

## **Acquisition of teacher assessment literacy by pre-service teachers: A review of practices and program designs**

**Cherry Zin Oo**

*Yangon University of Education, Myanmar*

**Dennis Alonzo**

*University of New South Wales, Australia*

**Ria Asih**

*Universitas Muhammadiyah Malang, Indonesia*

Teacher assessment literacy research with pre-service teachers (PSTs) has highlighted that they are not well prepared to use assessments to support student learning. Thus, initial teacher education (ITE) programs need to ensure that PSTs are provided with a range of opportunities to acquire both theoretical and practical assessment knowledge and skills. We reviewed assessment programs reported in the literature to develop a framework that will better guide curriculum developers. A total of 12 studies were considered after an initial literature search from three databases that generated 1002 articles from 1998 to September 2020. We reported the characteristics of assessment programs, including their orientation, content focus, outcomes measures, approaches and duration. We also identified some aspects of ITE assessment programs that are least explored.

### **Introduction**

Teacher assessment literacy is a critical component of effective learning and teaching. In many educational institutions, the current level of teacher assessment literacy remains relatively low, including in Australia (Davison & Michell, 2014), the United States (DeLuca et al., 2013), Europe (Tsagari & Vogt, 2017), Hong Kong (Lam, 2015), Philippines (Hailaya, 2014), Indonesia (Ahmad, 2020; Zulaiha et al., 2020), and Myanmar (Hardman et al., 2016). Despite advanced conceptualisations of this construct and many efforts to engage teachers in professional development, teacher assessment literacy remains a big concern. This issue can be traced back to the quality of initial teacher education (ITE) programs (Verberg et al., 2016). Studies in ITE programs show that the curriculum design does not support pre-service teachers (PSTs) for more effective assessment knowledge acquisition and practical skills development (Grainger & Adie, 2014; Oo, 2020). There is a consensus across different contexts that the assessment component of ITE curriculum is insufficient for PSTs to be assessment literate (James & Pedder, 2006; Siegel & Wissehr, 2011). In particular, ITE programs have an insufficient emphasis on developing and administering the assessment and interpreting the results, using data to inform their teaching and improve student learning (Craven et al., 2014). This has been corroborated by PSTs who believe that they do not have sufficient support and preparation to improve their assessment knowledge and skills (McGee & Colby, 2014; Rwehera, 2017; Volante & Fazio, 2007).

To effectively address this issue, researchers have argued that ITE programs need to ensure that PSTs are provided with a range of opportunities to develop their theoretical

knowledge and acquire practical skills in assessment (Grainger & Adie, 2014; McGee & Colby, 2014). Despite making assessment an integral part of the ITE curriculum, research evidence continues to report that PSTs are not fully supported. This is largely due to the quality of the design and implementation of an assessment program for promoting acquisition of PST assessment literacy (Greenberg & Walsh, 2012) embedded in their learning experience.

Assessment experts and curriculum developers have varied views on what assessment programs best support PST assessment literacy. This is evident in different approaches used in assessment courses implemented in various higher education institutions (Xu & Brown, 2016). To develop a more responsive ITE assessment program, a scoping review of the research literature can help us understand current practices, to provide an effective model of designing an ITE assessment program. This is not to say that a single design is needed, but the program's development should be guided by a framework to ensure that all components of the program support PST assessment literacy development.

In this scoping review, we aimed to examine extant ITE assessment programs to answer the following research questions:

1. What are the characteristics of ITE assessment programs in terms of their orientation, content focus, outcomes measure, approaches and duration?
2. What are the limitations of the extant literature in terms of defining an effective ITE assessment program?

### **Assessment literacy: What knowledge and skills are needed?**

In building PST assessment literacy, we need to respond to the question: *What knowledge and skills are needed for teachers to be literate in assessment?* Clarifying this construct will provide an overarching framework for the content of ITE assessment programs and clear measuring of the outcomes from these programs.

The most common conceptualisations used are assessment *for* and *as* learning, those assessments being conducted by teachers regularly or daily to gather data aimed at improving learning were collectively referred to as formative assessment (FA). In contrast, assessment *of* learning is referred to as summative assessment (SA) that includes all type of teacher SA and large scale national and international testing. Researchers have identified the main difference between FA and SA as residing in their aims and effects on students' learning, not in such superficial factors as the format or timing of the assessments (Gipps, 1995; Sadler, 1989). However, with years of researching the distinction and effectiveness of these two types of assessment, many authors have argued that they are supporting each other in terms of functions (Davison, 2007; Taras, 2009) contrary to the claim of Sadler (1989), "many of the principles appropriate to summative assessments are not necessarily transferable to formative assessment, the latter requires a distinctive conceptualisation and technology" (p. 120).

Even Black and Wiliam, who published extensively on FA (1998a, 1998b), have acknowledged that the distinction between FA and SA is irrelevant when assessment is conceptualised in the broader pedagogical model (Black & Wiliam, 2018). A view which was earlier argued by Kennedy et al. (2006) that the distinction between formative and summative assessment is "no longer useful, even though such a distinction has resulted in some excellent research and development work on formative assessment" (p. 14), because summative assessment needs to be used formatively to improve student learning (Davison, 2013; Harlen, 2005; Kennedy et al., 2007).

This conceptualisation of assessment was formalised by the Assessment Reform Group in 1999 with their push to use assessment *for* learning to highlight that any forms and types of assessment should support student learning. They highlighted the term A/L rather than formative assessment as "the term 'formative' itself is open to a variety of interpretations and often means no more than that assessment is carried out frequently and is planned at the same time as teaching" (Assessment Reform Group, 1999, p. 7). Meanwhile, assessment *of* learning is used for the evaluative purpose of students' learning collected at the end of the unit or course.

Therefore, teachers should be prepared to engage all assessment strategies in effective learning and teaching towards improving student learning. In addition, teachers should be prepared to design assessment tasks for different levels of students' ability (Wurf & Povey, 2020) and provide personalised feedback to students (Walker et al., 2021). This definition implies that teachers need to have an adaptive disposition in using their assessment literacy to support individual students (Loughland & Alonzo, 2018). If teachers are literate in assessment, they can use these assessments appropriately. Alonzo (2016) highlighted that teachers' assessment literacy requires them not only to use assessment data to inform their teaching and to increase students' learning, but also to give information to other stakeholders. Alonzo provided strong theoretical and empirical evidence to the dimensionality of teacher A/L literacy. He stressed that "teacher assessment literacy comprises the knowledge and skills to make highly contextualised, fair, consistent and trustworthy assessment decisions to inform learning and teaching to effectively support both students and teachers' professional learning (p. 58)."

