

Journal of Educational Technology Development and Exchange (JETDE)

Volume 14 | Issue 1

2021

A Comparison of Assessment Book Authors' and Educators' Perspectives on Ethics Issues in Assessment: A Review Study

Jin Liu

Department of Educational Studies, University of South Carolina, liu99@mailbox.sc.edu

Robert L. Johnson

Department of Educational Studies, University of South Carolina, rjohnson@mailbox.sc.edu

Xumei Fan

Department of Educational Studies, University of South Carolina, fan9@mailbox.sc.edu

Ruiqin Gao

Department of Educational Studies, University of South Carolina, rgao@email.sc.edu

Follow this and additional works at: <https://aquila.usm.edu/jetde>



Part of the [Educational Technology Commons](#)

Recommended Citation

Liu, Jin; Johnson, Robert L.; Fan, Xumei; and Gao, Ruiqin (2021) "A Comparison of Assessment Book Authors' and Educators' Perspectives on Ethics Issues in Assessment: A Review Study," *Journal of Educational Technology Development and Exchange (JETDE)*: Vol. 14 : Iss. 1 , Article 1.

DOI: 10.18785/jetde.1401.01

Available at: <https://aquila.usm.edu/jetde/vol14/iss1/1>

This Article is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Journal of Educational Technology Development and Exchange (JETDE) by an authorized editor of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.

A Comparison of Assessment Book Authors' and Educators' Perspectives on Ethics Issues in Assessment: A Review Study

Cover Page Footnote

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

A Comparison of Assessment Book Authors' and Educators' Perspectives on Ethics Issues in Assessment: A Review Study

Jin Liu

liu99@mailbox.sc.edu

Robert L. Johnson

rjohnson@mailbox.sc.edu

Xumei Fan

fan9@mailbox.sc.edu

Ruiqn Gao

rgao@email.sc.edu

Department of Educational Studies

University of South Carolina

Abstract: *Assessment book authors' perspectives on ethical assessment practices are not necessarily consistent with those of educators. This study's purpose was to explore similarities and differences between the two perspectives. Researchers presented scenarios of classroom assessment practices to gain insights into educators' perspectives on ethical issues. Fourteen scenarios that were common across three empirical research articles were selected. Educators had similar opinions on a scenario if 70% or more respondents selected "ethical" or "unethical" on one item. Twenty-five assessment-related books were reviewed to present the authors' views on the ethicality of classroom assessment practices. The results showed that assessment book authors and educators held similar views on five of the 14 scenarios. Findings might inform the professional development of in-service teachers and the training of pre-service teachers. The results can inform assessment book authors in the future development to address ethics issues in assessment and practitioners in educational technology to consider ethical issues in the process of designing assessment tasks.*

Keywords: assessment; assessment books; comparison; educators; ethics;

Introduction

Ethics have long been the focus of philosophers, clergy, researchers, and educators. Socrates, one of the fathers of Western thought, believed that ethics is “What we ought to do” (Plato, 2009, p. 352). Definitions of ethical conduct are embedded within cultures and govern people’s everyday lives. Brandt and Rose (2004) stated that ethics emphasize the “principles of conduct” that people choose to guide their behaviors and actions. The language of ethics includes terms such as fair, just, trust, and right (Sockett, 1990; Strike et al., 2005).

In the field of education, professional ethics is defined as the “norms, values, and principles that should govern the conduct of educational professionals” (Husu, 2001, p. 68). In writing about the profession of teaching, Christenbury (2008) noted broader principles put forth by the philosopher Thomas Aquinas, who proposed that the virtues of an ethical life include faith, hope, charity (or love), prudence, temperance, courage, and justice.

Researchers have investigated ethics issues in teacher education and teaching practice (e.g., Ehrich et al., 2011; Finefter-Rosenbluh, 2016; Maxwell et al., 2016; Shapira-Lishchinsky, 2016) for decades. Given that the practice of teaching is filled with uncertainty, especially for novice teachers, professional training programs provide opportunities for teachers to gain practice in ethical decision-making (Shapira-Lishchinsky, 2016). Especially, care and justice should be considered

in forming integrated approaches in the professional training of ethics (Finefter-Rosenbluh, 2016).

As assessment is a critical component of teaching, many researchers seek to gain insight into ethical practices associated with it. In the field of assessment, ethics has been described as rules of behavior or practices that a profession imposes on itself (Sax, 1974). The earlier studies on ethics and teaching offer insights to researchers on ethical issues in assessment. For instance, the caring dimension of ethics should be considered in assessment (Beets, 2012; Johnson et al., 2017). The following scenario illustrates the role of professional ethics in assessment.

A parent brought a suit to the United States Supreme Court in response to the practice of peer grading (Starr, 2002). In the suit, Owasso Independent School Dist. No. I-011 v. Falvo, the assessment practice involved students exchanging papers for grading and then reading the grades aloud, so the teacher could record students’ grades. The mother of one student appealed to the Supreme Court that The Family Educational Rights and Privacy Act (FERPA) was developed to protect the privacy of student education records. The Supreme Court ruling was that the practice of peer grading does not violate FERPA.

This scenario provides an instance in which a teacher’s classroom assessment practices can propose issues related to professional ethics. Presenting descriptive scenarios based on actual classroom

assessment practices appears to be a useful method for gaining insights into educators' perspectives about ethics in assessment practices. Earlier studies explored educators' perspectives about the ethics of classroom assessment practices among pre-service teachers, in-service teachers, and educational leaders (Green et al., 2007; Johnson et al., 2008; Liu et al., 2016). All these studies used scenarios (e.g., A teacher considers students' growth in assigning grades.) to investigate educators' perspectives about ethical issues in assessment. The results indicated the need for further discussion of ethics in assessment.

To deal with ethics issues in assessment, educators can invite input from teachers, colleagues, peers, and/or friends. However, these individuals might not acknowledge responsibility for the views they hold. Their suggestions might be subjective and influenced by personal feelings on any particular occasion. Teachers can also consult with school leaders/administrators for guidance on school and district policies, but school leaders might need guidance at times as school/district policies are not always clear or well defined. In addition, this option is not available to pre-service teachers. One common resource available to all educators is assessment-related books. The purpose of this study was to investigate the similarities and differences in the perspectives of educators (i.e., pre-service teachers, in-service teachers, and educational leaders) and assessment book authors on ethics issues in classroom assessment.

Literature Review

In this section, we first briefly reviewed ethics research in teaching, and how it might offer insight to student assessment. Next, we pointed out multiple questionable assessment practices and potential consequences. Then, we discussed how assessment book authors and educators view ethics issues in assessment. We identified the potential differences between the two perspectives and justified the need for comparing the two perspectives in the current study. The research questions and hypotheses were stated in the end.

Research of Ethics in Teaching and Student Assessment

In the field of teacher education and teaching, there are many discussions of how to help educators deal with ethical issues in practice. Maxwell and colleagues (2016) described the significance of offering ethics content in training pre-service teachers using a survey from the United States, England, Canada, Australia, and the Netherlands. Researchers also developed ethical frameworks or models to help educators deal with relevant issues (Ehrich et al., 2011; Warnick & Silverman, 2011). Specifically, culturally relevant pedagogies have been used in the moral decision-making process to help teachers address inequities in high schools (Mungal, 2020). Finally, recent studies pointed out early-career teachers had a lack of ethical sensitivity in the profession (Maxwell et al., 2020).

