New Teacher Assessment Literacy: Determining and Narrowing the Gaps

Christopher Adamson

University of Calgary

Assessment literacy is a fundamental prerequisite for effective student learning; therefore, determining and narrowing the gaps in teachers' assessment literacy is an important educational endeavour. The purpose of this research was to explore the gaps in new teacher assessment literacy within a rural Alberta school division. The researcher administered a modified assessment literacy inventory to teachers within their first four years of practice. Results have indicated that gaps in new teacher assessment literacy exist in four of the nine standards used; choosing assessment methods, developing assessment methods, administering, scoring, and interpreting results, and using assessment results in decision making. These findings reflect the need to improve the assessment literacy within this context and they hint at a more widespread issue. This article offers recommendations to narrow the gaps with tailored professional development through professional learning communities.

Keywords: assessment literacy, student learning, new teacher

Improving student achievement is a priority for all educational stakeholders and administration are continuously seeking effective approaches to reach these priorities. Literature indicates that student achievement is tied to sound assessment practices (Black, Harrison, Lee & Marshall, 2003; Campbell & Collins, 2007; Gregory, Cameron & Davies, 2000; Mertler, 2004; Mertler & Campbell, 2005; Stiggins, 2008) and that new teachers make assessment-related decisions without sufficient confidence or training (Brookhart, 2001; DeLuca, 2012; Popham, 2009). Gaps in new teacher assessment literacy exist and need to be identified, explored, and narrowed to ameliorate student achievement priorities.

The term assessment literacy was first introduced by Stiggins (1991) who stated that assessment literate educators effectively use assessments that produce clear, specific, and rich data that reflects a precisely defined learning target. Assessment literacy encompasses the knowledge and skills educators use to identify or create assessments optimally designed for various purposes, and the use of the evidence gathered to make appropriate decisions to improve instructional decisions and subsequently students' learning (Kahl, Hofman, & Bryant, 2013). The existing research attempts to identify the assessment literacy of preservice, new, and experienced teachers using both qualitative and quantitative methods (DeLuca & Klinger, 2010; Mertler, 2003; Remesal, 2011). Most qualitative methods determine teachers' perceived

confidence levels in assessment practices and the quantitative methods measure teachers' assessment literacy using measures such as the Assessment Literacy Inventory (ALI; Mertler & Campbell, 2005) or the Classroom Assessment Literacy Inventory (CALI; Mertler, 2003). The majority of these studies indicate that teachers need to improve their assessment literacy (Arter, 2001; Brookhart, 2001; Popham, 2006). Despite an increase in research regarding assessment literacy, "many of today's teachers know little about educational assessment" (Popham, 2009, p. 5) and teachers generally lack confidence in their assessment practices, and are in dire need of assessment training, resources, and opportunities (Volante & Fazio, 2007). However, for training to be effective, it must be tailored to narrow the specific gaps in the assessment literacy of teachers.

To address this problem, this study focused on determining the gaps in new teacher assessment literacy within a rural Alberta school division using a version of the ALI that was modified by the researcher. This article offers recommendations on how to provide tailored professional development (PD) to improve the assessment practices of these teachers. For the purpose of this study, new teachers are defined as those within their first four years of teaching. This study has the potential to impact teacher assessment practices and subsequently, student achievement. The results will be relevant to teacher participants, post-secondary faculty, school administrators, school division learning coaches, and researchers interested in the topic. Each stakeholder will be better equipped to make data-driven decisions regarding assessment-related PD with the purpose of narrowing the specific gaps in new teacher assessment literacy. Teachers use diagnostic and formative assessments to best meet the learning needs of their students. These same assessments will help determine and narrow the gaps in new teacher assessment literacy. The following research question guided this study: Where do the gaps in new teacher assessment literacy exist in this rural Alberta school division, as measured by the modified ALI?

Value of Assessment

Black and Wiliam (1998) determined that a focus by teachers on assessment practices produced a substantial increase in students' achievement, and Wiliam (2018) stated "the fact that assessment is essential for effective instruction is so obvious" (p. 42). Also, Stiggins (2002) indicated that to maximize student achievement, educators must pay far greater attention to the improvement of classroom assessment; both assessment of learning and assessment for learning. Stiggins (2006) outlined five indicators of sound assessment practices: clear purposes: assessment processes and results serve clear and appropriate purposes; clear targets: assessments reflect clear and valued student learning targets; sound design: expectations are translated into assessments that yield accurate results; effective communication: assessment results are managed well and communicated effectively; and student involvement: students are involved in their own assessment. A meta-analysis conducted by Kingston and Nash (2011) clearly outlined the evidence of the positive impact formative assessment practices has on student achievement and the importance of assessment training.

