
Title	The impact of conceptions of assessment on assessment literacy in a teacher education program
Author(s)	Christopher Charles Deneen and Gavin T. L. Brown
Source	<i>Cogent Education</i> , 3(1): 1225380
Published by	Taylor & Francis (Routledge)

Copyright © 2016 The Authors

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0>).

This is an article published by Taylor & Francis in *Pastoral Care in Education* on 13/09/2016, available online: <http://www.tandfonline.com/doi/pdf/10.1080/2331186X.2016.1225380>

Citation: Deneen, C. C., & Brown, G. T. L. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, 3(1): 1225380. <http://dx.doi.org/10.1080/2331186X.2016.1225380>



Received: 17 May 2016
Accepted: 09 July 2016
Published: 13 September 2016

*Corresponding author: Christopher Charles Deneen, Curriculum, Teaching and Learning, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616, Singapore
E-mails: Christopher.d@nie.edu.sg, cdeneen212@gmail.com

Reviewing editor:
Darlene Ciuffetelli Parker, Brock University, Canada

Additional information is available at the end of the article

EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

The impact of conceptions of assessment on assessment literacy in a teacher education program

Christopher Charles Deneen^{1*} and Gavin T.L. Brown²

Abstract: Assessment literacy is considered essential to modern teaching. Over time, assessment literacy has evolved to include both measurement and assessment for learning perspectives. At the same time, research into teachers' conceptions of the purpose and role of assessment demonstrates increasing evidence of the impact of teachers' conceptions on assessment practices. The conjunction of these two factors, assessment literacy and conceptions of assessment has not been adequately explored. This study addresses this need by examining the impact of a master's level teacher education course in educational assessment on student teachers' expressed literacy in and conceptions of assessment. Achievement data were collected and interviews conducted in a class of 32 pre-service and practicing teachers. Inferential analysis and qualitative coding were applied to the data. Analytical results included a strong, polarized affective component. These positive and negative affective conceptions appeared independent of level of academic achievement. Academic achievement appeared to play a role in allowing deeper articulation of conceptions, but did not accompany particular conceptual changes. These findings suggest that while fluency in factual knowledge (i.e. assessment literacy) was enhanced; conceptions of assessment that may influence application

ABOUT THE AUTHORS

Dr Christopher Charles Deneen is an assistant professor at The National Institute of Education, Nanyang Technical University in Singapore. Chris is particularly interested in how innovation and leadership in assessment may build relationships between assessment for and of learning. He focuses on the influence of change management, technology and stakeholder conceptions on achieving formative purposes, meaningful summative achievement and sustainable competencies.

Gavin T.L. Brown is a professor at University of Auckland in New Zealand. Gavin is especially interested in teachers' conceptions of assessment and their impact on performativity. Deneen and Brown are currently working on a 246,000 SGD-funded research project that employs mixed methods to examine policies, perceptions, and practices of AfL in Singapore secondary schools. The grant builds on and extends the work presented in this paper.

PUBLIC INTEREST STATEMENT

Understanding and use of assessment and data that comes from assessment have evolved into key elements of modern teaching and teacher education. We collectively refer to these abilities as assessment literacy. Teachers' conceptions of assessment are equally important as they can significantly impact teachers' assessment practice. Connections between assessment literacy and conceptions of assessment are not yet well researched, however. This paper seeks to address this by providing findings from a relevant research study.

Researchers examined the impact of a masters-level teacher education course in educational assessment on participants' literacy in and conceptions of assessment. Data was collected on 32 pre-service and practicing teachers enrolled in the course.

Findings suggest that while assessment literacy was enhanced; conceptions of assessment that may influence practical application of assessment literacy remained unchanged. Some variations between pre-service and practicing teachers are examined. Implications for teacher education and the practice of assessment are explored.

of assessment literacy were not changed through the mediating influence of the assessment course. Implications of the results and this apparent disconnection between assessment literacy, teacher education and practice are explored.

Subjects: Assessment; Assessment & Testing; Education; Educational Research; Teachers & Teacher Education

Keywords: assessment; conceptions of assessment; assessment literacy; teacher education; pre-service teachers; teacher education curriculum

1. Introduction

Teachers must be assessment literate; this priority is so important that it is “professional suicide” for teachers to ignore it (Popham, 2004). Teacher education programs remain a principal route to teacher preparation in the United States for both pre-service and practicing teachers. Assessment literacy is, therefore deemed a priority in modern teacher preparation and corresponding curricula. Historically, how well this priority translates to successful outcomes for graduates of these programs is debatable.

Teacher education programs’ attempts to define and achieve assessment literacy is a subject of increasing attention. Approaches to assessment literacy reflect several historical stages: a statistical interpretive stage, a stage emphasizing assessment for learning, and a modern interpretation of assessment literacy which draws from and synthesizes elements of prior stages. Accepted goals of current pre-service or advanced-level courses therefore include improved understanding of innovative assessment practices (e.g. portfolios, rubrics, performances, peer and self-assessment), attention to assessment as, for and of learning (AaL; AfL; AoL), and skills development in various assessment processes (i.e. design, administration, scoring, valid statistical interpretation, and action). These represent an ongoing evolution in identifying, defining, and operationalizing the most useful facets of assessment for teachers. These particular emphases may also define criteria and benchmarks for achieving the outcomes of assessment literacy through teacher education (Popham, 2010). The principal route to achieving these outcomes is having students engage with specific courses on assessment, or units within courses on assessment.