## **Issues in ITE assessment programs**

Issues regarding the assessment component in teacher education include insufficient or inappropriate assessment components in teacher preparation, concerns with the quality of assessment courses in teacher education, and insufficient practical knowledge in assessment courses (Brookhart, 1999; Greenberg & Walsh, 2012; Popham, 2011). In what follows, a detailed description of each issue is presented.

Insufficient and inappropriate assessment components in teacher education are one of the key issues that needs to be addressed. According to the findings of Poth (2013) who analysed 57 assessment course syllabi provided in Western Canadian teacher education programs, most assessment courses are related to summative purposes. Only a few courses covered broader assessment concepts, including in the general curriculum or

educational studies courses in teacher education (DeLuca & Klinger, 2010). The impact of this narrow focus on assessment literacy is reflected in the study of Ogan-Bekiroglu & Suzuk (2014) that shows PSTs did not have adequate training related to many areas in assessment.

Another issue is the competing focus of assessment components in teacher preparation. Xu and Brown (2016) indicated that current assessment courses have a different emphasis on assessment content. The general description and the basic concept of the assessment in the course cannot provide adequate support to be literate in assessment (Greenberg & Walsh, 2012). Greenberg & Walsh's analysis of 455 assessment courses from over 180 initial teacher education qualifications demonstrated that only 3% of these assessment programs showed adequate assessment preparation for student teachers in ITE. This issue has also been cited and argued that the assessment course needs to include more than a brief coverage of assessment (Oo et al., 2021; Popham, 2011).

There is also an issue related to the quality of the assessment courses. The assessment component of the teacher education curriculum should be clearly articulated and aligned with the learning objectives of the assessment context (Hill et al., 2014). Brookhart (2011) suggested that communicating assessment results and using assessment data needs to emphasise teacher preparation greatly. Therefore, Ogan-Bekiroglu and Suzuk (2014) pointed out that the “content and context of an assessment course in teacher education programmes are prominent issues” (p. 362).

Due to these issues, various approaches are used to build PST assessment literacy. Assessment programs are provided as a separate or stand-alone course on enhancing theoretical and practical knowledge. Assessment programs that focus mainly on enhancing theoretical knowledge prioritise on-campus learning over in-school assessment (Bloomfield et al., 2013). Strong support in research and practices in PSTs extending knowledge in their professional experience is still needed in teacher education (Ellis & Loughland, 2017). The amount of assessment knowledge acquired by PSTs cannot fully support their effective implementation of assessment in the classroom (Siegel & Wissehr, 2011).

An integrated curriculum course is an approach that embeds assessment concepts in other methods courses. This approach supports PSTs to learn how to develop integrated units of work, and design appropriate assessment strategies with specific key learning areas. However, these courses could not meet PST needs (Craven et al., 2014) because they provide limited opportunities for PSTs to apply their assessment knowledge and skills in an authentic classroom setting. The topics on assessment require a more focused approach to highlight its central role in learning and teaching rather than just a part of the learning activities (Popham, 2008). Assessment programs are also provided as professional development (PD) or intervention programs that target the specific needs of PSTs. For example, McGee and Colby (2014) explored the impact of an assessment course implemented in the form of PD. Their findings revealed that PSTs increased their assessment literacy in some areas after the program.

## Method

To answer the research questions, we conducted a literature review search and analysis from the initial data search through databases to the study selection for data synthesis according to the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) guidelines (Moher et al., 2009).

### Data sources and literature search

An initial search of the literature was conducted through three databases: *ProQuest*, *Scopus* and *Web of Science*. The time frame for this review was from 1998 to September 2020, starting at the time when Black and Wiliam (1998) published their seminal paper in formative assessment that became the basis for major educational reforms in many educational institutions across the world. The combination of keywords *assessment course*, *assessment program*, *assessment training*, *assessment workshop*, *assessment curriculum*, *assessment pedagog\**, and *teacher* were used to identify the papers in every database. The detailed search strategy syntax used for each database can be seen in Table 1. Articles were included in this review if they were published in peer-reviewed journals in English. There were no restrictions regarding the design of studies: quantitative, qualitative or mixed-methods.

Table 1: Search strategy syntax

Database	Syntax	No. of articles
ProQuest	noft ("Assessment Course" OR "assessment program" OR "assessment training" OR "assessment workshop" OR "assessment curriculum" OR "assessment pedagog*") AND noft(teacher)	472
Web of Science	TOPIC (("Assessment Course" OR "assessment program" OR "assessment training" OR "assessment workshop" OR "assessment curriculum" OR "assessment pedagog*") AND (teacher)) Refined by: DOCUMENT TYPES: (ARTICLE) Timespan: 1988-2020.	210
Scopus	TITLE-ABS-KEY (("Assessment Course" OR "assessment program" OR "assessment training" OR "assessment workshop" OR "assessment curriculum" OR "assessment pedagog*") AND (teacher)) AND PUBYEAR > 1997 AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))	281

### Study selection

The literature search based on the inclusion/exclusion criteria identified a total of 1002 articles (Figure 1). After removing the duplicates, the titles and abstracts of the articles were reviewed if they met the criteria. As the purpose of the review is to help contextualise the key features of the assessment programs for pre-service teachers, the following criteria were used: (1) use of terms “course”, “program”, “training”, “workshop”, “curriculum” or “pedagog\*” in defining their assessment program; (2)

context of pre-service teachers as the nature of the programs are varied between pre-service and in-service teachers; (3) peer-reviewed journal articles; (4) use of English language; and (5) access to full-text. After applying these criteria, 39 articles remained for full-text review.