Gaps in Assessment Literacy

According to Stiggins (1991), "most decision makers - educators and non-educators alike - are not sufficiently literate in the basics of assessment to know whether their achievement data

are sound or unsound" (p. 536); therefore, they struggle to make sound educational decisions. Despite this long awareness of a need for teacher competency in the area of assessment, the literature suggests that teachers and administrators continue to lack relevant assessment training and skills. The majority of teacher education programs do not adequately prepare preservice teachers to fulfill their expected classroom assessment responsibilities and new teachers continue to feel unprepared in this area (DeLuca & Bellara, 2013; Popham, 2009; Stiggins, 2010). Likewise, assessment training is rare in instructional leadership programs designed for administrators (Stiggins, 2008). This deficiency in new teacher assessment literacy can cripple the quality of education (Popham, 2009). This is particularly true in rural Alberta school divisions due to the high levels of staff turnover and the struggle to maintain a strong professional knowledge base regarding assessment. It is estimated that teachers spend between 30% and 50% of their professional time engaged in assessment practices (Stiggins, 1999) and that assessment literacy is increasingly being recognized as an integral part of teacher professionalism (Brookhart, 2002). Why are teachers' assessment practices inadequate if they are investing so much of their valuable time on this endeavour and where do these assessment related inadequacies exist?

Measuring Assessment Literacy

To address the issue of assessment illiteracy, we must first determine the current level of assessment literacy and any efforts to improve it. According to Gotch and French (2014), the most widely used measure was the Teacher Assessment Literacy Questionnaire (Plake, Impara, & Fager, 1993), later revised as the Classroom Assessment Literacy Inventory (Mertler, 2003), which has appeared in 11 studies. These tests were intended to be objective and measure teacher knowledge regarding the prescribed competencies and to identify strengths and weaknesses in their assessment literacy (Xu & Brown, 2016). Volante and Fazio (2007) indicated that the ALI developed by Mertler and Campbell (2005) could be used "as a diagnostic instrument geared toward the identification and remediation of classroom assessment misconceptions or weaknesses of in-service teachers" (p. 762). Mertler and Campbell (2005) indicated that when the ALI is used with preservice teachers the internal consistency of the measure demonstrates acceptable reliability within the measure (r_{KR}= .74). Also, the validity of the ALI has been tested by Hailaya, Alagumalai, and Ben (2014) and according to their research,

> The ALI has some psychometric qualities that make it useful for measuring teachers' assessment literacy. At the item level, the ALI can be a potential instrument in examining teachers' knowledge on classroom concepts and application, and can be used among in-service teachers. (Hailaya, Alagumalai, & Ben, 2014, p. 313)

Gotch and French (2014) expressed a need for increased work on assessment literacy measures and indicate that both the CALI and ALI fail to incorporate aspects of formative assessment. Formative assessment is a crucial part of classroom assessment that is often missing and according to Brookhart (2011), "what's new in formative assessment is the importance of students as formative decision-makers who need information of a certain type (descriptive) at a certain time (in time to act) in order to make productive decisions about their own learning" (p.

4). In addition, the extent to which students are involved in their own assessment is determined by their understanding of classroom learning outcomes; therefore, teachers must clarify and share learning intentions and criteria for success with their students (Wiliam, 2014). Providing effective feedback is also key to effective formative assessment (Hattie & Timperley, 2007). Wiliam, Lee, Harrison, and Black (2004) reported an effect size of 0.32 on student achievement when teachers participated in collaborative PD focused on developing formative assessment strategies. Therefore, due to the importance of formative assessment and feedback, a modified version of the ALI that incorporates both formative assessment and feedback was used for this study.

Gaps in Assessment Training

Research has continued to show that teachers' assessment and evaluation practices are largely incongruent with recommended best practice (Galluzo, 2005; Mertler, 2003; Xu & Brown, 2016), which is alarming given the growing trend towards assessment-based accountability models within North America (Volante & Fazio, 2007). According to Stiggins (2002), there continues to be relatively little emphasis on assessment directed PD in North America. Unfortunately, assessment practices covered in course work and practica are incomplete or superficial in many teacher education programs, leaving graduates unable to effectively meet the demands of today's classrooms (Kahl et al., 2013). In addition, performance measures used to measure the assessment literacy of preservice teachers are inadequate at determining candidates' mastery (Kahl et al., 2013). Teacher educator programs need more than a brief mention of assessment in a course and must place more emphasis on developing the assessment literacy of their preservice teachers (Popham, 2011; Siegel & Wissehr, 2011). Findings by DeLuca and Klinger (2010) support the need for assessment training with specific learning topics identified to develop the assessment literacy of preservice teachers.