A growing body of research suggests that in the practice of assessment, how teachers conceptualize assessment is at least as important as the above characteristics. Teachers’ conceptions about the purposes of assessment influence implementation of assessment practices at all educational levels (Barnes, Fives, & Dacey, 2015; Brookhart, 2011; Deneen & Boud, 2014; Fulmer, Lee, & Tan, 2015). Positive conceptions of assessment (e.g. assessment should enhance students’ learning) have been shown to precipitate beneficial assessment practices; negative conceptions of assessment (e.g. assessment is bad for students or irrelevant to learning) may play a significant role in teachers resisting or subverting assessment policies and intended practices (Brown, 2008; Deneen & Boud, 2014). The utility of teacher education programs is built on the presumption of enhancing practice; thus, it is essential to understand the relationship of teachers’ conceptions of assessment to assessment literacy and a teacher education program’s approach to enhancing that literacy.

This paper reports on a mixed-methods study that examines how and to what degree a graduate-level course designed to enhance assessment literacy among practicing and pre-service teachers interacts with and mediates conceptions of assessment. This was accomplished with a small but diverse sample of New York City pre-service and practicing teachers.

2. Review of literature

2.1. *The evolution of assessment literacy*

Assessment in teacher education programs is taught as an amalgamation of skills and knowledge introduced to bring about fluency; that is, demonstrated understanding and application of assessment practices within authentic educational contexts. This is commonly abbreviated as assessment literacy (Popham, 2004, 2008). Becoming assessment literate requires developing understanding in theory and application of diverse assessment practices and skills to appropriately administer and interpret assessments at classroom and jurisdictional levels (Stiggins & Chappuis, 2005). This is not, however, a values-free definition. As teacher education has changed over the years, assessment literacy has transitioned through what might be loosely grouped as three stages. Each stage reflects certain beliefs about what matters most in assessment.

A quarter century ago, teacher education provided little formal training in assessment practice (Stiggins, 1991). In response, standards were designed to help guide teacher education programs and professional development in assessment (AFT, NCME, & NEA, 1990). Early efforts to foster effective assessment practice focused on measurement theory and skills. Subsequent research demonstrated that this approach to assessment literacy had limited formative utility, especially in developing teachers' capacity to provide effective feedback and communicate assessment results toward positive change (Brookhart, 2011; Leighton, Gokiert, Cor, & Heffernan, 2010; Plake & Impara, 1997).

The next stage of assessment literacy saw the ascension of AfL as a priority for teacher education and teacher practice (Stiggins, 2004). Increasing attention was given to the longitudinal impact of assessment on students; Ecclestone and Pryor (2003) suggest that these long-term interactions may be understood as a students' assessment careers. The balance of formative and summative intentions within these careers are posited as having significant consequences for students' success. Accompanying this belief was a widely disseminated message that summative measurement, especially through standardized testing was bad for students and their learning. The most vehement proponents of this perspective suggested that this was not an issue of illiteracy or malpractice, but a more fundamental problem with conceptualizing assessment as measurement (Gibbs & Simpson, 2004; Shepard, 2000). Teacher education emanating from this second stage of assessment literacy therefore pushed back against a central tenet of the preceding stage. Measurement skills should be deemphasized or even dispensed with as benchmarks of assessment literacy and replaced with those oriented towards assessment for learning. Given the rise of test-based accountability during this same time period, this tenet of stage two assessment literacy created a schism between what many teacher education programs promoted as desirable and what teachers encountered through assessment practice.

An emphasis on AfL was a positive step; wholesale rejection of assessment as measurement, however is neither feasible nor professionally authentic. The third or current stage of assessment literacy therefore represents an attempt to mend this schism, synthesize the best priorities of earlier stages and prepare teachers to negotiate increasingly complex tensions and demands. The procedures and results of recent large-scale accountability testing programs (e.g. Race to the Top, No Child Left Behind) carry career-defining consequences for teachers and school leaders (Guskey, 2007; Ravitch, 2011). Teachers and administrators are expected to make curriculum, planning, and pedagogical decisions based on standardized test data disaggregated at school and class level (Hallinger, Murphy, & Hausman, 2013; Popham, 2008). Utilizing assessment data at various levels is, therefore a key professional skill (Fulmer, 2011). This necessitates particular interpretations of assessment literacy (Brookhart, 2011). Teachers and administrators must navigate federal accountability and state-level testing systems to prepare students, report to stakeholders, and preserve their job; teacher education must, therefore prepare them for the literate performance of these tasks. Third stage assessment literacy must meet these demands.

AfL remains a profound imperative, though. AfL requires that teachers and students perform as active assessors, using interpretations of achievement to build learning and learner capacities (Earl,

2012; Stiggins & Chappuis, 2012). The juxtaposition of high stakes accountability testing against Afl creates tension between “the pressure to have one’s students do well on a test so the district will look good in media coverage, and students’ learning needs” (Brookhart, 2011, p. 5). Thus, the modern, third stage of assessment literacy is more about balancing tensions and navigating competing purposes than expressing one particular theoretical perspective (Brown & Harris, 2016).