Research of ethics in teaching and teacher education may offer insights into

ethics research in assessments due to the close relationships between student assessment and teaching. Researchers have addressed similar considerations in the field of assessments. For instance, ethical models (frameworks) helping educators deal with the ethical issues in student assessment were developed (Gao et al., 2019). Confronting the issues of equity in instruction and student assessment is an everyday part of a teacher's role (Johnson et al., 2008). The fairness of student assessments has become a major area of interest (Liu et al., 2016; Rasooli et al., 2019).

Questionable Assessment Practices and Consequences

Although efforts have been made to help teachers deal with ethics issues in assessments, many are not prepared to respond appropriately when faced with issues in assessments. Unethical assessment practices may influence students, teachers, parents, principals, and school boards. The following is an example of how questionable assessment practices may affect the stakeholders negatively.

The biology teacher decided to assign zeroes for semester projects to the students who plagiarized from the Internet (Carroll, 2002). The zeroes resulted in many students receiving failing grades for the science course. Even though the teacher had support from her principal and superintendent, parents complained to the school board, which, in turn, directed the teacher to assign partial credit to students and decrease the project's value from 50 percent to 30 percent of the final course grade.

Who was impacted in this scenario? The teacher lost respect from her students and resigned from her position. The students and school lost a biology teacher. The authority of the principal and superintendent was diminished. Parents and students learned the consequences of plagiarism are negotiable. In other words, few people were left untouched by the incident.

Teachers confront other questionable practices in classroom assessment. For instance, teachers' personal biases including generosity, severity, and central tendency errors (McMillan, 2010) may influence students' grades. All these practices involve score pollution and students may not obtain an accurate result of their mastery level of knowledge in learning.

Educators' reactions to the increasing accountability in high-stakes tests have contributed to unethical assessment practices. To increase students' test scores in high-stakes tests, the practice of "teaching to the test" occurs in the United States (e.g., Berliner, 2011). Examples of unethical practices in the assessment include assessing students with specific questions from a standardized test or a parallel form of the test (Mehrens & Kaminski, 1989). A recent survey confirmed that test score manipulation occurred in at least 37 states and Washington D.C. in the past four academic years (National Center for Fair & Open Testing, 2013). The researchers documented more than 50 ways of inflating scores from pre-testing to post-testing phases.

Questionable practices can result in tough consequences. From the perspective of student learning, students may not acquire the knowledge that is aligned with the learning standards in classrooms. In addition, parents, teachers, and schools may not be able to obtain accurate assessment results of students. The investigation by a state's law enforcement division, dismissal from the educators' professional society, and the loss of employment may be imposed on those who violate the rules of conduct.

Educators' Perspectives

Researchers (e.g., Johnson et al., 2008; Liu et al., 2016) have used web-based surveys that present various scenarios to analyze educators' perspectives on ethics issues in assessment. They consistently identified the divided opinions of educators (in-service teachers, pre-service teachers, and educational leaders) on making ethical judgments of assessment situations in the United States (Green et al., 2007; Johnson et al., 2008). Recently, this line of studies went beyond the context of the United States, indicating the worldwide concern of relevant issues. Liu and colleagues (2016) compared Chinese and U.S. pre-service teachers' perspectives about ethical issues in assessment and their findings indicated that the pre-service teachers from both countries had divided opinions on multiple scenarios. Such divided opinions were identified across cultures as well. Similar conclusions were obtained in another comparative study on the same topic. American pre-service teachers had different opinions on communications about grading and using

multiple assessment methods compared to Chinese pre-service teachers (Fan et al., 2019). In China, faculty members in higher education also reported divided opinions in certain classroom assessment issues with a low agreement level with assessment experts' views (Fan et al., 2020).

Resonating with Maxwell and colleagues (2020) who called for further investigation of teachers' perceptions and reaction to ethical situations in teaching, continuous conversations are necessary to help educators address assessment issues. As mentioned in the introduction, assessment book resources may offer useful guidance for educators on related issues.

Assessment Book Authors' Perspectives

Standards related to assessment have been developed to help guide educators' assessment practices. The Classroom Assessment Standards for PreK-12 Teachers (Joint Committee on Standards for Educational Evaluation, 2015) lists foundational standards for developing and implementing high-quality assessment procedures, standards to guide selection, development, and communication of classroom assessment results, and quality standards offering teachers guidance on sound classroom assessments. For instance, the five quality standards include cultural and linguistic diversity, unbiased and fair assessment, reliability and validity of the assessment, and reflection to improve the overall quality of assessment.

Many authors have written about

ethical practices of assessment in books. Green and Johnson (2010) presented two general guidelines on the subject: do no harm and avoid score pollution. The “do no harm” principle directs teachers to avoid acting in a manner that results in physical or mental harm to students. For example, when a teacher had students grade one another’s papers and call out grades for recording in the teacher’s grade book, students with lower grades might be hurt as their privacy and dignity were not well protected. “Avoid score pollution” requires that scores from any assessment should represent students’ actual mastery of the content; any other factors unrelated to mastery, such as neatness and late work, should not influence scores. Also, Beets (2012) argued that the principle of ethical caring, similar to “do no harm”, should be woven into student assessment. Teachers might reach different conclusions with different principles in particular assessment practices due to the controversial nature of ethics. For example, certain students don’t perform well on quizzes and tests despite their hard work and great effort, which could cause teachers to consider the effort in grading simply to keep students from being harmed by low grades. Regardless of its good intent, this grading practice causes score pollution. We believe “avoid score pollution” should function as a primary principle to ensure accuracy in assessment. Distinguishing low achievers from high achievers can help teachers make appropriate decisions about providing individualized assistance to students, which might seem to ignore the low achievers’ feelings and harm them

in the short term. However, students are more likely to benefit from the follow-up instructions and improve their learning in the long term.

With Socrates defining ethics as ‘how we ought to live’ (Plato, 2009, p. 352d), certain authors on classroom assessment did not use the term “ethics” and described what teachers ought to do, placing assessment practices under the umbrella of professional responsibilities and ethics. For instance, McMillan (2010) described the do’s and don’ts of effective grading, which can be used to judge whether relevant practices are ethical or unethical.

Comparison of the Views of Educators and Assessment Book Authors

Most assessment books intend to offer guidance to educators by discussing what they ought to do in practice. However, there are two major issues. First, current guidelines are not adequate to help educators resolve ethical issues in practice. Even if the assessment practices were discussed explicitly by the assessment book authors, educators might not strictly follow the guidelines suggested in the books. Researchers have noticed a disagreement between teachers’ grading practices and assessment experts’ recommendations. Stiggins et al. (1989) first noted that teachers did not follow the textbook authors’ recommendations in high-school grading practices. Later, multiple empirical studies received similar conclusions with pre-service teachers and in-service teachers (e.g., Brookhart, 1993; Campbell & Evans, 2000; Randall

& Engelhard, 2010). Similar conclusions were received based on our comparison of educators' perspectives about appropriate assessment practices and assessment book authors' recommendations. For example, while assessment books suggest that effort should not be considered in assessment (e.g., Brookhart & Nitko, 2008; Chappuis et al., 2012), educators have different opinions (Green et al., 2007; Johnson et al., 2008; Liu et al., 2016).

To our knowledge, no studies have been conducted to compare perspectives of assessment book authors and those of educators through a review study via diverse assessment practices. Due to the potential disagreement between assessment book authors and educators and the lack of discussions on certain assessment practices, a review study that integrates the arguments of stakeholders on the topic might initiate a framework to guide educators in terms of ethics in assessment practices. In the current study, we conducted a comprehensive review of book authors' opinions on ethical issues in classroom assessment practices as well as those of educators from different demographic groups.

The study focused on identifying differences between the two perspectives without judging the appropriateness. The study results may inform assessment book authors of the potential gap between assessment practices and what their assessment books recommended. The results of the study could also be helpful to the planning of professional developments for all educators, including in-service teachers as well as pre-service teachers.