As graduates leave their teacher education programs, the research suggests additional assessment training is required to become assessment literate professionals. Therefore, effective PD for new teachers must fill this void. A current systematic review by Vangrieken, Meredith, Packer, and Kyndt (2017) indicates the best approach to teacher learning is through structured collaboration with their peers complemented with targeted learning sessions; not the one-shot workshops teachers typically experience. Some evidence indicates that schools with embedded collaboration also have higher levels of student achievement (Goddard, Goddard, & Tschannen-Moran, 2007). Koh (2011) also provided evidence to support this approach by reporting positive effects for students when their teachers showed improvement in assessment literacy after attending PD with a focus on assessment.

Gaps in the Literature

Despite the small pockets of success, limited research has been conducted to better understand the assessment literacy of classroom teachers (Volante & Fazio, 2007). Based on the research by Gotch and French (2014), it appears that current assessment literacy measures are weak at assessing teachers' ability to assess.

Furthermore, it is apparent that PD initiatives are most effective when a comprehensive understanding of the specific gaps exists and training is tailored to address those gaps (Stiggins, 1991; Volante & Fazio, 2007). According to Mertler (2009) there is little research on teachers' current assessment practices from which to construct PD structures aimed at promoting teacher assessment literacy. Stiggins (1991) stated that "it is essential that we provide high-quality, efficient in-service training for all those teachers and administrators who have completed their degree programs without relevant assessment training" (p. 358). Current and future educators in need of assessment training can be identified using a similar approach to the method used in this study and once identified, tailored PD can be implemented. This approach to addressing the gaps in assessment literacy presents a deficiency in the literature where this study can be located.

The emerging theme is evident; sound assessment is important for student learning and value exists in determining and narrowing the gap of new teacher assessment literacy within this rural Alberta school division.

Research Methods

The inclusion criteria for participants involved teachers in the rural Alberta school division who were within their first four years of teaching due to the importance of providing PD to new teachers as they begin to establish their professional practice. Teachers with five or more years of experience were excluded as data sources because more experienced teachers often have a better understanding of assessment due to their years of practice. In addition, teachers at the researcher's school were excluded from the study due to the researcher's administrative position.

An explanation of the research study and an invitation to participate was communicated via email to teachers who qualified to be participants based on the inclusion criteria. Individuals were given two weeks to complete and return the consent forms. The participants included 10 teachers ($n = 10$). Two participants were in their first year, three in their second year, two in their third year, and three in their fourth year. Their teaching assignments ranged from kindergarten to grade twelve and 60% of the participants were female, 40% were male.

Data Collection and Instrumentation

Once the participants were identified by signed consent forms, they completed the survey under the direct supervision of the researcher at division office and data for the study was collected using the modified assessment literacy inventory (MALI). The original ALI is a user-friendly version of the Teacher Assessment Literacy Questionnaire (Plake, Impara, & Fager, 1993). However, new modifications to the original ALI were necessary to account for an important development in educational assessment; formative assessment and feedback (Brookhart, 2011); these modifications were made by the researcher. Black and Wiliam (1998) also highlighted the importance of formative assessment for improving student achievement. In addition, studies have revealed that student achievement improves when teachers emphasize proper questioning techniques, feedback without grades, peer assessment, self-assessment, and use of formative data to alter instructional strategies (Black et al., 2004; Black & Wiliam, 1998).

The original ALI contained 35 questions that were developed to mimic real-world applications of assessment standards. The ALI consists of five scenarios, each followed by seven

questions. The questions are related to the seven *Standards for Teacher Competence in the Educational Assessment of Students* (1990), in Mertler and Campbell (2005). The modification for this study involved removing one of the scenarios and adding eight questions related to formative assessment and feedback; two questions for each remaining scenario, resulting in a survey of 36 questions. Therefore, each of the nine standards has four related questions. The new questions regarding formative assessment and feedback were created in alignment with the standards F1 (assessment purpose), F2 (learning expectations), U2 (effective feedback), and U3 (instructional follow-up) in the new revised document titled *Classroom Assessment Standards for PK-12 Teachers* (JCSEE, 2015, as cited in DeLuca, LaPointe-McEwan & Luhanga, 2015). Table 1 outlines how the MALI aligns to the *Standards for Teacher Competence in the Educational Assessment of Students* (American Federation of Teachers, 1990) and specific formative assessment and feedback standards from the *Classroom Assessment Standards for PK-12 Teachers* (JCSEE, 2015, as cited in DeLuca, LaPointe-McEwan & Luhanga, 2015). In addition, three questions related to the educational background of the participants were asked following the MALI.