2.2. *Teacher conceptions of assessment*

For the purposes of this paper, the term conception is inclusive of attitudes, perceptions, dispositions and other terms that suggest belief about a phenomenon. Considerable research has reported on how teachers think about teaching, learning, knowledge, and curriculum (see Bennett, 2011; Fives & Buehl, 2012). Research has begun to focus on teachers’ thinking about the nature and purpose of assessment. Strong emerging evidence suggests that assessment outcomes and practices may be influenced by teachers’ conceptions of assessment (Barnes et al., 2015; Fulmer et al., 2015). As assessment literacy continues to evolve, conceptions of assessment should, therefore be taken into account. Brown (2008) has proposed that teachers’ varied conceptions of assessment may be aggregated into four major purposes (i.e. informing educational improvement, evaluating students, evaluating schools and teachers, and irrelevance). Improvement focuses on assessment informing teachers and students regarding what students need to learn next and how this may guide learning. Evaluation has to do with appraising student performance against standards, assigning scores/grades and awarding qualifications. School evaluation has to do with using assessment results to determine the performance of teachers and schools. Irrelevance has to do with either rejecting assessment as having meaningful connection to learning or believing it to be bad for students.

Research on conceptions of assessment suggests that there is a role to be played in assessment literacy (Benson, 2013; Brown, 2004; Ludwig, 2013). While teachers’ conceptions of assessment is a growing field of research globally (Barnes et al., 2015; Fulmer et al., 2015), research on how teachers conceptualize assessment within the context of the United States requires further attention. Research into teachers’ interactions with district, state, and federal accountability testing systems has focused more on functional interactions or systemic conceptions, rather than looking directly at teachers’ conceptions of assessment. Shepard (2006) highlighted the potency of beliefs and the importance of taking them into account in understanding how teachers enact classroom assessment. A teacher who approaches the act of assessment with the belief that the purpose of assessment is punitive (e.g. administrators punishing teachers for low scores) or disempowering to student or teacher (e.g. formative assessment overridden by the requirements of mandated standardized testing) may possess the requisite understanding and skills to be considered assessment literate; their practice of assessment, however will likely reflect these negative conceptions as much as their skills and knowledge. If teacher preparation is meant to lead to specific and salutary practices, then teacher preparation must address conceptions of assessment.

Addressing conceptions of assessment has significant implications for teacher education. Changes in teachers’ beliefs or attitudes usually take longer than changes in knowledge or skill (Guskey, 2007). Thus, planning and implementing conceptual change would require a generous budget of time and attention. Degree and certification programs in teacher education, however may last as little as 1–2 years. Skills and knowledge are, for the most part developed within courses. Capstone projects and practicums require students’ synthesis of skills and knowledge across a program and occur at the program end; their nature and timing discourage a singular focus on conceptions of assessment. Although they may pose problems in duration, current implementation of teacher education curricula suggests that courses in assessment, more than any other aspect of teacher education, must address conceptions and beliefs (Brookhart, 2011; Shepard, 2006). This appears to pose a fundamental problem: do teacher education curricula offer sufficient time and space for conceptual change to take place? Examining conceptual change in assessment within the common model of teacher education would allow some understanding of the utility of this model in achieving this change; this in turn would inform understanding of how this model may promote or inhibit developing further stages of assessment literacy that are inclusive of teacher conceptions.

2.3. Assessment in NYC

New York City has developed one of the most aggressive accountability testing regimes in the United States. The federal No Child Left Behind Act of 2001 (NCLB) demanded local testing and accountability systems be established and maintained (McGuinn, 2006). An essential function established under NCLB and continued under the auspices of its successor, Race to the Top is the report of Adequate Yearly Progress (AYP) that holds schools and their administrators accountable for longitudinally tracked variations in testing scores (US Department of Education, 2011). AYPs have tremendously increased the role of assessment and more specifically, grade-level testing as instruments of school-based accountability. Poor student test score results have been used directly and indirectly to dismiss numerous school administrators (Shulman, Sullivan, & Glanz, 2008). Although NYC received an easement of requirements, high stakes remain the dominant factor in determining school quality (Schwartz, 2012). Assessment is, therefore a contested space, with teachers and administrators attempting to balance AfL with high stakes demands. Given national and international trends, NYC provides a focused, concentrated version of the educational tensions that many teachers and administrators experience. Thus, a research study set in this context may have meaningful implications for stakeholders operating in many US and global jurisdictions.

3. Context of study

Participants were enrolled in a 12-week graduate-level course on assessment. The course addressed a diverse range of topics ranging from understanding and interpreting large-scale test results to effective use of AfL techniques in the classroom. Thus, the course may be understood as emanating from a stage-three assessment literacy perspective. The course is taught as part of a Masters in Education (MEd) program at a large urban university in New York City. The program has approximately 2,500 enrolled students forming a diverse racial and socioeconomic profile. There were 32 participants in the assessment course under study.

The course aim, as identified in the course outline was to enhance assessment literacy; course outcomes were structured toward students demonstrating achievement relevant to this aim. The course was designed using constructive alignment (Biggs, 1996) so that teaching, assessment and other learning engagements aligned with the intended aim and supporting outcomes. A key focus of the course was negotiating the tension between accountability and improvement. Course materials were selected using this criterion. The central text by Popham (2013), for example explicitly addresses this tension. Supplementary readings and course activities focused on enhancing both AfL and assessment-oriented statistical skills.

The course instructor has significant experience teaching in the field of assessment and has published research on relevant assessment issues (e.g. Deneen, Brown, Bond, & Shroff, 2013; Deneen & Deneen, 2008). Both the instructor and the course have received positive student, peer, and program evaluations in earlier runs of the courses. The course design was subjected to internal peer review and external evaluation via NY Regents and Middle States program accreditation prior to implementation and periodic internal and external reevaluation, thereafter.