Research Questions and Hypothesis

Our study was guided by four research questions: (1) What are educators' perspectives of appropriate (ethical) decisions? (2) What are assessment book authors' perspectives of appropriate (ethical) assessment decisions? (3) What are the similarities and differences between educators' and assessment book authors' perspectives? (4) Do assessment book authors comprehensively address assessment practices described in the scenarios? We expected to find both similarities and differences between the two perspectives and identify book resources that are lack discussions on certain assessment scenarios.

Methods

Twenty-five Assessment Books (Qualitative Data)

Assessment books addressing appropriate classroom assessment practices were the focus of this study. To identify a list of books for the analysis, we thoroughly searched books about classroom assessment from Google Scholar. We also expanded our search to books focusing on measurement and student evaluation given classroom assessment can be considered part of these broader topics. We then included two books on assessment standards and referred to another recently published review article by Fives et al. (2016) on classroom assessment to ensure we covered similar assessment book resources and/or assessment book writers. In the interest of more recent information,

we refined our focus on assessment books published after 2010. Finally, 25 assessment books that represent a broad picture of classroom assessment were selected as the qualitative data (see Table 1).

Table 1. Textbooks Reviewed for Guidance on Ethics and Classroom Assessment Practices

Resource Number	Classroom Assessment Texts Reviewed
1	American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational & Psychological Testing. (2014). <i>Standards for educational and psychological testing</i> . Washington, DC: American Educational Research Association.
2	Airasian, P. W. (2000). <i>Assessment in the classroom: A concise approach</i> . New York, NY: McGraw-Hill.
3	Banks, S. R. (2012). <i>Classroom assessment. Issues and practices</i> . Long Grove, IL: Waveland Press.
4	Brookhart, S. M., & Nitko. A. J. (2008). <i>Assessment and grading in classrooms</i> . Upper Saddle River, NJ: Pearson Education.
5	Brookhart, S. M., & Nitko. A. J. (2018). <i>Educational assessment of students (8th ed.)</i> . Upper Saddle, NJ: Pearson
6	Chappuis, J., & Stiggins, R. J. (2017). <i>An introduction to student involved assessment for learning (7th ed.)</i> . New York, NY: Pearson.
7	Chappuis, J., Stiggins, R. J., Chappuis, S., & Arter, J. A. (2012). <i>Classroom assessment for student learning: Doing it right—using it well</i> . Upper Saddle River, NJ: Pearson Education.
8	Frey, B. B. (2013). <i>Modern classroom assessment</i> . Los Angeles, CA: Sage Publications.
9	Green, S., & Johnson, R. (2010). <i>Assessment is essential</i> . New York: McGraw-Hill.
10	Gronlund, N. E., Linn, R. L., & Miller, M. D. (2013). <i>Measurement and assessment in teaching</i> . New York, NY: Pearson.
11	Hopkins, K. D. (1998). <i>Educational and psychological measurement and evaluation</i> . MA: Allyn & Bacon,
12	Joint Committee on Standards for Educational Evaluation (2015). <i>The classroom assessment standards for PreK-12 teachers</i> . Kindle Direct Press.
13	Kuhs, T., Johnson, R., Agruso, S., & Monrad, D. (2001). <i>Put to the test: Tools and techniques for classroom assessment</i> . Portsmouth, NH: Heinemann.
14	McMillan, J. H. (2010). <i>Classroom assessment: Principles and practices for effective instruction (5th ed.)</i> . Boston, MA: Pearson.

15	Oosterhof, A. (2009). <i>Developing and using classroom assessments (4th ed.)</i> . Upper Saddle River, NJ: Pearson.
16	Ory, J.C., & Ryan, K. E. (1993). <i>Tips for improving testing and grading</i> . Newbury, CA: Sage.
17	Popham, W. J. (2017). <i>Classroom assessment: What teachers need to know (8th ed.)</i> . Boston, MA: Pearson.
18	Russell, M. K., & Airasian, P. W. (2012). <i>Classroom assessment: Concepts and applications</i> . New York, NY: McGraw Hill.
19	Sax, G. (2010). <i>Principles of educational and psychological measurement and evaluation (4th ed.)</i> . Belmont, CA: Wadsworth
20	Taylor, C., & Nolen, S. (2008). <i>Classroom assessment: Supporting teaching and learning in real classrooms (2nd ed.)</i> . Upper Saddle River, N.J.: Pearson/Merrill/Prentice Hall.
21	Thorndike, R. M. & Thorndike-Christ, T. M. (2010). <i>Measurement and evaluation in psychology and education (8th ed.)</i> . New York, NY: Pearson.
22	Waugh, C.K. & Gronlund, N. E. (2013). <i>Assessment of student achievement (10th ed.)</i> . New York, NY: Pearson.
23	Williams, A. M. & Irvin, J. L. (1991). <i>Measurement and evaluation in education and psychology (4th ed.)</i> . Belmont, CA: Wadsworth Publishing.
24	Witte, R. H. (2012). <i>Classroom assessment for teachers</i> . New York, NY: McGraw Hill
25	Worthen, B. R., White, K. R., Fan, X., & Sudweeks, R. R. (1999). <i>Measurement and assessment in schools (2nd ed.)</i> . Needham Heights, MA: Allyn & Bacon.

Three Empirical Research Articles (Quantitative Data)

The authors used the following three published research articles to provide insight into the perspectives of educator respondents:

- Green, S., Johnson, R., Kim, D., & Pope, N. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23(7), 999–1011.

- Johnson, R., Green, S., Kim, D., & Pope, N. (2008). Educational leaders' perceptions about ethical assessment practices. *The American Journal of Evaluation*, 29(4), 520-530.

- Liu, J., Johnson, R., & Fan, X. (2016). A comparative study of Chinese and United States pre-service teachers' perceptions about ethical issues in classroom assessment. *Studies in Educational Evaluation*, 48, 57-66.

The three empirical research articles provided educators’ perspectives about the ethics of classroom assessment practices, thus allowing the comparison of educators’ perspectives with those in the assessment books. In these studies, researchers conducted a web-based survey composed of 36 assessment scenarios. Educators in K-12 were asked to review each scenario and indicate the ethicality of each item (ethical or unethical) based on their understanding of professional knowledge and specific situations. First, scenarios that were used in each study were reviewed and a table was generated with common scenarios used in all three studies. Fourteen scenarios were retained for subsequent analysis. Although six scenarios out of the 14 had slightly different wordings, they assessed had the same assessment situation based on our review. We used the wording of the 14 scenarios from the latest publication (i.e., Liu et al., 2016) to show items with

different wordings.

In these studies, respondents from different demographic groups were asked to review the scenarios and indicate the ethicality of the situation by selecting “ethical” or “unethical.” The percentages of educators rating a scenario as “ethical” were calculated based on the frequency of “ethical.” This information was used to show the perspectives of respondents from the following three groups: in-service teachers (Green et al, 2007), educational leaders (Johnson et al., 2008), and pre-service teachers (Green et al., 2007; Liu et al., 2016). The number of participants in each group and percentages of respondents who selected “ethical” for each scenario was added to Tables 2-5 as the main quantitative data resource. It is noted that the latest paper published in 2016 examined the perspectives of pre-service teachers in the U.S. and China and the current study only considered information about the U.S. population.