Table 1

Modified Assessment Literacy Inventory Alignment to Standards (1990 and 2015)

Description	Standard	Item Alignment
Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.	Choosing Assessment Methods	1, 10, 19, 27
Teachers should be skilled in developing assessment methods appropriate for instructional decisions.	Sound Design	2, 11, 20, 29
The teacher should be skilled in administering, scoring and interpreting the results of both externally-produced and teacher produced assessment methods.	Scoring	3, 12, 21, 30
Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.	Using Results	4, 13, 22, 31
Teachers should be skilled in developing valid pupil grading procedures which use pupil assessments.	Grading	5, 14, 23, 32
Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.	Communicating Results	6, 15, 24, 33
Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.	Ethical Assessment	7, 16, 25, 34

Classroom assessment practices should have a clear purpose that supports teaching and learning. Learning expectations should form the foundation for aligning classroom assessment practices with appropriate instruction and learning opportunities for each student.	Foundations of Formative Assessment	8, 17, 26, 35
Classroom assessment practices should provide timely and useful feedback to improve student learning. Analysis of student performance should inform instructional planning and next steps to support ongoing student learning.	Use of Formative Assessment	9, 18, 27, 36

Note. Adapted from American Federation of Teachers (1990).

Data Analysis

Following data collection, descriptive statistics regarding minimum, maximum, *mean*, standard deviation (*SD*), levels of achievement, and percentage (%) were used to analyze the data.

Results

The minimum score, maximum score, *mean* (maximum of four), and *SD* for each of the nine standards are presented in Table 2. In addition, the minimum and maximum total score of all participants is displayed with associated *mean* and *SD*.

The results revealed that the highest and lowest total score of all participants was 28 and 15, respectively, with a *mean* of 23.6 and *SD* of 4.06. The *mean* scores based on the MALI ranged from 2.0 to 3.7, with a *SD* between 0.48 and 1.23. The standard with the highest *mean* was standard 9: use of formative assessment (*mean* = 3.7; *SD* = 0.48). The standards with the lowest *means* were standard 2: sound design (*mean* = 2.0; *SD* = 0.82) and standard 3: scoring (*mean* = 2.0; *SD* = 1.05).

Table 2

Participant Scores on the Modified Assessment Literacy Inventory

Description	Standard	Min.	Max.	Mean	SD
Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.	Choosing Assessment Methods	0	4	2.2	1.03
Teachers should be skilled in developing assessment methods appropriate for instructional decisions.	Sound Design	1	3	2	0.82
Teachers should be skilled in administering, scoring and interpreting the results of both externally-produced and	Scoring	0	4	2	1.05

teacher produced assessment methods.

Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.	Using Results	1	3	2.2	0.79
Teachers should be skilled in developing valid pupil grading procedures which use pupil assessments.	Grading	1	4	2.8	1.03
Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.	Communicating Results	0	4	2.8	1.23
Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.	Ethical Assessment	1	4	2.7	0.95
Foundations 1 - Assessment Purpose Foundations 2 - Learning Expectations	Foundations of Formative Assessment	2	4	3.2	0.63
Use 2 - Effective Feedback Use 3 - Instructional Follow up	Use of Formative Assessment	3	4	3.7	0.48
	Total score out of 36 points	15	28	23.6	4.06

To get a better understanding of the degree of assessment literacy based on the nine standards, participants' scores were divided into three levels; low, medium, and high (Table 3). The low level of achievement represents the number of participants who scored lower than 50 percent on a standard (zero or one out of the four questions correct), medium represents participants who scored 50 - 75 percent on a standard (two or three out of the four questions correct), and high represents participants who scored higher than 75 percent on a standard (all four questions correct). The number of participants that achieved a total score on a low, medium, and high level are also represented in Table 3.

The levels of achievement based on the nine standards of the MALI revealed that participants had scores for eight standards at the medium level and one at the high level (standard 9: use of formative assessment). The standards that were achieved at the lowest levels were the following; standard 2: developing assessment methods, standard 3: administering, scoring, and interpreting results, and standard 4: using assessment results in decision making.

Table 3

Levels of Achievement on the Modified Assessment Literacy Inventory

	Assessment Literacy					
	Low		Medium		High	
Standard	Number	Percent	Number	Percent	Number	Percent
1. Choosing an assessment method	1	10%	8	80%	1	10%
2. Developing assessment methods	3	30%	7	70%	0	0%
3. Administering, scoring, and interpreting results	2	20%	7	70%	1	10%
4. Using assessment results in decision making	2	20%	8	80%	0	0%
5. Using valid student assessments to determine levels of student achievement	1	10%	6	60%	3	30%
6. Communicating assessment results to all stakeholders	1	10%	6	60%	3	30%
7. Recognizing unethical, illegal, and inappropriate assessment methods and uses of assessment information	1	10%	7	70%	2	20%
8. Identifying the purpose of assessment and learner expectations	0	0%	7	70%	3	30%
9. Using feedback and instructional follow up	0	0%	3	30%	7	70%
Total out of 10 participants	1	10%	7	70%	2	20%

Table 4 summarizes the data by highlighting the standards that were achieved at the lowest level (gaps in assessment literacy) and begins to answer the research question. The *mean*, *SD*, and percentage of participants that achieved at a low level for the given standards are displayed.