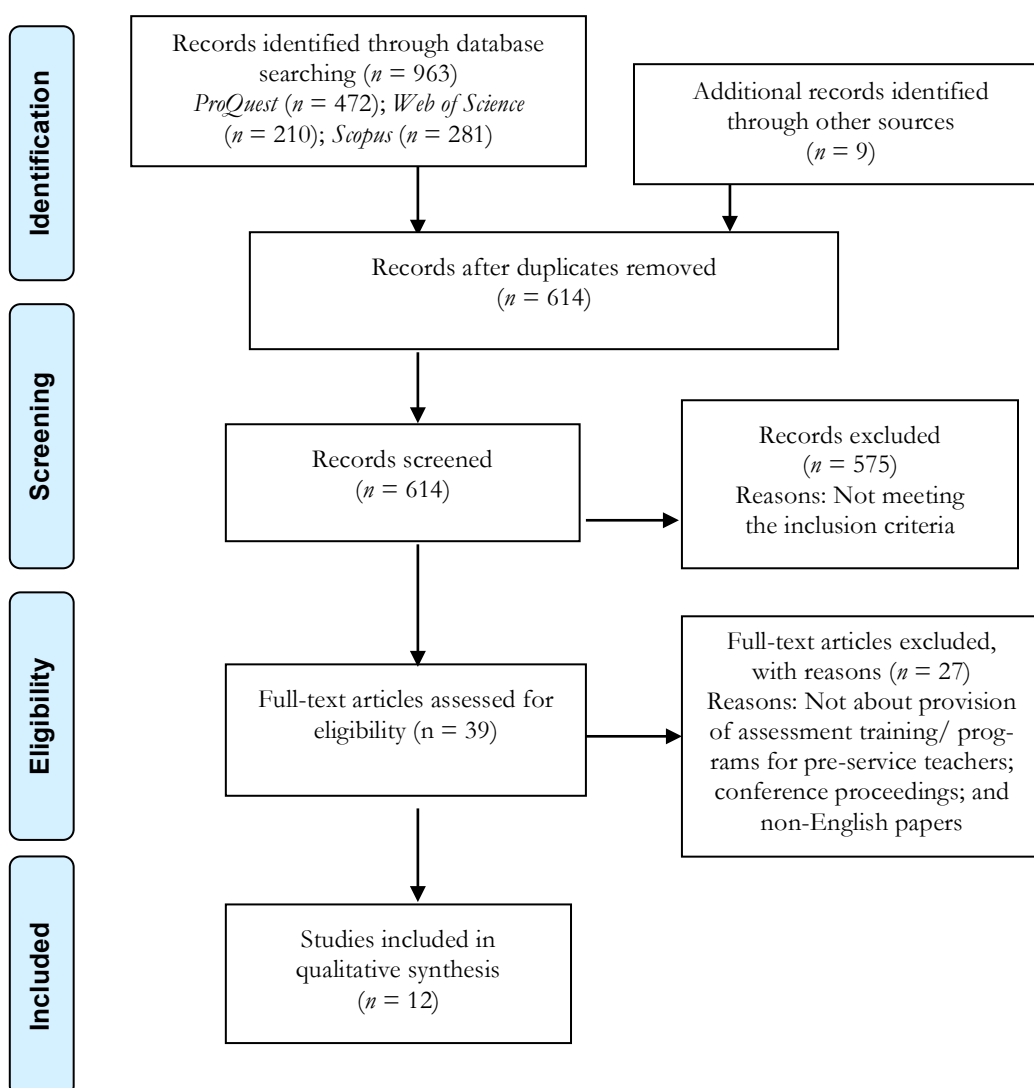


Figure 1: Study selection flow diagram

### Approach to analysis and synthesis

The 39 full-text articles were reviewed if there was clear information about an ITE assessment program. Papers that did not clearly articulate an assessment program for PSTs were excluded from final analysis and synthesis. In addition, articles were excluded if the full text was presented in another language with only their abstracts in English.

Following the same shortlisting and consensus-building process above, 12 articles remained for the purpose of qualitative evidence synthesis.

Three stages of thematic synthesis in systematic reviews as highlighted by Thomas and Harden (2008) were used. These three stages are (i) coding text: the line-by-line coding that was done using *NVivo* software; (ii) developing 'descriptive' themes; and (iii) generating analytical themes. The first author developed descriptive and analytical themes that were reviewed by the second and third authors. Then, these data-driven themes and coding were reviewed again by all authors. The results of the thematic synthesis are presented in the following section.

## Results

This section presents the results of thematic analysis of the literature on assessment programs for PSTs to answer the research questions.

### Characteristics of included articles

This section presents the approaches of assessment programs used in the 12 studies to build PST assessment literacy, the skills/assessment types the program focused upon, and the length of the program (Table 2). Geographically, there were articles from the USA (21%), South America (14%), Canada (21%), Europe (14%), and the Middle East (29%).

Table 2: Characteristics of the included articles

Authors	Country	Content orientation	Approaches	Skills focus /Assessment types	Outcomes measures	Duration
Brevik et al. (2017)	Europe	Theory-driven; Policy-driven; Classroom practice-driven	Integrated curriculum unit	A/L/FA	Changes in assessment skills	One term
Childs & Lawson (2003)	Canada	Audience-oriented; Theory-driven	Intervention/ Workshop	Large scale /Standardised assessment	Changes in perception/conception about assessment; Changes in assessment knowledge	Two months
DeLuca & Klinger (2010)	Canada	Theory-driven; Policy-driven; Classroom practice-driven	Stand-alone course	A/L/FA A/L/SA	Changes in confidence in assessment	One term

Deneen & Brown (2016)	USA	Audience-oriented; Policy-driven; Classroom practice-driven	Stand-alone course	A/L/FA A $\phi$ L/SA Large scale /Standardised assessment	Changes in perception/conception about assessment; Changes in assessment knowledge Changes in assessment skills	12 weeks
Giraldo & Murcia (2019)	Colombia, South America	Audience-oriented; Classroom practice-driven	Intervention/Workshop	A/L/FA A $\phi$ L/SA	Changes in assessment knowledge	16 weeks
Izci & Caliskan (2017)	Middle East	Policy-driven; Classroom practice-driven	Stand-alone course	A/L/FA	Changes in perception/conception about assessment	14 weeks
Levy-Vered & Alhija (2018)	Middle East	Policy-driven; Classroom practice-driven	Stand-alone course	A/L/FA A $\phi$ L/SA	Changes in perception/ conception about assessment	28 hours
McGee & Colby (2014)	USA	Audience-oriented; Classroom practice-driven	Stand-alone course	A/L/FA A $\phi$ L/SA Large scale /Standardised assessment	Changes in assessment knowledge	One assessment course
Ogan-Bekiroglu & Suzuk (2014)	Middle East	Theory-driven; Classroom practice-driven	Stand-alone course	A/L/FA A $\phi$ L/SA	Changes in assessment knowledge; Changes in assessment skills	One assessment course
Reeves & Chiang (2017)	USA	Theory-driven Classroom practice-driven	Intervention/Workshop	Large scale /Standardised assessment	Changes in perception/conception about assessment; Changes in assessment skills	6 hours
Sluijsmans et al. (2002)	Europe	Audience-oriented; Theory-driven; Classroom practice-driven	Intervention/Workshop	A/L/FA	Changes in perception/conception about assessment; Changes in assessment skills	4 hours
Yurdabakan (2012)	Middle East	Audience-oriented; Classroom practice-driven	Intervention/Workshop	A/L/FA A $\phi$ L/SA	Changes in assessment skills	8 weeks

\*A/L = Assessment *for* learning; A $\phi$ L = Assessment *of* learning; FA = Formative assessment; SA = Summative assessment



## **Content orientation**

Four key content designs in assessment programs emerged from the 12 articles: audience-oriented; theory-driven; policy-driven; and classroom practice-driven in designing assessment programs.