Table 2. Summary Table – Category 1: Scenarios Agreed on by Both Assessment Book Authors and Educators

Scenarios	Green et al. (2007)		Johnson et al. (2008)	Liu et al. (2016)	Assessment Book Authors’ Opinions, Resources, and Sample Quotes
	Pre-service Teachers (N=114)	In-service Teachers (N=55)	Educational Leaders (N=65)	Pre-service Teachers (N=173)	
1. Two teachers teach different sections of the same course. One teacher assigns very few report card grades of "A" because of her belief that students' work is rarely perfect.	18.4%	23.6%	27.7%	6.9%	Unethical Resources: #5, #9, #10, #12, #14, #17, #20, #22 “Objective assessment of performance can be threatened by personal bias and the halo effect.” (Resource #22-page 169).

2. An early elementary teacher uses observational checklists, and anecdotal notes, and interviews (student conferences) in assessing students.	99.1%	100%	100%	97.1%	Ethical Resources: #5, #8, #9, #10, #12, #13, #14, #18, #20, #22 “Parents and guardians of very young children depend on the information they receive from teachers to monitor their child’s progress in school, to know when and what to work on at home, and to know whether their child needs special assistance. Teachers can supplement collections with observational records or other kinds of assessments.” (Resource #20; page 392-393).
3. A teacher tells students which materials are important to learn in preparing for a class test.	97.4%	100%	95.4%	94.2%	Ethical Resources: #9, #10, #12, #14, #16, #20, #24 “Grading plans stated at the beginning of the course should not be changed without thoughtful consideration and a complete explanation to the students. Students should be informed about which course activities will be considered in their final grade. Information about the importance or weight of exams, quizzes, homework sets, papers, and projects should also be provided. Advise students of the relative importance of the topics covered in the course.” (Resource #16; page 114).
4. A teacher allows a student with a learning disability in language arts to use a tape-recorder when the student answers the essay questions on social studies tests.	93.9%	94.5%	96.9%	88.4%	Ethical Resources: #1, #4, #5, #9, #10, #12, #14, #17, #20, #21 #24 “Special attention to issues related to individuals of diverse linguistic backgrounds or with disabilities may be needed when developing, administering, scoring, interpreting, and making decisions based on test scores.” (Resource #21; page 267).
5. A teacher spends a class period to train students in test-taking skills (e.g., not spending too much time on one problem, eliminating impossible answers, guessing).	89.5%	90.9%	95.4%	94.2%	Ethical Resources: #1, #2, #3, #5, #9, #10, #14, #17, #18, #19, #25 “Teaching to the test involves teaching pupils the general skills, knowledge, and behaviors they need to answer the question on the test. This is an appropriate and desirable practical it is what good teaching testing are all about.” (Resource #2; page 105).

Note: Percentages indicate the number of educators considering an assessment practice as ethical. Item 2 was worded slightly different in three articles.

Table 3. Summary Table – Category 2: Scenarios Considered to be Unethical by Assessment Book Authors but with an Opposite Opinion by Educators

Scenarios	Green et al. (2007)		Johnson et al. (2008)	Liu et al. (2016)	Assessment Book Authors' Opinions, Resources, and Sample Quotes
	Pre-service Teachers (N=114)	In-service Teachers (N=55)	Educational Leaders (N=65)	Pre-service Teachers (N=173)	
6. A teacher considers students' growth in assigning grades.	79.30%	80.00%	76.90%	69.40%	Unethical Resources: #3, #9, #12, #14, #15, #16, #22, #24 (#1, #4, #11) "Using improvement as a basis for grading may be unfair. When grades are based on improvement, a student who initially knows a minimum amount of course material and learns quite a bit, receives a high grade. All students receiving an A in a course should have a similar grasp of the course material. When improvement and effort are used as a basis of grading, students with very different proficiency levels may all receive an A. Consequently, an A no longer means superior achievement or mastery of the course material." (Resource #16; page 126).
7. A teacher considers student effort when determining grades.	88.6%	78.2%	78.5%	74.0%	Unethical Resources: #3, #7, #9, #10, #12, #14, #15, #16, #17, #19, #20, #22, #24 (#1, #4, #11) "Introducing 'effort' points into the record of academic achievement will artificially inflate the grade when it comes time to calculate one, and it will no longer accurately represent level of achievement." (Resource #7; page 315).
8. In preparation for the district achievement testing, a teacher uses Scoring High, a commercially available publication with the same format and skills as the district achievement test. The booklet does not include the same questions as the district test.	86.6%	89.1%	93.8%	87.3%	Unethical Resources #5, #8, #9, #17, #19, #23 "Teachers should not try to improve student performance by developing items that are parallel those on standardized tests, nor should teachers administer one form of an examination when the district is to administer a second form as part of its testing program. Sometimes teachers do this to "preview" student performance on the regularly scheduled test, but the practice is unethical since it gives spuriously high scores to students having the advantage of practice." (Resource #19; page 43).

Note: Percentages indicate the number of educators considering an assessment practice as ethical. Items 7 and 8 were worded slightly different in different articles.

Table 4. Summary Table – Scenarios Considered to be Unethical by Assessment Book Authors but with Split Opinion by Educators

Scenarios	Green et al., (2007)		Johnson et al., (2008)	Liu et al., (2016)	Assessment Book Authors' Opinions, Resources, and Sample Quotes
	Pre-service Teachers (N=114)	In-service Teachers (N=55)	Educational Leaders (N=65)	Pre-service Teachers (N=173)	
9. A teacher always checks the name of the student whose essay test he is grading.	42.1%	61.8%	75.4%	30.6%	Unethical Resources: #5, #9, #10, #12, #16 “Score students fairly by removing from the scoring process anything that would cause unfair results. Examples include using objective items and a scoring key when appropriate, having students place their names on the back of their essay examinations, so you are not influenced by the name, scoring all student responses to one question before moving on to another, scoring performance tasks with a scoring rubric, periodically rescoring a sample of student responses as a check against your initial scoring, and having a colleague rescore a sample or all of your papers.” (Resource #5; page 94).
10. Based on her review of the district's mathematics frameworks, a teacher creates learning activities with specific math problems that are on the district's annual achievement test.	63.2%	32.7%	38.5%	75.7%	Unethical Resources: #5, #8, #9, #17, #18, #19, #21, #23, #25 “Teachers should not examine the content of standardized tests to determine what is to be taught in their classrooms.” (Resource #19; page 43).
11. A teacher weights homework heavily in determining report card grades.	57.9%	56.4%	50.8%	30.1%	Unethical Resources: #9, #11, #20 “Homework should be viewed primarily as an instructional activity, not an activity that has direct implications for evaluation.” (Resource #11; page 317).
12. To enhance self-esteem, an elementary teacher addresses only students' strengths when writing narrative report cards.	39.5%	43.6%	36.9%	25.4%	Unethical Resources: #5, #6, #9, #12, #16, #23 (#1) “Report both students' strengths and areas of need so that strengths can be built upon and areas of need addressed.” (Resource #12; page 537).

Note: Percentages indicate the number of educators considering an assessment practice as ethical. Items 9 and 10 were worded slightly different in different articles.

Table 5 Summary Table – Category 4: Scenarios with Split Opinion from Educators and Undecided Opinions from Assessment Book Authors

Scenarios	Green et al., (2007)		Johnson et al., (2008)	Liu et al., (2016)	Assessment Book Authors' Opinions, Resources, and Sample Quotes
	Pre-service Teachers (N=114)	In-service Teachers (N=55)	Educational Leaders (N=65)	Pre-service Teachers (N=173)	
13. To minimize guessing on a multiple-choice test, a high school teacher announces she will deduct more points for a wrong answer than for leaving the answer blank.	29.8%	32.7%	32.3%	19.7%	Undecided
14. While administering a mid-term achievement test, a high school teacher notices that a student has skipped a problem and is recording all of her answers out of sequence on the answer form. The teacher shows the student where to record the answer she is working on, and instructs the child to put the answer to each question with the same number on the answer sheet.	68.4%	70.9%	60.0%	86.7%	Undecided

Note: Percentages indicate the number of educators considering an assessment practice as ethical. Items 13 and 14 were worded slightly different in different articles.