Teaching and learning were designed to facilitate student outcome achievement; assessment was designed to provide opportunities to demonstrate evidence of this achievement. Assessment within the course was designed for formative and summative purposes; there were three assessments spaced evenly throughout the course, with the first assessment occurring after the first three sessions. Extensive instructor feedback was provided and each assessment was returned within two weeks of assignment submission. Preemptive and dialogic feedback (Carless, 2007) were also integrated into the formal course design, which allowed for guidance on assignments prior to final submission and scoring. Summative assessments were marked according to rubrics that had been subject to external validation and shared well in advance with students. Cross marking and outside moderation were conducted on the final assignment.

The assessments served as three data points for collection of student work. Course assessments consisted of:

- (1) An analytical paper on an assessment-related article
- (2) An assessment tool constructed by the student with an accompanying self-critique and description of use
- (3) A final examination consisting of a take-home problem-based exercise in which students were asked to apply course knowledge to a hypothetical assessment scenario

4. Methodology

This study was guided by three research questions:

- (1) What are the conceptions of assessment held by a sample of pre-service and practicing teachers in a graduate teacher education program?
- (2) In what ways did these conceptions change through a course on educational assessment?
- (3) What is the relationship between participants' conceptions of assessment and their assessment literacy, as expressed through course assessment results?

4.1. Participants

Thirty-two practicing and pre-service teachers were enrolled in the course. Based on course completion, student achievement data from 30 participants were used. Pre- and post-course interviews were conducted with six of the 30 participants. This constitutes a sub-sample of opportunity, in that the 30 students whose data were used were offered interview participation; six volunteered. Interview participants represented a range of course achievement. The interview sample, while small, represents a diverse population of practicing and pre-service teachers. Three participants were, at the time of data collection pre-service teachers and three were practicing teachers. Four participants are women, two are men, three are naturalized Americans, two of whom were from Eastern Europe, and one from Russia. One participant self-identified as African-American. All participants formally consented to have their data used as part of the study; all identified participants in the study were assigned pseudonyms. Research procedures were approved via an institutional review process and conducted according to recognized ethical standards.

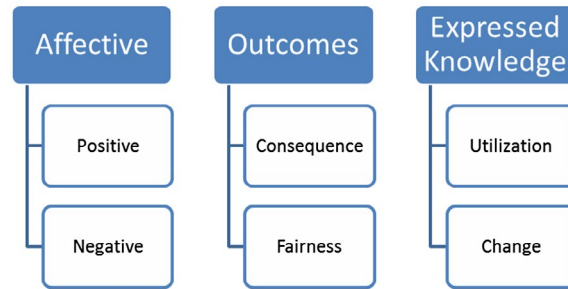
4.2. Data collection

Course assessment results were used from all 30 participants. Researchers conducted twelve 45-min interviews, with each of the six interview participants interviewed twice. The first interview occurred during week two of the assessment course, the second interview was in the week after the final meeting of the class, 13 weeks from the beginning. Interviews were semi-structured and drew on ethnographic techniques (Spradley, 1979). Interview protocols and questions were structured around the initial four conceptions of assessment identified by Brown (2004); specific items from the *Teachers' Conceptions of Assessment* abridged inventory (Brown, 2006) were used to initiate dialog. As per ethnographic interviewing procedure, prompts and opportunities were given for participants to express their own understandings and emphases.

4.3. Analysis

The course instructor and a researcher from outside the institution conducted analysis (first and second author of this paper). Analysis of assessment results was conducted using basic inferential statistics. Analysis of transcribed interview data was strongly influenced by the sequential and iterative coding process of phenomenography. Transcript data were initially organized according to relevance and frequency by the first author. From the results of this, a data pool (Marton, 1986) of material was established. The second author reviewed results of the initial coding pass; adjustments were then discussed and finalized. From this adjusted data pool, conceptual categories were assembled according to variation of concepts, structural relationships among concepts and the surrounding transcribed context (Åkerlind, 2005). This stage was initially performed by the first author;

Figure 1. Framework of interview analysis results.



procedures and results were reviewed and adjusted through dialog with the second author. Finally, criterion attributes of each conceptual category were established collaboratively. This allowed for the construction of a final hierarchical “outcome space” (Åkerlind, 2005) from which this study draws its findings (Figure 1). While the process was sequential, it was also highly iterative; the researchers revisited data, categories and codes multiple times, both independently and together. Analysis was facilitated through the use of ATLAS Ti software.

5. Limitations

Small-scale intensive research carries some inherent limitations. These include generalizability and replicability of results. This study does not claim that results may be generalized to an entire teacher population; rather, the intention of this study was to explore the meeting place of assessment conceptions and assessment literacy in a meaningful context. It is recommended that results and conclusions be understood as (a) presenting a particular case and findings in a meaningful context and (b) as guidance for further, larger scale research.

6. Results

6.1. Knowledge and skills achievement

Of the 32 graduate students enrolled in the course, 30 completed the course. Performance was scored using an A+ to F grade scale. For the determination of means and trends, grade scale was converted to a numeric scale. D grades were given values 2–4, C grades 5–7, B grades 8–10, and A grades 11–13. Means for the reflection paper, project, and final exam ranged from 10.00 to 10.97 (B+ to A–). The mean total grade without participation was 10.34 ($SD = 2.30$). The correlation of the three assignments to the course total were $r = .61$ for the 1st assignment, $r = .89$ for Assignment two, and $r = .85$ for Assignment three, suggesting that the formative feedback after the first assignment helped students align their performance with the overall course goals. Hence, the class performed well on the assessed knowledge and skills of the course.

Interestingly, the achievement of the six interviewed students did not differ considerably from these same score values. Their mean performance on the three assignments ranged from 10.50 to 11.00, with a total mean of 10.75 ($SD = 1.49$), giving a trivial effect size difference of $d = .19$. This is evidence that interviewees’ achievement corresponds closely to that of the whole participant group of 30.

