. Shulman identified three bodies of knowledge needed by teachers: knowledge of content, knowledge of pedagogy, and, most importantly, knowledge of how content and pedagogy reciprocally affect each other. Mishra and Koehler argued that the knowledge needed by teachers includes technology in addition to pedagogy and content. Furthermore, just as Shulman considered the interactions between pedagogy and content to be essential, Mishra and Koehler asserted the importance of Technological Pedagogical Content Knowledge, or the interactions between technology, pedagogy, and content. See Figure 1 for a depiction of the components in the TPACK framework and their relationship to one another.

In addition to recognizing three distinct bodies of knowledge and the interactions between them, the TPACK framework also takes context into account (Kelly, 2008; Koehler & Mishra, 2008; Porras-Hernandez & Salinas-Amescua, 2013; Rosenberg & Koehler, 2014). That is, TPACK is not a sterile, theoretical treatment of knowledge so much as a type of knowledge situated in the real-world environments in which teachers work. Students, resources, and supports influence how TPACK develops among teachers as well as how teachers use their TPACK in their planning and instruction.

Methods of Assessing TPACK

Researchers have tried to measure TPACK since its introduction, and there are now multiple methods of assessing teachers' TPACK (Abbitt, 2011; Koehler, Shin, & Mishra, 2012). For example, a number of researchers have explored the potential of using semi-structured interviews to assess teachers' TPACK (Harris, Grandgenett, and Hofer, 2012; Niess, Lee, Sadria, & Suharwoto, 2006). An alternative to interviews is questionnaires, which allow teachers to respond in an open-ended manner. This method has been used by Niess et al. (2006) and by So and Kim (2009). Self-report surveys are also a common measure, with the survey by Schmidt et al. (2009) being widely used. Archambault and Crippen (2009) also developed a self-report survey to assess the TPACK of teachers in online course settings. Observation protocols, such as the one used by Niess et al. (2006), are promising for research as well as practice—administrators and professional development providers can use these to provide ongoing, or formative, assessment.

Finally, some researchers have explored assessing authentic tasks and activities for evidence of teachers' TPACK, a practice referred to as performance assessment. Tasks and artifacts assessed in this way include lesson planning and instruction (Akcaoglu, Kereluik, & Casperson, 2011), instructional plans (Trip, Graham, & Wentworth, 2009), and lesson plans. Harris, Grandgenett, and Hofer (2010) analyzed teachers' lesson plans using a rubric for which they established validity and reliability properties. We discuss this rubric in detail in the next section.

Performance Assessments for TPACK

Performance assessments evaluate teachers' TPACK by directly examining their performance in complex, authentic tasks (or representations of complex, authentic tasks). Some ask teachers to create and maintain a set of artifacts in a portfolio or journal, while others consist of scenario- or problem-based questions that are assessed by experts. Of six performance assessments of TPACK reviewed by Koehler et al. (2012), only one demonstrated validity and reliability: Harris et al.'s (2010) adaptation of the Technology Integration Assessment Instrument (TIAI; Britten & Cassady, 2005).

Preview version of this paper. Content and pagination may change prior to final publication.

The TIAI was originally designed for evaluating technology-integrated lesson plans generally, but Harris et al. (2010) modified it for the specific purpose of assessing TPACK. During the modification process, they solicited advice on how closely aligned it was with TPACK from six TPACK researchers and twelve local technology-using teachers and administrators. They then revised the rubric and then asked experienced, in-service, technology-using teachers to assess pre-service teachers' technology-integrated lesson plans across varying grade levels and content areas. They did this by training 15 experienced teachers for a 6-hour period in the use of the rubric, after which the teachers used the rubric to code pre-service teachers' lesson plans.

Unlike many TPACK assessments, Harris et al. (2010) thoroughly demonstrated the validity and reliability of this rubric. The rubric demonstrated construct validity through the approval of the six TPACK experts. It demonstrated face validity in feedback from the 15 experienced teachers, who supported its use to assess TPACK from lesson plans. Harris et al. (2010) demonstrated reliability through the calculation of four statistics. First, they used interrater reliability using a hierarchical linear model. Teachers were the level-2 variable, and their scores were the level-1 variables. An intraclass correlation coefficient explained how the variability across teachers' scores was due to teacher differences. Second, percentage agreement was calculated by pairing raters. Third, Cronbach's alpha demonstrated the internal consistency of participants' scores between each of the four rows of the rubric. Fourth, they computed test-test reliability by having raters code three lesson plans in different content areas and grade levels one month after their initial coding, followed by the calculation of percentage agreement across the two periods.