### **Audience-oriented assessment program**

Audience-oriented is defined as developing an assessment program based on the needs of a specific cohort. The programs are prepared to ensure the relevance and usefulness for targeted PSTs and to reflect the assessment system of a particular context. Two sub-themes of this key feature emerged: (i) needs of the target audience, and (ii) refinement of the program.

Thirty-three per cent of the articles developed the assessment programs to address PST assessment knowledge and skills. For example, an assessment workshop was provided in the study by Childs and Lawson (2003) as the PSTs requested it. This workshop was an addition to the prescribed assessment course to prepare classroom-ready teachers in Ontario. Giraldo and Murcia (2019) designed their classroom language assessment course grounded in the findings of the previous diagnostic study to meet the needs of participants and also professors.

Few authors mentioned the refinement of the programs for PSTs grounded in the perspectives of course instructors and stakeholders. For example, Sluijsmans et al. (2002) embedded peer assessment training in their second-year course "Designing creative lessons" which was a modification of the existing course. They considered the teachers' perspectives (course instructors) in redefining the course objectives and the tasks in their course. Similarly, Giraldo & Murcia (2019) modified the program's content derived from the diagnostic stage of their study.

### **Theory-driven assessment program**

Of the reported articles, 50% highlighted that their assessment programs were underpinned by a theoretical framework or teacher standards. The most common theoretical frameworks used are the Vygotskian sociocultural approach (Brevik et al., 2017), Tittle's framework dimensions of assessment (Ogan-Bekiroglu & Suzuk, 2014), Mandinach and Gummer's framework for teacher assessment literacy (Reeves & Chiang, 2017), and the standards for teacher competence in educational assessment of students (Childs & Lawson, 2003).

The Vygotskian sociocultural approach was used in framing assessment courses based on the ongoing relationship between PSTs, university, and real school settings (Brevik et al., 2017). The role of interactions in this approach is the main feature of the learning activities. Brevik et al. commented that "the Vygotskian framing therefore provides a perspective on the student teachers' use of assessment principles to develop their principles" (p. 168). In the study by Ogan-Bekiroglu and Suzuk (2014), their assessment

program was based on the second dimension of Tittle (1994), the knowledge, beliefs, intents, and actions of the assessment of both interpreter and user. As their study focused on improving PST assessment literacy and their implementation into practice, they viewed their program from a constructivist epistemology.

The study by Reeves and Chiang (2017) used five knowledge/skill domains of assessment data literacy proposed by Mandinach and Gummer (2016) to target “PSTs capacity to transform data into information (e.g., interpret data, understand data representations) and transform information into a decision (e.g., specify next instructional steps)” (p. 157). Although DeLuca and Klinger (2010) noted that the approaches to assessment programs can be varied depending on the needs of PSTs, they prepared assessment programs reflecting the Ontario College’s *Standards of Practice for the Teaching Profession*, that is, PSTs are expected to “use appropriate pedagogy, assessment and evaluation” to meet the needs of students and learning communities (p. 421). In the study by Sluijsmans et al. (2002), they prepared the program grounded in a peer assessment model including three main peer assessment skills.

### **Policy-driven assessment program**

Almost half of the included articles (42%) designed assessment programs to enact policy. For example, in the study by Izci and Caliskan (2017), the concept of assessment *for* learning was integrated into their teacher education program as mandated by Turkey’s Ministry of National Education. Similarly, the assessment program in Levy-Vered and Alhija (2018) was developed in response to the recommendation of the Ministry of Education in Israel to promote the use of alternative assessments. Brevik et al. (2017) prepared an assessment program based on the demands of the Norwegian Education Act. This policy-driven approach considers the course as evolving within the context of policy directives that shape teacher assessment practices (Alonzo, Labad, Bejano & Guerra, 2021).

### **Classroom practice-driven assessment program**

There are assessment programs that were based on school assessment practices. Three themes emerged from the articles: (1) assessment tasks and assignments; (2) students’ actual output; and (3) assessment practices in a real classroom.

Of the included articles, 83% highlighted that assessment tasks and assignments were included in their assessment programs. The assessment strategies needed in schools influenced the assessment design. Some of the programs include developing *an assessment scenario; an assessment task and a scoring rubric* (Izci & Caliskan, 2017); *an assessment plan* (Ogan-Bekiroglu & Suzuk, 2014); *a data use schema* (Reeves & Chiang, 2017); *a content-based assessment and a performance assessment* (McGee & Colby, 2014); and *multiple-choice assessments and rubrics for language assessments* (Giraldo & Murcia, 2019). In addition, some programs included analysing actual classroom assessment tasks and practices as part of the programs. For example, the assessment course mentioned used by Levy-Vered and Alhija (2018) asked PSTs to critically analyse a teacher-made achievement test before they

designed their own assessment. In the study of Brevik et al. (2017), the tasks for PSTs are to observe and analyse the formative assessment practices in a video-based classroom.

Assessing actual students' work is used in assessment programs. Twenty-five percent of the included articles highlighted that real students' work was combined in the program (Brevik et al., 2017; Izci & Caliskan, 2017; McGee & Colby, 2014). This addresses the issue of extended classroom practice when opportunities to include a practicum experience for PSTs are limited. The process involved analysing students' responses if they met the learning outcomes and then "offer pedagogical ways for how to help low-level students to develop the concepts" (Izci & Caliskan, 2017, p. 467). In addition, PSTs participated in further instruction for students based on the analysis of student performance (McGee & Colby, 2014). In the study by Brevik et al. (2017), PSTs assessed students' written text and/or their video recorded presentation, drew up criteria to assess, and gave feedback on students' work at the end.

Assessment practices in classrooms is the authentic practical sides of assessment in programs. Only 14% of the included articles described that assessment practices in actual classrooms were embedded in their assessment programs. The practicum experience is crucial for improving PST assessment literacy. DeLuca and Klinger (2010) highlighted that PSTs gained confidence in using assessment from their practicum experience. Likewise, Ogan-Bekiroglu and Suzuk (2014) provided opportunities to PSTs for their practical application of knowledge and skills. Consequently, PSTs can implement feedback from their instructors (Griffiths et al., 2021) to improve their assessment practices.