6.2. Conceptions of assessment

Analysis of interview data yielded three major code categories with corresponding child codes (Figure 1). Emotional language was woven into the participants’ responses and coalesced into two affective codes: positive and negative. Outcomes of using assessment were key for interviewees. Interestingly, interviewees expressed less distinct variation regarding particular achievement outcomes. They were more concerned with two issues: the consequences of assessment for students (including non-achievement consequences such as well-being) and fairness, meaning the degree to which assessment outcomes yielded equitable results. The third category, expressed knowledge concerned use of

vocabulary, skills, and concepts directly relevant to the course. Participants evidenced change between the first and second interviews, suggesting the course had some mediating influence, notably in one category. The pattern of this change varied slightly between pre-service and practicing teachers, especially regarding whether conceptions formed prior to the course changed.

6.2.1. *Affective*

Participants expressed strong feelings about assessment. Language polarized around positive and negative expression. No one took an emotionally neutral stance on assessment.

Positive feelings toward assessment linked strongly with AfL priorities. For example, in her first interview, Iris, a pre-service teacher said,

If you maintain a positive attitude and you are reinforcing positive feedback, you build good self-esteem in a child, and that's what assessment should do.

This quote illustrates a general trend among participants: they expressed positive conceptions about assessment as improving students, but often in non-academic or indirect ways (i.e. "good self-esteem"). While positive feelings did focus on direct academic capacity building, this was rarely linked to summative intentions (e.g. determining an achievement score).

Negative responses did not emerge as often, but when they did, they were vivid. Negative language most often accompanied personal experiences, dating back into childhood. In her first interview, Svetlana, a 20-something pre-service Russian émigré vividly described her negative experiences with assessment as a student:

It was horrible. I came from the Russian state education, with the punishment system and the test results. As a child, I felt embarrassed if I didn't pass something or if I spelled words wrong in dictations. If my math results were the worst in the class, I felt so bad. It just killed my personality and I got so many complexes and (have) so many bad memories about being tested.

Svetlana was a high achiever across all the course assessments and earned a final grade of A. Svetlana's affective language in her post-course interview did not waver, though. In discussing her plans on running assessment in her classroom, Svetlana expresses pessimism and a new connection to the category, Outcomes and specifically the code, Fairness:

It's upsetting ... They are checking the abilities for all the students, the same age, the same group of kids and so on; but at the same time, they're all different. That's what isn't fair.

For both pre-service and practicing teachers, affective language often appeared alongside discussion of personal experiences with assessment, often dating back into childhood. For practicing teachers, affective language also accompanied discussions of professional experience. In the second interview, pre-service teachers tended (as Svetlana did) to tether their affective language to new expressed knowledge. Interestingly, neither pre-service nor practicing teachers demonstrated affective change between the first and second interviews.

6.2.2. *Outcomes*

Participants expressed variation in their concepts of the intended functions, mechanics and desirable outcomes of assessment. There were, however, two stable codes that emerged: (1) assessment carries consequences and (2) assessment must operate in a manner that is fair to students. Participants used consequential language alongside positive and negative emotional reactions (i.e. affective codes). Just as no one took an emotionally neutral stance on assessment, none of the participants considered assessment to be irrelevant.

Fairness centered on recognizing and honoring the needs and circumstances of the individual student. As two practicing teachers explained:

Nkem: Fairness is always vital, but what's fair for one person may not be fair for another person.

Ivan: I don't think anything should be ignored, but you should always take certain things into account. Like if a teacher is assessing a student, the teacher might notice a student had a difficult day or didn't get any sleep or didn't eat breakfast, and see something is weird.

Conversely, absence of fairness in assessment was defined by an overemphasis on traditional, summative assessment. As Svetlana explained in her second interview:

This is the data and we need it but as the teacher, I would prefer to grade my students not based on their unit tests or standardized test; I would rather include different ways to assess and incorporate that into their grades.

Participants were particularly concerned with consequences. They tied consequences closely to holding the individual teacher accountable for allowing assessment to demonstrate individual strengths and meet diverse needs.

Julie (pre-service teacher): As a teacher it's your job to make sure that your assessment allows students to demonstrate individual strengths... so that you're able to teach every single child. You should be held accountable if you were not able to reach the student.

Julie, along with two other pre-service teachers suggests that accountability should center on assessment facilitating the teacher being able to "reach the student." Interestingly, this was not a conception expressed by any of the practicing teachers. As with the affective codes, participants did not express significant change between their two interviews.

6.2.3. *Expressed knowledge and skills*

Both pre-service and practicing teachers evidenced growth in expressed knowledge and skills. The interview data suggest some differences between the two groups, however. During the first set of interviews, pre-service teachers' vocabulary was frequently vague or general. What knowledge they demonstrated emerged around pre-service participants' personal histories and experiences as students. By the end of the course, however pre-service participants demonstrated more concretized language and used definitions and concepts congruent with the course curriculum.

The two codes for this category were Utilization and Change. Change was evidenced through achievement and interview data for both groups. For pre-service teachers, Utilization also represented an ontological shift from discussing assessment they experienced as students to assessment practices they intended to employ as teachers.

In her first interview, Iris expresses the positive conception of assessment as helping students' learning. As the earlier provided quote illustrates, her focus is on helping self-esteem. In her second interview, Iris discusses how assessment may help students' academic achievement:

If I do low stakes assessment, like a quiz or even some questioning, this gets them ready. I can try and match whatever I'm giving them with a larger assessment; that could be my unit test or maybe a grade-level test. Then by the time they're up to that, they are better prepared.