These validity and reliability measures demonstrated that this rubric could be used with trained raters to assess the levels of TPACK demonstrated in pre-service teachers' lesson plans. Harris et al. (2010) cautioned that the rubric had not yet been used with in-service teachers or with experienced lesson plans. They also suggested that lesson plans considered for coding with this rubric be written in detail so that raters can make well-informed choices. Finally, they suggested that supplementing coded lesson plans with interviews might help triangulate teachers' TPACK. In the next section, we discuss how the research by Harris et al. informed our attempt to evaluate another kind of performance assessment: artifacts or created works in teachers' digital portfolios.

Initial Attempt to Develop Performance Assessment for Teachers' TPACK Using Artifacts from Digital Portfolios

Our experience designing and teaching a master's level teacher portfolio course led us to explore whether we could assess the levels of teachers' TPACK through the portfolios that teachers create. Our initial attempt (Koehler et al., 2014) to test the validity and reliability of Harris et al.'s (2010) performance assessment when applied to portfolio artifacts was guided by four research questions:

1. What skills do teachers choose to represent in their portfolios?
2. Do teachers connect their technology competencies to their content areas or pedagogical approaches?
3. How amenable are the artifacts teachers submit to the TIAI?
4. What levels of TPACK are evidenced by the artifacts teachers submit?

Guided by these research questions, we coded a total of 75 artifacts from 25 portfolios; although each of these portfolios had more than three artifacts, we selected from each one the three that seemed most amenable to analyzing teachers' TPACK through the adapted rubric. We did not analyze "fit," which was part of Harris et al.'s (2010) rubric, because we believed that the TPACK category included the same construct. We trained with the rubric by using it to evaluate a sample of the previously selected teachers' artifacts and then began coding. Because Harris et al. (2010) had already demonstrated the validity of the measure, we concentrated on calculating reliability statistics. We used Cohen's kappa, which takes into account chance agreement and percentage agreement.

Sample	Variable	Percentage Agreement	Cohen's Kappa and Interpretation
n=75 various artifacts	TCK	45.3%	0.234 (slight)
	TPK	41.3%	0.157 (poor)
	TPACK	36.0%	0.120 (poor)

Table 1: Reliability for our initial attempt to infer teachers' TPACK from digital portfolios using Harris, Grandgenett, and Hofer's (2010) rubric; we interpreted Cohen's kappa using guidelines based upon a review of

Preview version of this paper. Content and pagination may change prior to final publication.

current research by Sim and Wright (2005): 0 - .20 = poor; .21 - .40 = slight; .40 - .60 = moderate; .61 - .80 = substantial; .81 - 1.00 = excellent.

Based on these discouraging results, we identified three potential areas for improvement: we could (1) limit future assessment to common artifacts for coders or artifacts designed to elicit TPACK; (2) use a different rubric or develop a new rubric ourselves; or (3) select different raters, perhaps those with content-area expertise (Koehler et al., 2014).

Purpose and Research Questions

The purpose of this study is to test the validity and reliability of a performance assessment for measuring the TPACK evidenced by artifacts included in teachers' portfolios. We chose to focus on artifacts (rather than the rubric or the raters) as a specific area for improvement because the rubric had already been extensively tested (Harris et al., 2010), and the raters were deemed as competent. Furthermore, this decision aligned with Harris et al.'s (2010) advice that those using their rubric to code lesson plans make sure the artifacts have sufficient detail; we had experienced this problem in our initial attempt, as the artifacts we coded typically did not include detailed information on TPACK. Therefore, we limited our focus to two kinds of artifacts that had more detail than the artifacts we examined in our previous attempt. However, we recognize that this decreases the external validity of the rubric. Teachers' portfolios do not consist of just one artifact or one type of artifact: They consist of varied artifacts, and coding only two artifacts therefore limits the generalizability of our results and may mean that those interested in inferring teachers' TPACK from their portfolios may need to require the inclusion of specific artifacts or types of artifacts by teachers. The following research question guides our study: Can the TPACK-adapted TIAI be used to reliably measure the TPACK demonstrated in two types of artifacts designed to elicit teachers' TPACK?