## **Approaches in implementing ITE assessment programs**

A closer look at the assessment approaches used to build PST assessment literacy reveals that there are six articles about a stand-alone course, one article reported an integrated curriculum unit and six articles used intervention/workshop (Table 2). All assessment approaches are designed to cover assessment concepts that are deemed important by the course developers.

The stand-alone courses (DeLuca & Klinger, 2010; Deneen & Brown, 2016; Izci & Caliskan, 2017; Levy-Vered & Alhija, 2018; McGee & Colby, 2014; Ogan-Bekiroglu & Suzuk, 2014) were provided as either elective or mandatory courses in an ITE curriculum. An elective course raises two important concerns: it is not deemed as an important course that every PST should take, and consequently, not all PSTs who graduate have the same level of assessment knowledge and skills.

Integrated curriculum courses (Brevik et al., 2017) embed assessment concepts to support PSTs to design appropriate assessment strategies for specific key learning areas. This approach is consistent with the nature of assessment which is context-driven and the design should be carefully planned within the context of content and pedagogical knowledge (William, 2013). However, it raises the issue of the sufficiency of time allocated for this course to cover important assessment knowledge and skills.

The third approach reported is the use of intervention programs or workshops enabling PSTs to enhance their existing assessment knowledge and skills (Childs & Lawson, 2003; Giraldo & Murcia, 2019; Reeves & Chiang, 2017; Sluijsmans et al., 2002; Yurdabakan, 2012). This approach is often used when assessment courses are not an integral part of the ITE curriculum.

### **Skills focused/assessment types**

In terms of specific knowledge and skills-focus of the programs, the papers reported assessment *for* learning (A/L), formative assessment (FA), assessment *of* learning (A/L), summative assessment (SA), and large scale or standardised assessment. Most assessment programs covered more than one type of assessment (Table 2). Seventy-eight percent of the articles emphasised A/L and/or FA, while 14% included only large scale/standardised assessment, provided as an additional program with the existing ones. For example, the study by Childs and Lawson (2003) provided large scale assessment as another workshop, because was requested by course instructors.

The varied content-focus of the programs is due largely to the competing conceptualisations of effective assessment practices (Alonzo, 2016) and policy requirements (Davison, 2013). It is evident that the construct of teacher assessment literacy needs to be clearly articulated.

### **Measures of outcomes reported**

There are various methods used to measure the impact of assessment programs including quantitative methods (questionnaire or survey, assessment instrument, checklists), qualitative methods (interviews, open-ended questionnaire, peer-assessment tasks), and mixed methods (questionnaire, interviews, project assignment, lecture notes, lesson plans, and class observation). Four key measures are used to report the impact of the program. These include:

1. Changes in perception/conception of assessment
2. Changes in confidence in assessment
3. Changes in assessment knowledge
4. Changes in assessment skills

Fifty percent of the studies reported changes in PST perception/conception of assessment after the program. Giraldo and Murcia (2019) used pre- and post-design to measure PST changes in perception, and found that PST perception towards designing assessment tasks changed radically. Perceptions evolved “from limited views to an intricate and professionalising process-oriented endeavour” (p. 254). Similarly, PST perceptions in the studies by Izci and Caliskan (2017) and Levy-Vered and Alhija (2018) changed towards using assessment tasks that can improve student learning, away from assessment tasks focused on memorisation. PSTs in the study by Sluijsmans et al. (2002) described positive changes in their perception of assessment and instruction as evidenced in the results of pre- and post-test.

One study used the change in confidence in assessment to report the impact of the program. DeLuca & Klinger (2010) reported the questionnaire results of PST confidence levels from experimental and control groups that PSTs who participated in a formal assessment program were significantly more confident than those who did not. They also commented that their assessment program had a greater impact on their confidence level regarding assessment theory and practice.

Another measure is the changes in PST assessment knowledge. Thirty-three percent of the articles highlighted that assessment programs have a positive impact on PST assessment knowledge. The pre- and post-test results in the study by Deneen and Brown (2016) and McGee and Colby (2014) show an increase on PST assessment knowledge. McGee and Colby (2014) described a positive significant change in all subscales: choosing methods; sound design; scoring; using results; grading; communicating results, and ethical assessment.

Fifty percent of the included articles used improvement in assessment skills. PSTs changed not only their assessment knowledge but also their assessment skills at the end of the program (Deneen & Brown, 2016; Ogan-Bekiroglu & Suzuk, 2014). This is evident in the work of Deneen and Brown (2016) that PSTs have developed well-aligned and well-designed assessment tasks as a result of engagement in the program. Sluijsmans et al. (2002) and Yurdabakan (2012) used experimental design to measure the impact of the program and both studies have documented an improvement on PST assessment skills.

### **Duration of the program**

Regarding duration of the assessment programs, the articles described the length of the program as hours, weeks, months, term or semester. In describing the length of the programs, the range is from 4 hours to 28 hours, from 8 weeks to 16 weeks (4 months), and one term or course. In some interventions/workshops, the course duration is shorter than the stand-alone/separate course (e.g., Reeves & Chiang, 2017; Sluijsmans et al., 2002). The different duration of the programs raises a critical issue on how much time is needed to build assessment literacy.

### **Limitations of the extant literature**

Taken as a body of research, we have identified a few areas that have not been reported in the 12 studies. The results of our review highlight that there is a little reported research on improving assessment literacy for PSTs, which may explain why graduate teachers have low assessment literacy (Maclellan, 2004; Oo et al., 2021; Volante & Fazio, 2007) and feel inadequate to take on their role (Hill & Evers, 2016; Ogan-Bekiroglu & Suzuk, 2014).

There was no paper reporting the different factors that influence PST assessment literacy development. There is evidence that teachers' assessment literacy development is influenced by various factors including school culture (Alonzo, Leverett, et al., 2021; Charteris & Dargusch, 2018; Evers, 2014), non-cognitive skills (Oo, 2020), policies (Alonzo, Labad, et al., 2021) and sociocultural contexts (Willis & Klenowski, 2018). There

is an emerging evidence that PST assessment literacy development is not only dependent on their beliefs and values in using assessment but largely influenced by their supervising teachers during practicum (Oo et al., 2021). This is an important area to explore to ensure that enabling factors are leveraged during the program's implementation while managing limiting factors.

In addition, the definition of assessment literacy is limited to knowledge and skills, but it has been proven that teaching skills require a much broader definition of knowledge including confidence (Beswick et al., 2012) and beliefs (Ball et al., 2008). This broader conceptualisation of assessment literacy would ensure that PSTs will better understand what constitutes an assessment literate teacher.