Table 4

Gaps in New Teacher Assessment Literacy

Standard	Mean	SD	Percentage
2. Developing assessment methods	2	0.82	30%
3. Administering, scoring, and interpreting results	2	1.05	20%
4. Using assessment results in decision making	2.2	0.79	20%
1. Choosing an assessment method	2.2	1.03	10%

Discussion

Gaps in new teacher assessment literacy exist, and the results support this statement. The gaps in new teacher assessment literacy within the rural Alberta school division were primarily within four of the nine standards used in the MALI: standard two: developing assessment methods, standard three: administering, scoring, and interpreting results, standard four; using assessment results in decision making, and standard one: choosing an assessment method.

In the study performed by the original creator of the ALI, (Mertler, 2009), standard two was also the lowest achieved standard among the research participants; however, standard two showed the second greatest amount of improvement when reassessed following a two-week in-service training in classroom assessment. Standard two: developing assessment methods, incorporates questions that discuss validity coefficients, reliability coefficients, item difficulty value, and other assessment vocabulary that the participants possibly were not familiar with. This may explain why it was the lowest achieved standard. In addition, survey questions included the associated numerical values for these terms and perhaps participants couldn't make sense of the statistics.

The work by Stiggins (2010) provided overwhelming evidence that student achievement is strongly related to their teachers' ability to identify or create high-quality classroom assessments and the meta-analysis by Hattie and Timperley (2007) revealed how the productive use of formative assessment data can have a positive impact on student achievement. The low level of achievement on the four standards is concerning since it suggests that participants struggle to develop and select appropriate assessment methods and make sound decisions based on assessment data. However, the lack of assessment-based courses, assessment PD, teaching experience, and available time to digest assessment data appropriately likely accounts for the low achievement on the other standards. New teachers have not had the opportunity to experience the difference between unsound and sound assessment nor do they feel confident or appropriately prepared for assessment when they graduate from their teacher education programs (Volante & Fazio, 2007). Although assessment self-efficacy was not a focus of this study, this statement supports a large body of literature that advocates greater support for new teachers to improve the quality of teaching and teacher retention, which is a struggle in this rural Alberta school division (Volante & Fazio, 2007).

Formative Assessment

The standards that were achieved at the highest level were standards eight and nine; identifying the purpose of assessment and learner expectations, and using feedback and instructional follow-up, respectively. According to Black and Wiliam (1998), the effect size of formative assessments is larger than most found for educational intervention; therefore, it is promising that participants achieved standards eight and nine at a high level. Nonetheless, this may have occurred given the fact that these two standards and associated questions were additions to the original ALI; therefore, these questions had not been exposed to any item analysis (validity testing, item difficulty index, item discriminatory index, distractor analysis, etc.). As a result, the questions are not as rigorous as the other 28 in the MALI. This could explain the high level of achievement that participants experienced on these standards.

Limitations

In addition to the psychometric limitations of standards eight and nine, this study is limited by other factors. The sample size ($n = 10$) should encourage readers to be cautious when making interpretations and generalizations of the whole population based on the results of this study. Although a sample is useful when estimating trends for a population, increasing the duration and scope of this study would allow for a larger sample size and a more accurate picture of new teacher assessment literacy in this rural Alberta school division. Furthermore, generalizations cannot be made for the province or country at large since the results represent a unique subgroup of Alberta. However, there is still value in looking at individual participant scores. When individual participant scores are analyzed and interpreted it becomes apparent to teacher participants, post-secondary faculty members, school administrators, school division learning coaches, and other relevant stakeholders where the gaps in individual new teacher assessment literacy exist. Although, "measuring teachers' mastery of the principles drawn from the knowledge base can give an estimate of teacher assessment literacy, transfer to the practical realm is not guaranteed" (Xu & Brown, 2016, p. 153). Therefore, to get a more holistic view of new teacher assessment literacy, the MALI data should be paired with qualitative data from classroom observations of practical assessment skills.

Implications and Recommendations

With this information, post-secondary faculty members, school administrators, school division learning coaches, and central office executive staff can tailor PD to target the needs of individual participants. A core component of effective PD involves teachers engaging in ongoing professional development that is tailored to their needs (Desimone, 2009). The significance of this study is to demonstrate the purpose and opportunity of measurement tools like the MALI to serve as a diagnostic assessment for teachers. The results of this study have implications for the structure of future PD models given the fact that the data can be used to tailor PD with the intent of improving teachers' assessment literacy and subsequently, student achievement. This endeavour is of interest to all educational stakeholders.