Method

In this section, we describe how we established the reliability of our attempt to code specific artifacts. In doing so, we specifically discuss the sample, the procedure, and the data analysis of this study.

Sample

We analyzed two kinds of artifacts designed to elicit teachers' TPACK, both of which are featured in classes offered by the Master of Arts in Educational Technology program at Michigan State University. This program is designed for in-service teachers and other educators—the students who complete these assignments can therefore also be considered teachers.

- Dream IT: An assignment in a hybrid course (that is, a course with face-to-face and online aspects) on approaches to educational research in which students develop a grant proposal focused on the thoughtful integration of technology into teaching and learning. Students must consider TPACK when developing the proposal when describing how the technology will be used in practice. In addition, the assignment emphasizes the importance of tailoring the grant proposal for the context in which students function.
- Learning Technology Initiative: An assignment in an online course on technology, leadership, and education requiring students to solve a problem of practice, a difficult-to-solve issue that students face in their occupations, through a technology-integrated solution. Students must consider TPACK when completing a table that specifies what knowledge teachers would need when implementing this solution. This table is organized by the seven components of the TPACK framework (TK, PK, CK. TPK, TCK, PCK, and TPACK). The proposal also emphasizes the contextual factors that need to be considered in order for the initiative to be a success.

Students' Dream IT artifacts were all completed on individual websites; we therefore collected the URLs for the Dream IT artifacts completed during the Summer 2014 and Summer 2013 semesters (*n*=35) and added them to a spreadsheet. We then reviewed the Dream IT artifacts for suitability for coding by checking for broken links and completed assignments; this caused us to remove eight of them, reducing our sample to 27 artifacts. We also collected the Word documents for the Learning Technology Initiative artifacts completed during the Spring 2014

Preview version of this paper. Content and pagination may change prior to final publication.

semester (n=21) from the course submission system. All of the Learning Technology Initiatives were completed, allowing us to code all 21 artifacts.

To segment the data, we created individual spreadsheets for the Dream IT and Learning Technology Initiative artifacts and then added rows to each spreadsheet for each individual artifact. We then identified the portions of each of the two types of artifacts that were amenable to coding for evidence of teachers' TPACK. For the Dream IT artifacts, we only coded the students' descriptions of TPACK: These segments were typically one to two paragraphs in which students discussed the relationships among technology, pedagogy, and content with respect to their grant proposals. Thus, in the case of the Dream IT artifacts, we inferred teachers' TPK and TCK (as well as their TPACK) from their descriptions of how this framework informed their grant proposals. In the case of the Learning Technology Initiative artifacts, we coded only the table students created to present their descriptions of the knowledge teachers would need with respect to each of the seven areas of TPACK.

Procedure

Two coders evaluated each artifact for three aspects of the TPACK-adapted TIAI rubric from Harris et al. (2010), as shown in Table 2. As in our previous study, we did not analyze "fit"—a part of Harris et al.'s (2010) rubric—because we believed that the TPACK category included the same construct. Also, we added clarifications to the criteria to make explicit their connection to the different aspects of TPACK (e.g., Curriculum Goals & Technologies can be thought of as teachers' TCK). Note that Harris et al. (2010) referred to the content aspect of TPACK as "curricular goals" and the pedagogy aspect of TPACK as "instructional strategies." As an example, if an artifact exhibited strong alignment between the use of technology and curriculum goals it was coded "4" for Curriculum Goals & Technologies. If an artifact exhibited support from technology with respect to instructional strategies, it was coded "3" for Instructional Strategies and Technologies. Finally, if the alignment of technology, instructional strategies, and curriculum goals was exemplary, it was coded "4" for Technology Selections.