Methodologically, no paper reported the impact of the program using effect size. We need a more rigorous empirical methodology to provide evidence of the effectiveness of the programs. No paper reports what “works” as an effective program but what have been reported were the outcomes of research on the design and content imposed by the researchers. Moreover, there is no critical evaluation of what the best measures for reporting on the effectiveness of the program.

Furthermore, there no longitudinal study has ascertained whether PST assessment literacy gained from ITE has been applied in their teaching. The consequential validity (Slomp et al., 2014) of the programs is critically important to ensure that the knowledge and skills gained by PSTs are applied in schools.

## Discussion

Building upon the findings of this literature review, this paper proposes a framework for designing and implementing an assessment program (Figure 2). This framework consists of six key features that are critical components for designing the assessment program.

The framework will guide curriculum developers by addressing the questions provided. The first key question we need to consider when developing a PST assessment program is, “*What would be the basis for the program?*”. What is reported in the literature is either audience-oriented, theory-driven, policy-driven, or classroom practice-driven. We also need to consider each aspect of these findings. For example: “*What are the assessment literacy needs of PSTs?*” and “*What assessment practices are articulated in the assessment policy?*”. The orientation of the program could be a combination of addressing the assessment policy requirements, adhering to a strong theorisation of effective assessment practices, and addressing the assessment knowledge and skills needs of PSTs. As assessment is a context-dependent construct, no single content orientation is effective.

Regarding the content focus of the assessment program, three types of assessment are found in this review: FA/AJL, SA/AoL, and large scale or standardised assessment. Curriculum developers need to consider, “*What assessment knowledge and skills are needed in their own context?*”. Formative assessment and AJL is the most common assessment type

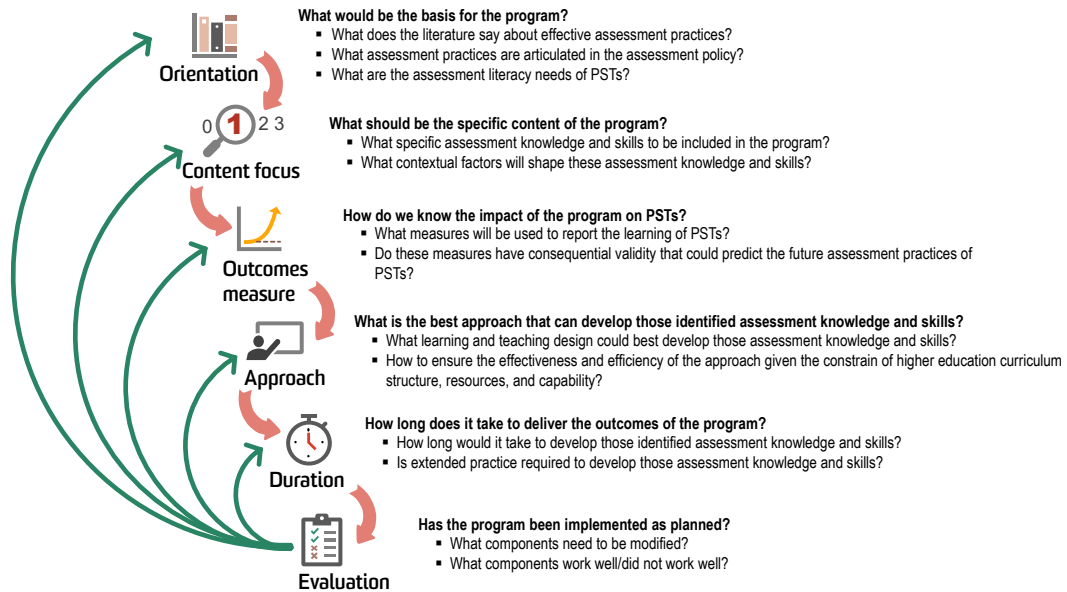


Figure 2: A framework for designing and implementing the assessment program in ITE  
(Use web or PDF reader 'zoom in' function to view details)

found in this literature. This finding is consistent with the literature that highlights A/L in teacher preparation is the highest potential of increasing student outcomes (Black & Wiliam, 1998b; Hattie, 2009). However, with the current conceptualisation of assessment, where the distinction between FA and SA is becoming irrelevant (Black & Wiliam, 2018) because both assessment types, including high-stakes tests, can be used to support student learning and address accountability requirements, a broader content focus is recommended. Therefore, we also need to consider “*What contextual factors will shape these assessment knowledge and skills?*”.

In terms of the outcome measures of the program, this review found four outcome measures: changes in perception/conception of assessment; changes in confidence in assessment; changes in assessment knowledge; and changes in assessment skills. This variability in outcomes reported raises the questions, “*How do we know the impact of the program on PSTs?*”, particularly “*What measures will be used to report the learning of PSTs?*” and “*Do these measures have consequential validity that could predict the future assessment practices of PSTs?*”. Change in assessment practices is considered as one of the critical outcome measures to prepare classroom-ready teachers (BOSTES, 2016; Volante & Fazio, 2007). However, as assessment literacy is a broad construct involving not only knowledge and skills-based but also non-cognitive skills, the measures should be expanded to account for this broader conceptualisation of assessment literacy.

Having identified the content focus and measures to be used, curriculum developers need to identify the best approach to develop that knowledge and skills. Our literature review shows three approaches, separate/stand-alone course, integrated curriculum unit, and

intervention/workshop. Among these approaches, “*What learning and teaching design could best develop those assessment knowledge and skills?*”. There is no consensus as to what is the best approach as each approach used in various studies reported its effectiveness. However, owing to the centrality of preparing PSTs to become classroom-ready teachers, curriculum developers need to consider their theoretical knowledge acquisition and practical skills development while building their confidence to design and implement various assessment strategies.

*How long does it take to deliver the program?* The duration of the program is largely dependent on the outcomes aimed for, the structural context of ITE program and the regulatory requirements of accrediting body. In developing those identified assessment knowledge and skills, our review shows that classroom practices, including assessment tasks and assignments, students’ actual output and assessment practices in real classroom, were combined in the program. Therefore, we also need to consider “*Is extended practice required to develop those assessment knowledge and skills?*”

One of the missing components of programs reported in the literature is conducting an evaluation to answer the question: *Has the program been implemented as planned?* For every implementation of any assessment program, a clear evaluation must be conducted to identify the aspects of the program that work well, and those aspects that need further improvement. This is an iterative process that multiple sources of data will be used to provide feedback for every aspect of the framework to ensure continuous improvement.