Post-secondary faculty members must begin to close the gap in new teacher assessment literacy by providing more rigorous assessment training in teacher education programs. A

measure such as the MALI would help identify the gaps and thus allow faculty members to tailor their courses to the specific standards of assessment literacy in need of improvement. Preservice teachers would benefit from purposeful introductions to competing assessment philosophies and "specific course work focused on assessment, embedding assessment topics in content and methods courses, and providing real-world opportunities to enable candidates to apply what they have learned" (Kahl et al., 2013, p. 3). In addition, post-secondary faculty members must increase the congruency between their teaching and their practice on the topic of assessment. When faculty members explicitly model sound assessment practices, preservice teachers have an opportunity to experience consistent and effective formative assessment as learners and are far more likely to practice formative assessment in their own classrooms as a result (DeLuca & Volante, 2016; Poth, 2012). Using formative assessment as a means of teaching preservice teachers will potentially trickle down to improve the learning of K-12 students across the entire educational system (DeLuca & Volante, 2016). Munroe, Mitton-Kükner, and Graham (2015) indicated that a fundamental prerequisite for the above recommendations is effective collaboration and collegial team support from like-minded educators. Munroe et al. (2015) stated "meeting regularly throughout the semester to plan course activities, to debrief preservice teachers' responses, and to discuss our marking contributes to our abilities to be adaptive, reflective, open-minded, and organized" (p. 121).

The use of professional learning communities to complement or replace more traditional professional development activities is an effective way to tailor PD for new teachers using the data obtained from this type of research (Popham, 2009). In addition, "in-service teachers may need to utilize daily classroom practices as sites for their assessment literacy development by implementing assessment for learning" (Xu & Brown, 2016, p. 154). Teachers acquire assessment knowledge in the field through reflection and collaboration about classroom experiences (Howley, Howley, Henning, Gilla, & Weade, 2013) and this supports the professional learning community (PLC) approach to narrowing the gaps in new teacher assessment literacy. Incorporating PLCs into a school's culture and structure is an effective way administrators can encourage teachers to share and create common assessments, review research on the power of formative assessments, use evidence of student learning to improve teaching, and be accountable for their results. In order for PLCs to be effective, teacher teams with common professional goals need to exist and by including teacher mentors, new teachers will have access to an invaluable tacit knowledge base. Furthermore, PLCs must meet frequently throughout the year and focus on topics specific to the needs of those teachers and their students. Measurements such as the MALI can influence the focus of PLCs to ensure the gaps in new teacher assessment literacy are being rectified. Teachers must also periodically monitor the implementation of the practices developed in PLCs and the related teacher and student outcomes to determine if positive associations exist between them. PLCs involve teacher teams engaging in a structured collective inquiry to decide on the work that will most benefit their students and when administrators increase their involvement in these teams, they can have a greater impact on teacher practice and subsequently, student learning (DuFour & Mattos, 2013; Hewson, 2013).

Once effective PLCs are established, according to DuFour and Mattos (2013), teachers will be more likely to do the following: take collective responsibility for student learning; share teaching practices; make results transparent; improve their professional practice and student

achievement; share leadership; and express higher levels of professional satisfaction. Teachers that engage in a structured collaborative model for PD improve their practice at greater rates and schools that are characterized by a collaborative culture have higher levels of student achievement (Ronfeldt, Farmer, McQueen, & Grissom, 2015).

Conclusion

According to this literature review, it is evident that sound assessment practices are tied to improved student learning and teachers have varying levels of assessment literacy. Measuring the assessment literacy of a teacher is a difficult task. By interpreting the data from the MALI, educators can make suggestions on where and how to improve new teacher assessment literacy. Gaps exist in new teacher assessment literacy and participants in this rural Alberta school division are no different. The findings reflect the need to improve the assessment literacy of new teachers in this rural Alberta school division, and they raise the question: how assessment literate are all teachers; not just those within their first four years?

Stiggins (1991) indicated that we spend a lot of resources training teachers to produce learning but very little to train teachers to assess the outcomes of those very efforts, and this appears to still be true. Establishing a better understanding of what assessment literacy is and how it can be measured is crucial to improving teacher assessment literacy via tailored PD and subsequently, student learning. A measurement tool that stands up to validity testing and is aligned with contemporary standards, such as the ones developed by the Joint Committee on Standards for Educational Evaluation (JCSEE, 2015, as cited in DeLuca, LaPointe-McEwan & Luhanga, 2015) is necessary. The standards issued by the JCSEE (2015) are the only educational standards approved by the American National Standards Institute which means "great care has been taken to follow a strict process of review to develop standards that are widely supported by researchers and practitioners who are knowledgeable about classroom assessment" (JCSEE, 2015, location 73, as cited in DeLuca, LaPointe-McEwan & Luhanga, 2015). An updated version of the ALI based on the most current assessment standards and paired with classroom observations of practical assessment skills would be beneficial in assessing teacher assessment literacy. This would result in more up-to-date standards and provide valuable data which would allow post-secondary faculty members, school administrators, school division learning coaches, and central office executive staff to make appropriate data-driven decisions regarding assessment-related PD.