A Mixed-methods Approach

We used a convergent parallel mixed method, in which researchers converge or merge quantitative and qualitative data to provide a comprehensive analysis of the research problem (Creswell & Creswell, 2018). Two parallel forms of data were collected, analyzed separately, and then merged. This method was chosen for the following reasons: researchers compared different perspectives drawn from the quantitative and qualitative data; qualitative and quantitative data were analyzed concurrently, and the same emphasis was placed on both types of data. We were able to compare and interpret the main findings from educators' views (quantitative data) and assessment

textbook authors' views (qualitative data). Next, we provided detailed information on how we analyzed both types of data.

First, we categorized the percentages of respondents from the three research studies for interpretation of results. We used 70% as the cut-off point for the percentages, which means a majority of the respondents (i.e., 70% or more) had similar opinions (ethical or unethical) on one scenario. Otherwise, we concluded that a scenario was associated with split opinions among respondents.

The next step was to review the 25 books to investigate how authors discussed ethical issues related to the fourteen scenarios. Thematic analysis

was conducted with the following six steps to build the trustworthiness of results (Nowell et al., 2017). First, we familiarized ourselves with the qualitative data (i.e., all book resources). The first author and the third author reviewed all selected book resources to get familiar with the book structure and major content. Then we engaged in deep discussions of the potential concepts the texts would address ethics issues in assessment. The second author, a senior researcher on this topic, was involved in the discussions to brainstorm ideas.

Next, the first and the third authors started to generate initial codes. The first and the third authors conducted a full-text analysis of the books to identify relevant discussions of content related to ethical issues in the assessment described in the 14 scenarios. We searched related keywords (e.g., effort, growth, or standardized test preparation) in each scenario to find relevant discussions in the selected assessment books. For those books without explicit discussions of ethics in assessments, we referred to descriptions of what teachers ought to do in assessment practices under the umbrella of professional ethics. We utilized a shared word file to record all relevant quotes including page numbers and references for each scenario. We reviewed each book resource twice to make sure all relevant content is identified. The first three authors met monthly to update the coding process.

After all book resources were thoroughly reviewed, the third step was taken to search for themes. For

each scenario, the first and the third authors reviewed all quotes and used a conventional context analysis (Hsieh & Shannon, 2005) to identify patterns of assessment book authors' opinions independently. In most instances, multiple books discussed the same issues in a scenario to achieve a stronger viewpoint on the ethics of certain assessment practices. Three or more book resources were needed to form a uniform opinion of the assessment book authors on a specific scenario. Many scenarios were discussed by authors directly, such as effort in assessments. Chappuis and colleagues (2012, p. 315) stated that "including effort points into the record of academic achievement will artificially inflate the grade when it comes time to calculate one, and it will no longer accurately represent the level of achievement." Thus, this scenario was unethical due to score pollution arising from taking effort into consideration of grading. For certain scenarios not directly discussed in book resources, we referred to relevant quotes to obtain authors' perspectives. For instance, "objective assessment of performance can be threatened by personal bias and the halo effect" (Waugh & Gronlund, 2013, p. 169), indicating an unethical assessment practice. Codes agreement levels were examined to show a similarity of coding between the two authors.

For any discrepancy identified, the second and the fourth authors were involved in the discussion to agree on conflicting quotes. After discussion, an agreement was reached on all quotes. Then, all authors came up with the

conclusion that experts shared similar opinions on most scenarios after several rounds of deep discussions. Then we started to develop themes by considering the quantitative results and the qualitative statements together.

The fourth phase was to refine the relevant themes (Braun & Clarke, 2006). The first and the third authors focused on scenarios that were without decided opinions from book authors. All book resources were reviewed one more time with an attempt to identify potential new book resources with the support of the fourth author. During this process, a few more quotes were identified. There were no major changes in the coding themes after this step.

Then, we finalized the themes in the fifth phase. Four themes were developed at this stage including scenarios agreed on by both assessment book authors and educators; scenarios considered to be unethical by assessment book authors but with an opposite opinion by educators; scenarios considered to be unethical by assessment book authors but with the split opinion by educators and scenarios with split opinion from educators and undecided opinions from assessment book authors. The final step was to write up the results. Table results were used to inform the readers of the major results of two perspectives with sample quotes and reference numbers from assessment book authors. Then all authors constructed the interpretation of the findings based on the descriptions with the thematic categories.

Results

As shown in Tables 2-5, in-service teachers, pre-service teachers, and educational leaders from the three studies had similar opinions on eight scenarios (Item1: unethical; Items 2-8: ethical) and split opinions on six scenarios (Items 9-14). Assessment book authors agreed with one another on most scenarios (i.e., Items 1, 6-12: unethical; Items 2-5: ethical), whereas two scenarios were classified as undecided in terms of the opinions of educators and assessment book authors (Items 13-14). Overall, educators had more split opinions than assessment book authors on ethics issues in assessment. Assessment book authors were more likely to notice unethical assessment practices than educators. Four categories were formed after comparing the opinions of educators and assessment book authors on selected scenarios. Detailed discussions were provided for each scenario under its corresponding category.

Category 1. Scenarios agreed on by both assessment book authors and educators (5 scenarios).

As shown in Table 2, educators and book authors had similar opinions on five of the fourteen scenarios. Educators and book authors were more likely to agree with each other when the scenarios described ethical behaviors (Items 2 to 5). These scenarios included basic classroom assessment practices, such as considering assessment formats for students with a learning disability, providing multiple assessment opportunities for students, informing students about the significance of materials for test preparation, and

training students in various test-taking skills. Overall, over 88% of educators in all groups considered the practices described in Items 2 to 5 ethical, whereas fewer than 28% of educators in all groups considered the practice described in Item 1 ethical. The only unethical scenario that was agreed on by educators in different groups and assessment book authors described a practice that involves personal bias in classroom assessment (Item 1).

Category 2. Scenarios considered to be unethical by assessment book authors but with an opposite opinion by educators (3 scenarios).

Table 3 indicated indicates that three scenarios evoked different opinions between book authors and educators with two scenarios (Items 6 and 7) related to grading practices and one scenario (Item 8) related to standardized test preparation. A high percentage of respondents in three groups believed it is ethical to consider students' growth or effort in determining grades. However, multiple authors of assessment books argued that it was unethical to consider growth or effort in grading due to score pollution. Item 8 also showed disagreement between educator respondents and assessment book authors. According to assessment book authors, it is inappropriate to provide practice on a published parallel test. High percentages of educator respondents in all groups considered it is ethical for teachers to use test preparation materials that had the same type of items and content as covered in a state or district achievement test in preparing students for the test.

Category 3. Scenarios considered to be unethical by assessment book authors but with split opinion by educators (4 scenarios).

We identified uniform opinions held by assessment book authors for the four scenarios in this category (Table 4). However, educators in different demographic groups expressed different opinions. Assessment book authors believed that students' names should be removed before scoring. However, we did not find similar perspectives from educator respondents. Responses for the ethicality of the practices of checking students' names in grading ranged from 30.6% to 75.4%. Although more than half of the pre-service teachers stated that this was not an ethical behavior, their opinion was counted as split using the 70% cut-off point. Next, standardized test preparation was often considered a questionable practice, as reflected in Item 10. Although the scenario depicted a practice related to teaching to the test, which was considered unethical by authors of 10 different assessment books, in-service teachers, educational leaders, and pre-service teachers had split opinions on this item, with the designation of ethicality ranging from 32.7% to 75.7%.