Criteria	4	3	2	1
Curriculum Goals & Technologies: Curriculum-based technology use (TCK)	Technologies selected for use are strongly aligned with one or more curriculum goals.	Technologies selected for use are aligned with one or more curriculum goals.	Technologies selected for use are partially aligned with one or more curriculum goals.	Technologies selected for use are not aligned with any curriculum goals.
Instructional Strategies & Technologies: Using technology in teaching/learning (TPK)	Technology use optimally supports instructional strategies.	Technology use supports instructional strategies.	Technology use minimally supports instructional strategies.	Technology use does not support instructional strategies.
Technology Selection(s): Compatibility with curriculum goals & instructional strategies (TPACK)	Technology selection(s) are exemplary, given curriculum goal(s) and instructional strategies.	Technology selection(s) are appropriate, but not exemplary, given curriculum goal(s) and instructional strategies.	Technology selection(s) are marginally appropriate, given curriculum goal(s) and instructional strategies.	Technology selection(s) are inappropriate, given curriculum goal(s) and instructional strategies.

Table 2: TIAI-adapted Rubric from Harris et al. (2010)

Data Analysis

We calculated interrater reliability for both kinds of artifact using two different statistics: percentage agreement and Cohen's kappa. Cohen's kappa is similar to percentage agreement but corrects for chance agreement between coders.

Results

Results for the reliability of our coding of both Dream IT *(n=27)* and Learning Technology Initiative (*n*=21) artifacts are presented in Table 3. Although we coded nearly 50 artifacts, our efforts at coding additional artifacts are ongoing, and so these reliability statistics may not be indicative of the statistics we will report after coding additional artifacts at our presentation.

Sample	Variable	Percentage Agreement	Cohen's Kappa and Interpretation
Dream IT (*n*=27)	TCK	62.9%	0.505 (moderate)
	TPK	51.8%	0.235 (slight)
	TPACK	55.5%	0.337 (slight)
Learning Technology Initiative (*n*=21)	TCK	42.9%	0.131 (poor)
	TPK	47.6%	0.138 (poor)
	TPACK	52.4%	0.163 (poor)

Table 3: Reliability of two coders for Dream IT and Learning Technology Initiative artifacts; we interpreted Cohen's kappa using Sim and Wright's (2005) guidelines as discussed in Table 1.

Discussion

Our reliability using the TPACK-adapted TIAI differed between the Dream IT and Learning Technology Initiative artifacts. For the Dream IT artifacts, we demonstrated some success in reliably coding these artifacts: for TCK, the two coders agreed a majority of the time across the four codes of the rubric, even when accounting for chance agreement. For TPK and TPACK the results were less impressive, especially when chance agreement was taken into account in the calculation of Cohen's kappa. Despite the reliability for TPK and TPACK being less impressive than that for TCK, the percentage agreement and Cohen's kappa for both of these variables was higher than those calculated in our initial attempt (e.g., Koehler et al., 2014). Thus, coding only Dream IT artifacts improved the reliability of our use of the measure. For the Learning Technology Initiative artifacts, while percentage agreement for TPK and TPACK increased from our initial attempt, the Cohen's kappa statistic increased only for TPACK. We demonstrated little improvement in our coding of these artifacts compared to both our initial effort and coding of the Dream IT artifacts. Thus, coding only Learning Technology artifacts did not improve the reliability of our use of the measure.

Differences between the two artifacts may account for the differences in the reliability of our coding. The Dream IT artifacts were written explanations of how students' grant proposals aligned with the TPACK framework, while the Learning Technology Initiatives were students' entries in a table with rows for each of the seven aspects of TPACK. Furthermore, the data we coded for the Dream IT artifacts was focused on just one area of TPACK (TPACK itself—that is, the intersection between technology, pedagogy, and content knowledge), as described in the sample section of the methods. This greater focus, on TPACK itself rather than its constituent areas of knowledge, may have allowed the coders to focus on straightforward responses rather than on extraneous details, thereby coming to achieve greater reliability using the rubric. Taking this into consideration along with Harris et al's (2010) suggestion that lesson plans must be written in sufficient detail to be coded, we may only be able to code artifacts with certain characteristics: being written in sufficient detail and also being focused on TPACK.

There are additional steps we can take to improve the reliability of our use of the TPACK-adapted TIAI rubric by continuing to focus on the nature of the artifacts we are coding. First, we can discuss the nature of our disagreements in coding the two types of artifacts. Doing so may reveal differences in how the coders used the rubric that can be reconciled, so that the reliability of the rubric can be improved in future rounds of coding. Second, we can explore different aspects of the two artifacts that we coded in this study: We segmented the Dream IT artifacts so that we coded only descriptions of TPACK and the Learning Technology Initiatives so that we coded only entries in the table about TPACK. Due to our comparative success coding the more-focused descriptions of TPACK among the Dream IT artifacts, making the data segments smaller may enhance the reliability of the rubric as long as they maintain a sufficient level of detail. Finally, we could explore different types of artifacts, from those not explicitly connected to TPACK, to those that make use of different forms of media, such as video, in order to further

Preview version of this paper. Content and pagination may change prior to final publication.

our understanding of the characteristics of artifacts that are amenable to coding. As our efforts at coding additional artifacts are ongoing, these steps provide direction for our efforts at enhancing the reliability of our use of the rubric.