## Conclusion

This paper highlights the characteristics of ITE assessment programs in terms of their orientation, content focus, outcomes measure, approaches and duration. Despite the peculiarities of each program, all studies reported their effectiveness in improving a specific area of PST assessment literacy. From our scoping review, it is evident that this body of knowledge is limited and hence more studies are needed to provide a richer conceptualisation of ITE assessment programs.

As an output of this review, we develop a framework for designing and implementing an ITE assessment program. This framework should be taken as a guide rather than a rigid structure to follow. There should be strong regard to the context of educational system in which the assessment program would be implemented.

This paper has its limitations, like most research. In terms of searching papers, we only collected publications from *ProQuest*, *Web of Science*, and *Scopus*. We considered these publishers as they offer strong indexing quality, written in English that made their journals readable for international audiences. However, this choice limited our scoping where we did not review printed books and documents from the government and international agencies about teacher assessment literacy. This may be an area that can be considered in future systematic literature reviews. In addition, we did not discuss the relationship between teacher assessment literacy and subject knowledge (e.g., STEM, languages,



humanities, and social sciences) as we put more focus on the general ITE assessment programs. Further research that investigates the link between teacher assessment literacy and subject knowledge may be carried out as the content focus of ITE assessment programs depends on the subject focus. Lastly, building pre-service teachers' knowledge and skills in high-stake assessment, particularly the national and international examinations, warrants further investigations. These high-stake assessment and other broader assessment terminologies can also be used in the future reviews.

## Acknowledgment

We would like to thank Dr Carlito Tabelin of the University of New South Wales Sydney for helping us with Figure 1.

## References

- Ahmad, I. F. (2020). Alternative assessment in distance learning in emergencies spread of coronavirus disease (COVID-19) in Indonesia. *Jurnal Pedagogik*, 7(1), article 1136. <http://doi.org/10.33650/pjp.v7i1.1136>
- Alonzo, D. (2016). *Development and application of a teacher assessment for learning (AfL) literacy tool*. PhD thesis, The University of New South Wales. <http://unsworks.unsw.edu.au/fapi/datastream/unsworks:38345/SOURCE02?view=true>
- Alonzo, D., Labad, V., Bejano, J. & Guerra, F. (2021). The policy-driven dimensions of teacher beliefs about assessment. *Australian Journal of Teacher Education*, 46(3), article 3. <https://doi.org/10.14221/ajte.2021v46n3.3>
- Alonzo, D., Leverett, J. & Obsioma, E. (2021). Leading an assessment reform: Ensuring a whole-school approach for decision-making. *Frontiers in Education*, 6, article 631857. <https://doi.org/10.3389/educ.2021.631857>
- Assessment Reform Group (1999). *Assessment for learning: Beyond the black box*. Nuffield Foundation and University of Cambridge. [https://www.nuffieldfoundation.org/sites/default/files/files/beyond\\_blackbox.pdf](https://www.nuffieldfoundation.org/sites/default/files/files/beyond_blackbox.pdf)
- Beswick, K., Callingham, R. & Watson, J. (2012). The nature and development of middle school mathematics teachers' knowledge. *Journal of Mathematics Teacher Education*, 15(2), 131-157. <https://doi.org/10.1007/s10857-011-9177-9>
- Black, P. & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148. <https://doi.org/10.1177/003172171009200119>
- Black, P. & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551-575. <https://doi.org/10.1080/0969594X.2018.1441807>
- BOSTES (2016). *Learning assessment: A report on teaching assessment in initial teacher education in NSW*. NSW Government.

- <https://educationstandards.nsw.edu.au/wps/wcm/connect/c204171e-a570-4947-8107-dc934ab2f70b/learning-assessment-report.pdf?MOD=AJPERES&CVID>
- Brevik, L. M., Blikstad-Balas, M. & Engelién, K. L. (2017). Integrating assessment for learning in the teacher education programme at the University of Oslo. *Assessment in Education: Principles, Policy & Practice*, 24(2), 164-184. <https://doi.org/10.1080/0969594X.2016.1239611>
- Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice*, 18(1), 5-13. <https://doi.org/10.1111/j.1745-3992.1999.tb00002.x>
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Charteris, J. & Dargusch, J. (2018). The tensions of preparing pre-service teachers to be assessment capable and profession-ready. *Asia-Pacific Journal of Teacher Education*, 46(4), 354-368. <https://doi.org/10.1080/1359866X.2018.1469114>
- Childs, R. A. & Lawson, A. (2003). What do teacher candidates know about large-scale assessments? What should they know? *Alberta Journal of Educational Research*, 49(4), 354-367. <https://doi.org/10.11575/ajer.v49i4.55029>
- Craven, G., Beswick, K., Fleming, J., Fletcher, T., Green, M., Jensen, B., Leinonen, E. & Rickards, F. (2014). *Action now: Classroom ready teachers*. Department of Education. <https://www.aitsl.edu.au/tools-resources/resource/action-now-classroom-ready-teachers>
- Davison, C. (2007). Views from the chalkface: English language school-based assessment in Hong Kong. *Language Assessment Quarterly*, 4(1), 37-68. <https://doi.org/10.1080/15434300701348359>
- Davison, C. (2013). Innovation in assessment: Common misconceptions and problems. In K. Hyland & L. L. C. Wong (Eds.), *Innovation and change in English language education* (pp. 263-275). Oxon: Routledge. <https://www.routledge.com/Innovation-and-change-in-English-language-education/Hyland-Wong/p/book/9780415826877>
- DeLuca, C., Chavez, T., Bellara, A. & Cao, C. (2013). Pedagogies for pre-service assessment education: Supporting teacher candidates' assessment literacy development. *Teacher Educator*, 48(2), 128-142. <https://doi.org/10.1080/08878730.2012.760024>
- 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>
- 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). <https://doi.org/10.1080/2331186X.2016.1225380>
- Ellis, N. J. & Loughland, T. (2017). 'Where to next?' Examining feedback received by teacher education students. *Issues in Educational Research*, 27(1), 51-63. <http://www.iier.org.au/iier27/ellis.pdf>
- Eyers, G. (2014). *Pre-service teachers' assessment learning: Change, development and growth*. PhD thesis, The University of Auckland. <http://hdl.handle.net/2292/23757>
- Gipps, C. (1995). Teacher assessment and teacher development in primary schools. *Education 3-13*, 23(1), 8-12. <https://doi.org/10.1080/03004279585200021>