Homework is a common classroom assessment practice (Item 11), but it was a different story if teachers weighted homework heavily in scoring. Four authors of assessment books expressed the view that homework is to provide practice of new learning and therefore should not be weighted heavily in determining report card grades. The educators were split for the opinions on the practices

with responses ranging from 30.1% to 57.9% across studies. Item 12 depicted the practice of only addressing students' strengths in narrative report cards. This scenario was also considered unethical by assessment book authors. However, respondents in different demographic groups had a split opinion with the choice of the ethicality of scenario ranging from 25.4% to 43.6%.

Category 4. Scenarios with split opinion from educators and undecided opinions from assessment book authors. (2 scenarios).

Two scenarios showed undecided opinions from the assessment book authors (Table 5). There were not enough book resources to justify whether these assessment practices were ethical or unethical. As expected, there was no uniform opinion from educators either. Item 13 was related to grading practice in which a high school teacher announced she would deduct more points for a wrong answer than for leaving the answer blank. A low percentage (ranging between 19.7% and 32.7%) of respondents in the three published articles considered it ethical. Item 14 described an assessment practice in test administration. A comparatively high percentage of respondents (ranging from 60.0% to 86.7%) in the three studies considered it is ethical to direct a student to record his or her answer in the correct sequence.

Except for the two scenarios without clear opinions on assessment book authors, results in Tables 2 to 4 showed that only one book resource (Resource #9) addressed all 12 scenarios due to authors'

research interests in ethics issues. Four book resources (Resources #5, #10, #12, and #14) addressed half of the assessment scenarios, and three book resources (Resources #1, #17, and #20) addressed five or six assessment scenarios. These book resources were on assessment-related standards (Resource #1 or #12) or more likely to be a recent publication (e.g., Resources #5 and #17). The rest book resources (n=17) addressed four or fewer assessment scenarios.

Discussions

Studies related to ethics have started to include classroom assessment practices. We delineated the various perspectives of educators and assessment book authors related to this topic. Our review indicated that assessment book authors held similar viewpoints on most of the scenarios, whereas educator respondents as a group did not share high agreement levels on multiple assessment situations. Two groups had similar opinions on 5 out 14 scenarios. We also identified that most book resources did not provide a comprehensive view of assessment situations as multiple assessment scenarios were not addressed. Following are some lessons that could be beneficial to stakeholders in different groups.

Lessons Learned

Educators and assessment book authors tended to have similar opinions on commonly used assessment practices, which was reflected in the first five scenarios. These ethical decisions based

on scenarios were generally consistent with the Association of American Educators code of ethics for educators (Association of American Educators, 2010). For instance, the code clearly states that the educator endeavors to present facts without personal prejudice. In other words, teachers should not report few “As” because of personal beliefs (Item 1). Educators should make a constructive effort to help students learn. This code is consistent with providing multiple assessment opportunities to children (Item 2), communicating with students about important materials in test preparation (Item 3), and training students in test-taking skills (Item 5). Finally, the educators treat each student considerately and justly, which is highly related to offering accommodations to students with learning disabilities (Item 4). It is no surprise that almost all educators make ethical decisions about these scenarios as they are related to the basic ethics code in the teaching profession. The review results indicated that these scenarios have been widely discussed by assessment books with an average of 9.4 resources per scenario (see the number of resources in the last column of Table 2).

Meanwhile, differences between educators and assessment book authors were common as shown in Items 6 to 12. In the second category, assessment book authors and educators expressed opposite opinions. Educators favored considering growth and effort in grading. However, assessment book authors argued that students with very different proficiency levels may all receive an “A” if growth

and effort are considered (Items 6 and 7). Consequently, the clarity of the meaning of an “A” is lost; that is, it no longer means superior achievement or mastery of the course material (Ory & Ryan, 1993, p. 126). To ensure an informative assessment, educators should be aware that scoring should not be influenced by irrelevant factors including student effort, growth, behavior, and attendance (Brookhart & Nitko, 2008, p. 270; JCSEE, 2015, p. 505; Oosterhof, 2009, p. 218; Taylor & Nolen, 2008, p. 399).

In Category 3, in-service teachers and pre-service teachers had split opinions on all the scenarios, which indicated that the educators who presently teach, and those who will work with students in the future, have different perspectives of the ethical issues described in the scenarios. School leaders had split opinions on three of the four scenarios. They agreed with each other on one scenario (i.e., Item 9), but their opinions were opposite from those of the assessment book authors. Although school leaders do not work with students directly in the classroom, their perspectives might affect the classroom assessment practices of teachers. On the other hand, assessment book authors have offered clear guidelines on these items. Brookhart and Nitko (2018) stated that students’ names should be removed in the scoring process for fair grading results (Item 9). JSCEE (2015, p. 537) suggests that teachers report both students’ strengths and weaknesses so that strengths can be built upon and areas of need addressed (Item 12). Taylor and Nolen (2008) clearly described how

to weigh homework in grading. They stated that “Assessment experts generally recommend that homework and work done for practice or preparation be omitted from the grade since this type of work shows developing skills and knowledge rather than expertise” (p. 399).

The scenarios in Categories 2 and 3 were related to grading practices. In sum, training is necessary for in-service teachers and school leaders to improve professional competency in dealing with ethical issues in classroom assessment with the guidance of the assessment book authors. For pre-service teachers, instructors in higher education need to investigate how to develop classroom assessment courses to engage them in the conversations about possible ethical issues they may face in the future with the guidance of assessment textbooks.

Although standardized testing itself is not part of classroom assessment practices, test preparations play an important role in related practices (e.g., Items 5, 8, and 10) with each fitting into a different category. Mehrens and Kaminski (1989) listed different standardized test preparation practices from the most to least legitimate: “

1. general instruction on objectives not determined by looking at the objectives measured on standardized tests;
2. teaching test-taking skills;
3. instruction on objectives generated by a commercial organization where the objectives may have been determined by looking at objectives

measured by a variety of standardized tests (The objectives taught may, or may not, contain objectives on teaching test-taking skills);

4. instruction based on objectives (skills, subskills) that especially match those on the standardized test to be administered;

5. instruction on specially matched objectives (skills, subskills) where the practice (instruction) follows the same format as the test questions;

6. practice (instruction) on a published parallel form of the same test; and

7. practice (instruction) on the test. (p. 16)”

As the authors stated, most educators might agree that Scenario 1 or 2 is ethical, while Scenarios 6 and 7 are not ethical. However, Scenarios 3 – 5 are grey area situations. Educators and assessment book authors agreed on the necessity of practices for offering students training to improve students’ test-taking skills (Scenario 5 & Scenario 2). However, they had different opinions on other preparation practices, which indicates a need for professional training on standardized test preparations. The ultimate goal was to help students acquire the knowledge so that they are well-prepared for standardized tests.

Assessment book authors appear to provide little guidance on two assessment practices. Scenario 13 described a high school teacher who announces that she

will deduct more points for a wrong answer than for leaving the answer blank to minimize guessing on a multiple-choice test. Gronlund et al. (2013) offered the following suggestions for different types of tests:

“For liberally timed classroom tests, the ‘answer every item’ directions are favored. But for speed tests and when teachers want to discourage guessing, directions such as the following are a good compromise: ‘Answer all items for which you can find some reasonable basis for answering, even though you are not completely sure of the answer. Do not guess wildly, however, because there will be a correction for guessing.’” (p. 339).