Svetlana speaks from the perspective of a student in her first interview; in her second interview, she uses concretized concepts like standards to suggest how she will act as a teacher:

Assessments and standards; they are like your right hand and left hand. Of course, when you assess you have to have the standards. The teacher should set the standards before he or she makes tests. Each standard will guide how you assess and what do you assess. The standards should apply to assessments all the time.

Practicing teachers employed a more concretized vocabulary during the first set of interviews and positioned themselves from the onset as teachers. By the second interview, the practicing teachers more frequently expressed themselves using the vocabulary and concepts specifically emanating from the course. In his second interview, Nkem revisited the theme of fairness, but with greater conceptual depth, linking fairness to the idea of varying standards according to appropriate expectations:

It's fairness that says that you have an ability to perform at the same level. It doesn't mean that you have the skills to answer the same questions. If you have a child who has ADHD, you can't give that child the same test with the same time restraints as a child who has no challenges. That's where the fairness comes in. Am I assessing the child based on his ability to accomplish the assessment and the material given to him? Once you've done that, then the assessment is fair.

It appeared that both pre-service and practicing teachers acquired specific skills and knowledge that they did not possess prior to the course. This matches the results of mean and assignment correlation from the quantitative data. Both groups integrated the course material and vocabulary into the schema of their preexisting conceptions, knowledge, and skills. From a stage-three assessment literacy perspective, this is progress.

This category contrasts with other categories. Pre-service teachers demonstrated some conceptual change; practicing teachers demonstrated a shift in vocabulary and some change in conceptual clarity. The picture that emerged around the other coding categories was an absence of change for either group despite any mediating influence of the course. Consequence, Fairness, as well as Positive and Negative Affective codes remained constant for each participant between their first and second interview. It is worth noting, too there were multiple points at which Expressed Skills and Knowledge had an axial connection to one of these categories. Given these results, there is cause to question whether changes in current understandings of assessment literacy will lead to more literate practice beyond the course, in a professional context.

7. Discussion

The results of this study suggest that a course on assessment purposefully designed according to stage-three assessment literacy principles may appear to achieve desired goals. Appearances may be misleading, though. Quantitative and qualitative results suggest an increase in assessment literacy occurred for both practicing and pre-service teachers. These results include articulation of design, interpretation and use of assessment results (Earl, 2012; Popham, 2008; Stiggins & Chappuis, 2012). Participants appeared aware of some of the formative and summative tensions at work and their need to navigate these successfully; this corresponds to a desirable assessment-related outcomes for a teacher education program (Brookhart, 2011). Because the course constructed learning engagement and assessment together and in alignment with evidence-informed practices, legitimate inferences may be made that student achievement relates to these intended outcomes of enhanced assessment literacy. Thus, the course would seem to be successful, as measured against recognized best practices and priorities in assessment education (Stiggins & Chappuis, 2012).

This is a problematic interpretation. The same data suggests that the current definition of success may be insufficient. A growing body of research indicates that teachers' conceptions of assessment impact their application of skills and knowledge (i.e. assessment literacy) (Barnes et al., 2015; Brookhart, 2011; Deneen & Boud, 2014; Fulmer et al., 2015). Teachers' conceptions about the purposes of assessment have been shown to influence implementation of assessment practices at all educational levels (Barnes et al., 2015; Brookhart, 2011; Deneen & Boud, 2014; Fulmer et al., 2015). Positive conceptions of assessment (e.g. assessment is for enhancing students' learning) have been

shown to precipitate beneficial assessment practices, while negative conceptions of assessment (e.g. assessment is bad for students or irrelevant to learning) may play a significant role in teachers resisting or even subverting assessment policies and intended practices (Brown, 2008; Deneen & Boud, 2014). Brown (2004, 2006, 2008) has added to this framework by proposing an “anti-purpose” in which assessment is conceived as being irrelevant or bad for education. It is clear from the interviews in this study that negative evaluations of assessment exist especially when it becomes extremely accountability-oriented, whereas positive pedagogical views were intended by the course program. It is equally clear that while one code category and achievement results indicated growth, the other more conceptual categories indicated no growth. Thus, a disconnection is evident, with clear implications for the application of assessment literacy.

The contrasting result of this study is that despite gains in AL, participant conceptions of the purpose and nature of assessment fundamentally did not change. What did not occur was a significant change in their preexisting conceptions that assessment was both about evaluative accountability and was negative. The robustness of these conceptions may have to do with each participant’s personal history and personal relationship to assessment derived from their own assessment career (Ecclestone & Pryor, 2003). It is without doubt that, given the way high stakes evaluative assessments were being used in New York (Ravitch, 2011; Shulman et al., 2008), having a negative affective response was legitimate especially if participants were dealing with the fallout of school closures, teacher retrenchment, or students being kept in grade. Additionally, since testing results are used extensively in many jurisdictions to determine educational opportunities and resources, it is not surprising that participants in this study (e.g. Svetlana) might have strongly negative (colored by anxiety and worry concerning their own futures) ways of thinking about assessment. It is worth noting that research elsewhere in the United States has confirmed that teachers have conceptions of assessment that affect their capability of becoming assessment literate (Benson, 2013; Ludwig, 2013). Thus, this study affirms previous findings that individuals’ conceptions of assessment are difficult to change even when the fundamentals of assessment literacy may change (Guskey, 2007). As Brown (2004) argued, there seems to be little value in helping teachers gain greater assessment literacy, if they consider that assessment is either irrelevant or bad for students. This carries the theoretical implication that a new stage of assessment literacy may be required, in which conceptions of assessment are accounted for. A practical implication is that better benchmarks are needed for establishing the value of a course on assessment or more generally the appropriacy of teachers’ practices of assessment.