Teachers need to shift their paradigm to understand how assessment can drive instruction and positively impact student learning by ensuring their classroom assessments are assessments for learning, rather than of learning. (Popham, 2009; Volante & Fazio, 2007). Quality instruction involves sound decision making and sound decisions require quality assessments and data (Stiggins, 1991).

The ability to tailor PD to what in-service teachers specifically need in the area of assessment is a critical feature to ensure their training is successful (Volante & Fazio, 2007). The PLC process has two powerful motivators for changing adult behavior; irrefutable evidence of better results and positive peer pressure (Elmore, 2004; Fullan, 2010, Hattie, 2009). Furthermore, Popham (2009) stated that "it must be professional development that will supply

the nation's teachers with assessment-related skills and knowledge they need" and these "professional development programs focused on assessment need to be tailored" (p. 5). Investing in PD geared towards assessment will pay dividends in terms of improved teaching and learning (Black & Wiliam, 1998; Volante & Melahn; 2005) and may also help retain some of the most effective teachers (Volante & Fazio, 2007), which is crucial in any educational context including this rural Alberta school division. In addition, Volante and Fazio (2007) believed that "an ounce of assessment literacy promotion may act as a pound of retention cure" (p. 762).

Engaging teachers in structured collaborative PLCs with a focus on their assessment learning needs, identified by measures such as the MALI, is a step in the right direction to improving instruction and subsequently student achievement. Additional research focused on the implementation and effectiveness of the above recommendations on how to improve new teacher assessment literacy would be beneficial to all educational stakeholders.

References

American Federation of Teachers. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, *9*(4), 30–32.

Arter, J. (2001). Learning teams for classroom assessment literacy. *NASSP Bulletin*, *85*(621), 53-65. https://doi.org/10.1177/019263650108562107

Black, P., Harrison, C., Lee, C., & Marshall, B. (2003). *Assessment for learning: Putting it into practice*. Berkshire, England: McGraw-Hill International (UK).

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148. https://doi.org/10.1177/003172171009200119

Brookhart, S. M. (2001, March). The standards and classroom assessment research. *The role of the standards for teacher competence in educational assessment of students in developing the quality teacher assessors.* Paper presented in a symposium at the annual meeting of the American Association of Colleges for Teacher Education. Dallas, TX. Retrieved from https://files.eric.ed.gov/fulltext/ED451189.pdf

Brookhart, S. M. (2002). What will teachers know about assessment, and how will that improve instruction. In R. W. Lissitz & W. D. Schafer (Eds.), *Assessment in education reform: Both means and ends* (pp. 2-17). Boston, MA: Allyn and Bacon*.*

Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, *30*, 3–12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice*, *26*(1), 9-18. https://doi.org/10.1111/j.1745-3992.2007.00084.x

Elmore, R. (2004). *School reform from the inside out: Policy, practice, and performance*. Boston: Harvard Educational Press.

Fullan, M. (2010). *The moral imperative realized*. Thousand Oaks, CA: Corwin Press.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199. https://doi.org/10.3102/0013189x08331140

DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education*, *34*(5/6), 576-591. https://doi.org/10.1080/01626620.2012.730347

DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education, 64*(4), 356-372. https://doi.org/10.1177/0022487113488144

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419-438. https://doi.org/10.1080/0969594x.2010.516643

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2015). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 1-22. https://doi.org/10.1007/s11092-015-9233-6

DeLuca, C., & Volante, L. (2016). Assessment for learning in teacher education programs: Navigating the juxtaposition of theory and praxis. *Journal of the International Society for Teacher Education*, *20*(1), 19-31. Retrieved from https://files.eric.ed.gov/fulltext/EJ1177153.pdf

DuFour, R., & Mattos, M. (2013). How do principals really improve schools? *Educational Leadership*, *70*(7), 34-40.

Goddard, Y. L., Goddard, R. D., & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teacher College Record*, *109*(4), 877–896. Retrieved from https://education.illinoisstate.edu/downloads/casei/collaboration_studentachievement.pdf

Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice*, *33*(2), 14-18. https://doi.org/10.1111/emip.12030

Gregory, K., Cameron, C., & Davies, A. (2000). *Knowing what counts: Self-Assessment and goal setting*. Merville, BC: Connections.