Testing is used to measure students’ mastery of learning. The decision to deduct more points for a wrong answer than for leaving the answer blank might discourage students from attempting a test item and lead to difficult questions not being answered. The test will fail to examine students’ ability to use partial knowledge and analytic reasoning to find an answer. Furthermore, it would be difficult to analyze the functioning of some unanswered test items which could have been used for informing teaching and learning. We do not suggest that all test takers should guess the answers to the questions they do not know. Students’ guessing in a test is a complicated issue and might be impacted by student gender, cultural and educational background, and ability. Some advanced statistical techniques (e.g., item response theory) are available to detect “guessing” in multiple-choice questions instead of deducting

points directly, but this technique is too technical for teachers to use in their classrooms. Overall, correction for guessing is not recommended in classroom-based assessments.

Finally, educators mostly agreed on directing students’ answers in a correct sequence in test administration. Popham (1991) stated that in test administration, coaching students, or indicating in any way that their answers may be wrong should be considered as an inappropriate test administration practice. On the other hand, it can be argued that student errors in transferring answers to an answer sheet might result in inaccurate scores. In this instance, a student’s score does not reflect student learning, but the student’s ability to align answers with a response sheet. To resolve the contradiction, we suggest that teachers should train students in test-taking skills before the test so that this kind of error can be avoided during the test. This practice was described in Scenario 5.

Overall, various assessment book authors’ views were more uniform than those of educators. This might be due to the complexity of these issues for educators. Possibly, educators would consider contextualized information that could be related to each scenario. In contrast, assessment book authors are inherently more removed from practice and therefore more able to make black and white decisions in most situations. In other words, assessment book authors are less conflicted in making decisions. For instance, addressing only students’ strengths in narrative report cards

(i.e., Scenario 12) causes inaccurate analysis of student performance. From the perspective of providing formative feedback, weaknesses should be addressed so parents and students know how to improve on such areas. Assessment book authors opposed to reporting only strengths. However, teachers might care about students' feelings (e.g., self-esteem). This care may be expressed by reporting strengths only, especially for the low-performing students. School leaders, who work with a set of school expectations and district pressures, might have to follow certain rules in score reporting. Even pre-service teachers might have seen such examples of ethical issues in related courses and find it difficult to make decisions.

Limitations and Future Directions

The usage of 70% as the cut-off point to decide if respondents had similar opinions in their group was an arbitrary decision based on the researchers' judgment. In qualitative data analysis, researchers had to make judgments on authors' opinions based on relevant statements to make conclusions of the ethicality of assessment practices. Even when we used alternative wording in the review process, we might overlook the discussions of certain scenarios as authors might use slightly different wordings to discuss related issues. A more accurate method would be to email assessment book authors directly and ask their opinions of the selected scenarios in the future.

The three empirical articles reviewed only provided teachers' statuses to describe the samples without detailed sample characteristics, such as teaching experiences and cultural backgrounds. Therefore, we focused on how teachers of different statuses might have similar or different opinions with the assessment books in the current study. Future studies should be conducted to investigate how specific characteristics of teachers contribute to the similarities and differences between the two perspectives.

Until now, it is not feasible to investigate such issues with a more representative sample of educators in a state or across the country. However, respondents from different groups were included, which was considered to cross-validate the perspectives from different populations. We intend to investigate students' ethical perspectives on assessment issues in the future and compare their views with educators' perspectives. Admittedly, we only covered a handful of assessment situations with 14 scenarios. Future research should focus on asking teachers to describe specific assessment conditions that need to be resolved in practice. Then we can further evaluate if assessment textbook authors offer guidance on such issues. Among the selected assessment books, we did note certain assessment books were not recently published and searched for recent versions of assessment books for the most up-to-date information.

The current study described educators' opinions in a close-ended format (ethical or unethical). We were not clear of

why they had diverse opinions on most scenarios and different opinions with assessment book authors. Johnson et al., (2017) and Gao et al., (2019) developed and validated an ethical decision-making model that can be applied in considering ethical issues related to student assessment by breaking the ethical decision-making process into the following elements: conflict incidents, conflict elements, decision, justification, implication, and alternative suggestions. This model can be used to identify the underlying reasons for split opinions among teachers or educational leaders in professional development workshops and pre-service teachers in assessment courses. We are aware that teachers may not always have the freedom to act in the ways they deem to be ethical. Future research may investigate how other resources, such as opinions from colleagues, friends, and schools, influence educators' decision-making in open-ended discussions.

Finally, we would like to inform assessment book authors about our study results, which could help them learn the gaps between educators' perspectives and the content conveyed in their books. For instance, homework weight in grading and masking students' identity in essay grading were discussed in a few assessment books. They may consider adding related content in future editions. Due to the high-level involvement of educational technology in assessment activities, the results can offer insights to practitioners to consider relevant ethics issues. For example, they may consider designing anonymous options in the process of grading essays

for instructors to avoid grading bias.

We delved into an exploratory comparative study of assessment experts and educators regarding students' assessment practices. In addition to the consistently divided opinions identified from educators, we found that educators reported inconsistent views with assessment book authors on multiple assessment practices. The findings contribute to the theory and practice of educational assessment and professional ethics. It further suggests that assessment practices are complex, contextual, and dynamic, and the principles and guidelines in the assessment books might not sufficiently help teachers address all the issues in their assessment practices. We call for building a platform where teachers, administrators, book authors, students, parents, and other stakeholders could share views with the support of online technology, thus promoting ethical and fair assessment to ultimately support student learning.

Corresponding author:

Jin Liu
145 Wardlaw
Columbia, SC 29208
Email: liu99@mailbox.sc.edu

Acknowledgement:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Association of American Educators. (2010). Code of ethics for educators. Retrieved from <http://aaeteachers.org/index.php/about-us/aae-code-of-ethics>.
- Beets, P. (2012). Strengthening morality and ethics in educational assessment through Ubuntu in South Africa. *Educational Philosophy and Theory*, 44(2), 68-83. <https://doi.org/10.1111/j.1469-5812.2011.00796.x>
- Berliner, D. C. (2011). Rational responses to high-stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*, 41(3), 287-302. <https://doi.org/10.1080/0305764x.2011.607151>
- Brandt, D., & Rose, C. (2004). Global networking and universal ethics. *AI & Society*, 18(4), 334-343. <https://doi.org/10.1007/s00146-003-0289-3>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101. doi:10.1191/1478088706qp063oa
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), 123-142. <https://doi.org/10.1111/j.1745-3984.1993.tb01070.x>
- Brookhart, S. M., & Nitko, A. J. (2008). *Assessment and grading in classrooms*. Upper Saddle River, NJ: Pearson.
- Brookhart, S. M., & Nitko, A. J. (2018). *Educational assessment of students* (8th ed.). Upper Saddle, NJ: Pearson.
- Campbell, C., & Evans, J. A. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *The Journal of Educational Research*, 93(6), 350-355. <https://doi.org/10.1080/00220670009598729>
- Carroll, D. (2002, February 9). Piper teacher who resigned in plagiarism dispute is a hit with media. *The Kansas City Star*. Retrieved August 3, 2004 from <http://www.kansascity.com/mld/kansascity/news/local/2634836.htm?1c>.
- Chappuis, J., Stiggins, R., Chappuis, S., & Arter, J. (2012). *Classroom assessment for student learning: Doing it right-using it well* (2nd ed.). Boston: Pearson.
- Christenbury, L. (2008). A consideration of the ethics of teaching English. *English Journal*, 97(6), 32-38.
- Creswell, J. W., & Creswell, J.D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Thousand Oaks, CA: Sage.
- Ehrich, L. C., Kimber, M., Millwater, J., & Cranston, N. (2011). Ethical dilemmas: A model to understand teacher practice. *Teachers and Teaching: theory and practice*, 17(2), 173-185. <https://doi.org/10.1080/13540602.2011.539794>
- Fan, X., Johnson, R., Liu, J., Zhang, X., Liu, X., & Zhang, T. (2019). A comparative study of pre-service teachers' views on ethical issues in classroom assessment in China and the United States. *Frontiers of Education in China*, 14(2), 309-332. <https://doi.org/10.1007/s11516-019-0015-7>
- Fan, X., Johnson, R., Liu, X., & Gao, R. (2020). College students' views of ethical issues in classroom assessment in Chinese higher education. *Studies in Higher Education*, 1-15. <https://doi.org/10.1080/03075079.2020.1732908>
- Finefter-Rosenbluh, I. (2016). Behind the scenes of reflective practice in professional development: A glance into the ethical predicaments of secondary school teachers. *Teaching and Teacher Education*, 60, 1-11. <https://doi.org/10.1016/j.tate.2016.07.028>
- Fives, H., Barnes, N., Dacey, C., & Gillis, A.