This study shows that while a course in assessment designed to modern, defensible specifications may provide gains in skills and knowledge, this cannot be understood as sufficient. Teacher assessment practices will not be consistently or reliably enhanced if conceptions of assessment are not also enhanced.

Courses in assessment, perhaps more than any other aspect of teacher education, must address preexisting conceptions and beliefs and their causes (Brookhart, 2011; Shepard, 2006). Teacher education programs therefore need to integrate skills and knowledge with beliefs, values, and attitudes toward assessment in an effective conceptual change model. This will be challenging; changing beliefs within a relatively short time span is unrealistic (Guskey, 2007). It is not entirely surprising that even in a well-designed course, students would gain knowledge and skills without developing or shifting into conceptual understanding that handles the tensions between pedagogical and accounting functions. A practical implication of these results is that assessment requires greater longevity; the traditional course-based approach, however salutary is fundamentally too brief an engagement to allow for conceptual shift. Thus, a greater program-level emphasis on assessment literacy taking into account students’ conceptions, the ways in which beliefs form, and their impact on effective practice is needed. This might productively address tensions, create more impact in shifting conceptions of assessment from maladaptive to adaptive, and help move research and teaching on assessment literacy into its next stage.

8. Conclusions

Modern assessment literacy emerges from competing tensions. Assessment literacy may be at risk if teachers fear assessment and testing, have negative perceptions of assessment, lack adequate training, or face strong pressure to place accountability purposes over improvement purposes (Fives & Buehl, 2012; Popham, 2008; Ravitch, 2011; Stiggins, 2004). Can teacher education toward assessment literacy help teachers negotiate these tensions through enhancing assessment literacy? If we account for change in AL through evaluating expressed skills and knowledge through well-aligned and well-designed assessment tasks, the results of this study would seem to support that at least the skill and knowledge components can and do develop.

However, it seems problematic to discuss what we are interested in (i.e. sophisticated understanding and use of assessment to improve teaching and learning) as literacy, if we cannot predict what will change in practice. Neither current definitions of assessment literacy nor models of engagement within teacher education sufficiently ensure predictable, positive change in assessment practice. Hence, a more evolved definition of assessment literacy and new approaches to assessment in teacher education are needed to achieve impact. The current model of a course on assessment may simply not present sufficient time to precipitate this impact. A sustained program-level engagement with assessment may be needed. Through that engagement, students' conceptions of assessment must be addressed. Thus, teacher education programs may have to innovate structure and content of their curricula while researchers do the same with theoretical understandings of assessment literacy.

Changes to theory and teacher education are not enough if tensions in practice remain unresolvable. At the time of this study, it is entirely rational to conceive of assessment as a negative accountability tool, in the context of New York City's policies and administrative practices. We evolved from stage two assessment literacy because of a schism between a conceptual ideal and the real experience of teaching. To be an assessment literate teacher requires knowledge, skill, and beliefs appropriate to the context in which assessment is deployed. Focusing solely on theory and teacher education, without considering the power of policy and context is untenable. Rather than shaping teacher beliefs towards ruthlessly pragmatic but maladaptive practice, we should instead call for assessment policy and practice to better align with more equitable and beneficial outcomes. While this may never eliminate fundamental tensions, doing so may at least make such tensions bearable as teachers conceptualize and practice assessment.

Funding

The authors received no direct funding for this research.

Author details

Christopher Charles Deneen¹

E-mails: Christopher.d@nie.edu.sg, cdeneen212@gmail.com

Gavin T.L. Brown²

E-mail: gt.brown@auckland.ac.nz

ORCID ID: <http://orcid.org/0000-0002-8352-2351>

¹ Curriculum, Teaching and Learning, National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616, Singapore.

² Faculty of Education & Social Work, School of Learning, Development & Professional Practice, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.

Citation information

Cite this article as: The impact of conceptions of assessment on assessment literacy in a teacher education program, Christopher Charles Deneen & Gavin T.L. Brown, *Cogent Education* (2016), 3: 1225380.

References

Akerlind, G. (2005). Learning about phenomenography: Interviewing, data analysis and the qualitative research paradigm. In J. A. Bowden & P. Green (Eds.), *Doing*

Developmental Phenomenography (pp. 63–73). Melbourne, VIC: RMIT University Press.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: National Council on Measurement in Education.

Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs* (pp. 284–300). New York, NY: Routledge.

Bennett, R. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18, 5–25.

<http://dx.doi.org/10.1080/0969594X.2010.513678>

Benson, III, T. P. (2013). *Sixth through eighth grade teachers' conceptions (beliefs) about assessment and practices* (EdD Educational Leadership unpublished doctoral dissertation). Wingate University, Matthews, NC.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347–364.

<http://dx.doi.org/10.1007/BF00138871>

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30, 3–12.