Hailaya, W., Alagumalai, S., & Ben, F. (2014). Examining the utility of assessment literacy inventory and its portability to education systems in the Asia Pacific region. *Australian Journal of Education*, *58*(3), 297-317. https://doi.org/10.1177/0004944114542984

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to student achievement*. New York: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. https://doi.org/10.3102/003465430298487

Hewson, K. (2013). Time shift: Developing teacher teams. *Principal*, *92*(3), 14-17. Retrieved from https://www.naesp.org/sites/default/files/Hewson_JF13.pdf

Howley, M. D., Howley, A., Henning, J. E., Gilla, M. B., & Weade, G. (2013). Intersecting domains of assessment knowledge: School typologies based on interviews with secondary teachers. *Educational Assessment*, *18*(1), 26-48. https://doi.org/10.1080/10627197.2013.761527

Kahl, S. R., Hofman, P., & Bryant, S. (2013). *Assessment literacy standards and performance measures for teacher candidates and practicing teachers*. Prepared for the Council for the Accreditation of Educator Preparation. Dover, NH: Measured Progress.

Klinger, D., McDivitt, P., Howard, B., Rogers, T., Munoz, M., & Wylie, C. (2015). *Classroom assessment standards for prek-12 teachers: Joint committee on standards for educational evaluation* [Kindle Version]. Retrieved from Amazon.ca.

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, *22*(3), 255-276. https://doi.org/10.1080/10476210.2011.593164

Mertler, C. A. (2003). Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education, 33*(2), 76-92. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.556.4217&rep=rep1&type=pdf

Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, *33*(1), 49-64.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*(2), 101–113. https://doi.org/10.1177/1365480209105575

Mertler, C. A. & Campbell, C. (2005, April). *Measuring teachers' knowledge & application of classroom assessment concepts: Development of the assessment literacy inventory*. Symposium conducted at the annual meeting of the American Educational Research Association, Montreal, QC.

Munroe, E. A., Mitton-Kükner, J., & Graham, D. (2015). Inquiring into the assessment education of preservice teachers: A collaborative self-study of teacher educators. *In Education*, *21*(2), 110-126. Retrieved from https://journals.uregina.ca/ineducation/article/view/214

Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, *12*(4), 10-12. https://doi.org/10.1111/j.1745-3992.1993.tb00548.x

Popham, W.J. (2009). Assessment literacy for teacher: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11. https://doi.org/10.1080/00405840802577536

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, *46*(4), 265-273. https://doi.org/10.1080/08878730.2011.605048

Poth, C. A. (2012). What assessment knowledge and skills do initial teacher education programs address? A Western Canadian perspective. *Alberta Journal of Educational Research*, *58*(4), 634-656. Retrieved from https://journalhosting.ucalgary.ca/index.php/ajer/article/view/55670/42588

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education*, *27*(2), 472-482. https://doi.org/10.1016/j.tate.2010.09.017

Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal*, *52*(3), 475-514. https://doi.org/10.3102/0002831215585562

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, *22*(4), 371-391. https://doi.org/10.1007/s10972-011-9231-6

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, *72*(7), 534-39. https://doi.org/10.1177/0031721714553413

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, *18*(1), 23–27. https://doi.org/10.1111/j.1745-3992.1999.tb00004.x

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765. https://doi.org/10.1177/003172170208301010

Stiggins, R. J. (2008). *Assessment manifesto: A call for the development of balanced assessment systems*. A position paper. *Portland, OR: ETS Assessment Training Institute*. Retrieved from https://www.nyscoss.org/img/uploads/file/Assessment_Manifesto_Article_-_Rick_Stiggins.pdf

Stiggins, R. J. (2010). Essential formative assessment competencies for teachers and school leaders. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 233-250). New York, NY: Taylor & Francis.

Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning: Doing it right--using it well*. Portland, OR: Assessment Training Institute.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, *30*(3), 749. https://doi.org/10.2307/20466661

Volante, L., & Melahn, C. (2005). Promoting assessment literacy in teachers: Lessons from the Hawaii school assessment liaison program. *Pacific Educational Research Journal*, *13*(1), 19-34. Retrieved from http://hawaii.edu/crdg/perj/perj2005/perj2005.pdf

Wiliam, D. (2014, April). Toward a theory of classroom assessment as the regulation of learning. In *Formative assessment and contingency in the regulation of learning processes*. Symposium conducted at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Wiliam, D. (2018). How can assessment support learning? A response to Wilson and Shepard, Penuel, and Pellegrino. *Educational Measurement: Issues and Practice*, *37*(1), 42-44. https://doi.org/10.1111/emip.12192

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, *11*(1), 49-65. https://doi.org/10.1080/0969594042000208994

Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*, 149-162. https://doi.org/10.1016/j.tate.2016.05.010