-
- (2016). Assessing assessment texts: Where is planning? *The Teacher Educator*, 51(1), 70-89. <https://doi.org/10.1080/08878730.2015.1107442>
- Gao, R., Liu, J., Johnson, R., Wang, J., & Hu, L. (2019). Validating an ethical decision-making model of assessment using authentic scenarios. *Studies in Educational Evaluation*, 62, 187-196. <https://doi.org/10.1016/j.stueduc.2019.05.003>
- Green, S., & Johnson, R. (2010). *Assessment is essential*. New York: McGraw-Hill.
- Green, S., Johnson, R., Kim, D., & Pope, N. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23(7), 999-1011. <https://doi.org/10.1016/j.tate.2006.04.042>.
- Gronlund, N. E., Linn, R. L., & Miller, M. D. (2013). *Measurement and assessment in teaching*. New York, NY: Pearson.
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288. <https://doi.org/10.1177/1049732305276687>
- Husu, J. (2001). Teachers at cross-purposes: A case-report approach to the study of ethical dilemmas in teaching. *Journal of Curriculum and Supervision*, 17(1), 67-89.
- Johnson, R., Green, S., Kim, D., & Pope, N. (2008). Educational leaders' perceptions about ethical assessment practices. *The American Journal of Evaluation*, 29(4), 520-530. <https://doi.org/10.1177/1098214008322803>
- Johnson, R. L., Liu, J., & Burgess, Y. (2017). A model for making decisions about ethical dilemmas in student assessment. *Journal of Moral Education*, 46(2), 212-229. <https://doi.org/10.1080/03057240.2017.1313725>
- Joint Committee on Standards for Educational Evaluation [JCSEE] (2015). *The classroom assessment standards for PreK-12 teachers*. Kindle Direct Press.
- Liu, J., Johnson, R., & Fan, X. (2016). A comparative study of Chinese and United States pre-service teachers' perceptions about ethical issues in classroom assessment. *Studies in Educational Evaluation*, 48, 57-66. <https://doi.org/10.1016/j.stueduc.2016.01.002>
- Maxwell, B., Tremblay-Laprise, A., Filion, M., Boon, H., Daly, C., Hoven, M., Heilbronn, R., Lenselink, M., & Walters, S. (2016). A five-country survey on ethics education in preservice teaching program. *Journal of Teacher Education*, 67 (2), 135-151. <https://doi.org/10.1177/0022487115624490>
- Maxwell, B., Boon, H., Tanchuk, N., & Rauwerda, B. (2020). Adaptation and validation of a test of ethical sensitivity in teaching. *Journal of Moral Education*. <https://doi.org/10.1080/03057240.2020.1781070>
- McMillan, J. H. (2010). *Classroom assessment: Principles and practices for effective instruction* (5th ed.). Boston, MA: Pearson.
- Mehrens, W. A., & Kaminski, J. (1989). *Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? Educational Measurement: Issues and Practice*, 8(1), 14-22. <https://doi.org/10.1111/j.1745-3992.1989.tb00304.x>.
- Mungal, A. S. (2020). *Inequities and Ethical Dilemmas Beyond the Classroom: Joseph's Story*. *The Educational Forum*, 84(1), 29-47. <https://doi.org/10.1080/00131725.2019.1656316>
- National Center for Fair & Open Testing. (2013). Food safety shake-up needed in the USA. *The Lancet*, 375(9732), 2122. Retrieved from <https://www.fairtest.org/2013-Cheating-Report-PressRelease>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). *Thematic analysis*:

- Striving to meet the trustworthiness criteria. *International journal of qualitative methods*, 16(1), 1609406917733847.
- Oosterhof, A. (2009). *Developing and using classroom assessments* (4th ed.). Upper Saddle River, NJ: Pearson.
- Ory, J.C., & Ryan, K. E. (1993). *Tips for improving testing and grading*. Newbury, CA: Sage.
- Plato (2009) Republic, in: R. C. Solomon, C. W. Martin & W. Vaught (eds.), *Book I. Morality and the Good Life: An Introduction to ethics through the classical sources*, 5th ed., G. M.A. Grube, trans. Boston, MA: McGraw-Hill.
- Popham, J. W. (1991). Appropriateness of teachers' test-preparation practices. *Educational Measurement: Issues and Practice*, 1(1), 12-15. <https://doi.org/10.1111/j.1745-3992.1991.tb00211.x>
- Popham, W. J. (2017). *Classroom assessment: What teachers need to know* (8th ed.). Boston, MA: Pearson.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372-1380. <https://doi.org/10.1016/j.tate.2010.03.008>
- Rasooli, A., Zandi, H., & DeLuca, C. (2019). Conceptualising fairness in classroom assessment: Exploring the value of organisational justice theory. *Assessment in Education: Principles, Policy & Practice*, 26, 584-611.
- Sax, G. (1974). *Principles of educational measurement and evaluation*. Belmont, CA: Wadsworth.
- Shapira-Lishchinsky, O. (2016). From ethical reasoning to teacher education for social justice. *Teaching and Teacher Education*, 60, 245-255. <https://doi.org/10.1016/j.tate.2016.08.010>
- Sockett, H. (1990). Accountability, trust, and ethical codes of practice. In J. Goodlad; R. Soder, & K. Sirotnik (Eds.), *The moral dimensions of teaching* (pp. 224-250). San Francisco, CA: Jossey-Bass.
- Starr, L. (2002). Peer grading vs. privacy: The Supreme Court rules. Retrieved from http://www.educationworld.com/a_issues/issues279.shtml
- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, 8(2), 5-14. <https://doi.org/10.1111/j.1745-3992.1989.tb00315.x>
- Strike, K., Haller, E., & Soltis, J. (2005). *The ethics of school administration* (3rd ed.). NY: Teachers College Press.
- Taylor, C., & Nolen, S. (2008). *Classroom assessment: Supporting teaching and learning in real classrooms* (2nd ed.). Upper Saddle River, N.J.: Pearson/Merrill/Prentice Hall.
- Warnick, B. R., & Silverman, S. K. (2011). A framework for professional ethics courses in teacher education. *Journal of Teacher Education*, 62(3), 273-285. <https://doi.org/10.1177/0022487110398002>
- Waugh, C.K. & Gronlund, N. E. (2013). *Assessment of student achievement* (10th ed.). New York, NY: Pearson.
- Williams, A. M. & Irvin, J. L. (1991). *Measurement and evaluation in education and psychology* (4th ed.). Belmont, CA: Wadsworth Publishing.