<http://dx.doi.org/10.1111/emip.2011.30.issue-1>

- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11, 301–318.
<http://dx.doi.org/10.1080/0969594042000304609>
- Brown, G. T. L. (2006). Teacher's conception of assessment: Validation of an abridged version. *Psychological Reports*, 99, 166–170. doi:10.2466/pr0.99.1.166-170
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York, NY: Nova Science Publishers.
- Brown, G. T. L., & Harris, L. R. (Eds.). (2016). *Handbook of social and human conditions in assessment*. New York, NY: Routledge.
- Carless, D. (2007). Conceptualizing pre-emptive formative assessment. *Assessment in Education: Principles, Policy & Practice*, 14, 171–184.
<http://dx.doi.org/10.1080/09695940701478412>
- Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment & Evaluation in Higher Education*, 39, 577–591.
<http://dx.doi.org/10.1080/02602938.2013.859654>
- Deneen, C., Brown, G. T., Bond, T. G., & Shroff, R. (2013). Understanding outcome-based education changes in teacher education: evaluation of a new instrument with preliminary findings. *Asia-Pacific Journal of Teacher Education*, 41, 441–456.
<http://dx.doi.org/10.1080/1359866X.2013.787392>
- Deneen, C. C., & Deneen, J. (2008). *Assessment evaluation and measurement for classroom teachers*. Lanham, MD: Rowman & Littlefield.
- Earl, L. M. (2012). *Assessment as learning: Using classroom assessment to maximize student learning*. Newbury Park, CA: Corwin Press.
- Ecclestone, K., & Pryor, J. (2003). 'Learning careers' or 'assessment careers'? The impact of assessment systems on learning. *British Educational Research Journal*, 29, 471–488.
<http://dx.doi.org/10.1080/01411920301849>
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook: Individual differences and cultural and contextual factors* (Vol. 2, pp. 471–499). Washington, DC: APA.
<http://dx.doi.org/10.1037/13274-000>
- Fulmer, G. W. (2011). Estimating critical values for strength of alignment among curriculum, assessments, and instruction. *Journal of Educational and Behavioral Statistics*, 36, 381–402.
<http://dx.doi.org/10.3102/1076998610381397>
- Fulmer, G. W., Lee, I. C. H., & Tan, K. H. K. (2015). Multi-level model of contextual factors and teachers' assessment practices: An integrative review of research. *Assessment in Education: Principles, Policy, and Practice*, 22, 475–494. doi:10.1080/0969594X.2015.1017445
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3–31.
- Guskey, T. R. (2007). Multiple sources of evidence: An analysis of stakeholders' perceptions of various indicators of student learning. *Educational Measurement: Issues and Practice*, 26, 19–27. doi:10.1111/j.1745-3992.2007.00085
- Hallinger, P., Murphy, J., & Hausman, C. (2013). Conceptualizing school restructuring: Principals' and teachers' perceptions. In C. Dimmock (Ed.), *School-based management and school effectiveness* (pp. 22–40). London: Routledge.
- Leighton, J. P., Gokiart, R. J., Cor, M. K., & Heffernan, C. (2010). Teacher beliefs about the cognitive diagnostic information of classroom- versus large-scale tests: Implications for assessment literacy. *Assessment in Education: Principles, Policy & Practice*, 17, 7–21.
<http://dx.doi.org/10.1080/09695940903565362>
- Ludwig, N. (2013). *Exploring the relationship between K-12 public school teacher's conceptions of assessment and their classroom assessment confidence levels* (Ed.D. Unpublished doctoral dissertation). Regent University, Virginia Beach, VA.
- Marton, F. (1986). Phenomenography: A research approach to investigating different understandings of reality. *Journal of Thought*, 21, 28–49.
- McGuinn, P. J. (2006). *No child left behind and the transformation of federal education policy, 1965–2005*. Lawrence, KS: University Press of Kansas.
- Plake, B. S., & Impara, J. C. (1997). Teacher assessment literacy: What do teachers know about assessment? In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 53–58). San Diego, CA: Academic Press.
- Popham, J. W. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, 62, 82.
- Popham, W. J. (2008). *Transformative assessment*. Washington, DC: ACSD.
- Popham, W. J. (2010). *Everything school leaders need to know about assessment*. Newbury Park, CA: Corwin Press.
- Popham, W. J. (2013). *Classroom assessment: What teachers need to know* (7th ed.). New York, NY: Pearson.
- Ravitch, D. (2011). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.
- Schwartz, I. (2012). *Summary of the provisions of New York's elementary and secondary education act (ESEA) waiver request*. Albany, NY: State Department of Education, Office of P-12. Retrieved from <http://www.p12.nysed.gov/accountability/documents/ESEAWaiverFieldMemo053012FINAL.pdf>
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, 4–14.
<http://dx.doi.org/10.3102/0013189X029007004>
- Shepard, L. A. (2006). Classroom assessment. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 623–646). Westport, CT: Praeger.
- Shulman, V., Sullivan, S., & Glanz, J. (2008). The New York city school reform: Consequences for supervision of instruction. *International Journal of Leadership in Education*, 11, 407–425.
<http://dx.doi.org/10.1080/13603120802183905>
- Spradley, J. P. (1979). *The ethnographic interview*. Belmont, CA: Wadsworth Cengage Learning.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72, 534–539.
- Stiggins, R. J. (2004). *Classroom assessment for student learning: Doing it right—using it well*. New York, NY: Pearson Education.
- Stiggins, R. J., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice*, 44, 11–18.
http://dx.doi.org/10.1207/s15430421tip4401_3
- Stiggins, R. J., & Chappuis, J. (2012). *An introduction to student-involved assessment for learning*. Boston, MA: Pearson Education.
- US Department of Education. (2011). *Race to the top fund* (CFDA Number: 84.395). Retrieved from <http://www2.ed.gov/programs/racetothetop/index.html>



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Education (ISSN: 2331-186X) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