References

Abbitt, J. T. (2011). Measuring technological pedagogical content knowledge in preservice teacher education: A review of current methods and instruments. *Journal of Research on Technology in Education*, *43*(4), 281-300.

Akcaoglu, M., Kereluik, K. & Casperson, G. (2011). Refining TPACK rubric through online lesson plans. In M. Koehler & P. Mishra (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2011* (pp. 4260-4264). Chesapeake, VA: AACE.

Archambault, L., & Crippen, K. (2009). Examining TPACK among K-12 online distance educators in the United States. *Contemporary Issues in Technology and Teacher Education*, *9*(1), 71-88.

Britten, J. S., & Cassady, J. C. (2005). The Technology Integration Assessment Instrument: Understanding planned use of technology by classroom teachers. *Computers in the Schools*, *22*(3-4), 49-61.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.

Graham, C. R., Tripp, T., & Wentworth, N. (2009). Assessing and improving technology integration skills for preservice teachers using the teacher work sample. *Journal of Educational Computing Research*, *41*(1), 39-62.

Harris, J., Grandgenett, N., & Hofer, M. (2010). Testing a TPACK-based technology integration assessment rubric. In C. Crawford, D. A. Willis, R. Carlsen, I. Gibson, K. McFerrin, J. Price & R. Weber (Eds.), *Proceedings of the Society for Information Technology & Teacher Education International Conference 2010* (pp. 3833–3840). Chesapeake, VA: AACE.

Kelly, M. A. (2008). Bridging digital and cultural divides: TPCK for equity of access to technology. In AACTE Committee on Innovation and Technology (Eds.), *Handbook of technological pedagogical content knowledge (TPCK) for educators* (pp. 30–60). New York, NY: Routledge.

Koehler, M. J., & Mishra, P. (2008). Introducing TPCK. In AACTE Committee on Technology and Innovation (Eds.), *Handbook of technological pedagogical content knowledge (TPCK) for educators* (pp. 3–29). New York, NY: Routledge.

Koehler, M.J., Rosenberg, J.M., Greenhalgh, S., Zellner, A., & Mishra, P. (2014, March). *Can portfolio-based assessments demonstrate teachers' TPACK?* In J. Voogt and P. Fisser (Chairs), Artifacts demonstrating teachers' technology integration competencies. Symposium conducted the Society for Information Technology & Teacher Education International Conference 2014, Jacksonville, FL.

Koehler, M. J., Shin, T. S., & Mishra, P. (2011). How do we measure TPACK? Let me count the ways. *Educational technology, teacher knowledge, and classroom impact: A research handbook on frameworks and approaches*, 16-31.

Mishra, P., & Koehler, M. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *The Teachers College Record*, *108*(6), 1017-1054.

Niess, M., Lee, K., Sadri, P., & Suharwoto, G. (2006). *Guiding inservice mathematics teachers in developing a technology pedagogical knowledge (TPCK).* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Porras-Hernández, L. H., & Salinas-Amescua, B. (2013). Strengthening TPACK: A broader notion of context and the use of teacher's narratives to reveal knowledge construction. *Journal of Educational Computing Research*, *48*(2), 223-244.

Rosenberg, J. & Koehler, M. (2014). Context and Technological Pedagogical Content Knowledge: A content analysis. In M. Searson & M. Ochoa (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2014* (pp. 2412-2417). Chesapeake, VA: AACE.

Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological Pedagogical Content Knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education*, *42*(2), 123-149.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 4-14.

So, H. J., & Kim, B. (2009). Learning about problem based learning: Student teachers integrating technology, pedagogy and content knowledge. *Australasian Journal of Educational Technology*, *25*(1), 101-116.

Preview version of this paper. Content and pagination may change prior to final publication.