- Giraldo, F. & Murcia, D. (2019). Language assessment literacy and the professional development of pre-service language teachers. *Colombian Applied Linguistics Journal*, 21(2), 243-259. <https://doi.org/10.14483/22487085.14514>
- Grainger, P. R. & Adie, L. (2014). How do pre-service teacher education students move from novice to expert assessors? *Australian Journal of Teacher Education*, 39(7), article 9. <https://doi.org/10.14221/ajte.2014v39n7.9>
- Greenberg, J. & Walsh, K. (2012). *What teacher preparation programs teach about K - 12 assessment: A review*. <https://files.eric.ed.gov/fulltext/ED532766.pdf>
- Griffiths, M., Shean, M. & Jackson, D. (2021). Supervision in initial teacher education: A scoping review. *Issues in Educational Research*, 31(2), 476-494. <http://www.iier.org.au/iier31/griffiths-m.pdf>
- Hailaya, W. M. (2014). *Teacher assessment literacy and student outcomes in the Province of Tawi-Tawi, Philippines*. PhD thesis, University of Adelaide, Australia. <https://hdl.handle.net/2440/99098>
- Hardman, F., Stoff, C., Aung, W. & Elliott, L. (2016). Developing pedagogical practices in Myanmar primary schools: Possibilities and constraints. *Asia Pacific Journal of Education*, 36(sup1), 98-118. <https://doi.org/10.1080/02188791.2014.906387>
- Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, 16(2), 207-223. <https://doi.org/10.1080/09585170500136093>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://www.routledge.com/Visible-Learning-A-Synthesis-of-Over-800-Meta-Analyses-Relating-to-Achievement/Hattie/p/book/9780415476188>
- Hill, M. F., Ell, F., Grudnoff, L. & Limbrick, L. (2014). Practise what you preach: Initial teacher education students learning about assessment. *Assessment Matters*, 7, 90-112. <https://doi.org/10.18296/am.0126>
- Hill, M. F. & Evers, G. (2016). Moving from student to teacher. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment*. Routledge. <https://www.routledge.com/Handbook-of-Human-and-Social-Conditions-in-Assessment/Brown-Harris/p/book/9781138811553#>
- Izci, K. & Caliskan, G. (2017). Development of prospective teachers' conceptions of assessment and choices of assessment tasks. *International Journal of Research in Education and Science*, 3(2), 464-474. <https://www.ijres.net/index.php/ijres/article/view/241>
- James, M. & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138. <https://doi.org/10.1080/09585170600792712>
- Kennedy, K. J., Chan, J. K.-S., Yu, F. W. & Fok, P.-K. (2007). Assessment for productive learning: Forms of assessment and their potential for enhancing learning. Presented at *Student assessment and its social and cultural contexts: How teachers respond to assessment reform*. Redesigning Pedagogy - Culture, Understanding and Practice Conference. Singapore, 28-30 May. [https://www.eduhk.hk/fpece\\_project/QEF/Download area/1\\_Assess for productive learning \(Kerry paper\).pdf](https://www.eduhk.hk/fpece_project/QEF/Download%20area/1_Assess%20for%20productive%20learning%20(Kerry%20paper).pdf)
- Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197. <https://doi.org/10.1177/0265532214554321>

- Levy-Vered, A. & Alhija, F. N. A. (2018). The power of a basic assessment course in changing pre-service teachers' conceptions of assessment. *Studies in Educational Evaluation*, 59, 84-93. <https://doi.org/10.1016/j.stueduc.2018.04.003>
- Loughland, T. & Alonzo, D. (2018). Teacher adaptive practices: Examining links with teacher self-efficacy, perceived autonomy support and teachers' sense of adaptability. *Educational Practice and Theory*, 40(2), 55-70. <https://doi.org/10.7459/ept/40.2.04>
- MacLellan, E. (2004). Initial knowledge states about assessment: Novice teachers' conceptualisations. *Teaching and Teacher Education*, 20(5), 523-535. <https://doi.org/10.1016/j.tate.2004.04.008>
- Mandinach, E. B. & Gummer, E. S. (2016). *Data literacy for educators: Making it count in teacher preparation and practice*. Teachers College Press. <https://www.tcpress.com/data-literacy-for-educators-9780807757536>
- McGee, J. & Colby, S. (2014). Impact of an assessment course on teacher candidates' assessment literacy. *Action in Teacher Education*, 36(5-6), 522-532. <https://doi.org/10.1080/01626620.2014.977753>
- Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4). <https://doi.org/10.1136/bmj.b2535>
- Ogan-Bekiroglu, F. & Suzuk, E. (2014). Pre-service teachers' assessment literacy and its implementation into practice. *The Curriculum Journal*, 25(3), 344-371. <https://doi.org/10.1080/09585176.2014.899916>
- Oo, C. Z. (2020). *Assessment for learning literacy and pre-service teacher education: Perspectives from Myanmar*. PhD thesis, The University of New South Wales, Australia. <http://handle.unsw.edu.au/1959.4/65021>
- Oo, C. Z., Alonzo, D. & Davison, C. (2021). Pre-service teachers' decision-making and classroom assessment practices. *Frontiers in Education*, 6, article 628100. <https://doi.org/10.3389/educ.2021.628100>
- 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. (2013). What assessment knowledge and skills do initial teacher education programs address? A Western Canadian perspective. *Alberta Journal of Educational Research*, 58(4), 634-656. <https://journalhosting.ucalgary.ca/index.php/ajer/article/view/55670/42588>
- Reeves, T. D. & Chiang, J. L. (2017). Building pre-service teacher capacity to use external assessment data: An intervention study. *The Teacher Educator*, 52(2), 155-172. <https://doi.org/10.1080/08878730.2016.1273420>
- Rwehera, M. (2017). *Myanmar National Education Strategic Plan 2016/17-2020/21 (NESP): Appraisal report*. <https://www.globalpartnership.org/sites/default/files/2018-05-gpe-esp-appraisal-report-myanmar.pdf>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144. <https://link.springer.com/content/pdf/10.1007/BF00117714.pdf>
- 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>

- Slomp, D. H., Corrigan, J. A. & Sugimoto, T. (2014). A framework for using consequential validity evidence in evaluating large-scale writing assessments: A Canadian study. *Research in the Teaching of English*, 48(3), 276-302.  
<http://www.jstor.org/stable/24398680>
- Sluijsmans, D. M. A., Brand-Gruwel, S. & van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects on performance and perceptions. *Assessment and Evaluation in Higher Education*, 27(5), 443-454.  
<https://doi.org/10.1080/0260293022000009311>
- Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of Further and Higher Education*, 33(1), 57-69.  
<https://doi.org/10.1080/03098770802638671>
- Thomas, J. & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, article 45.  
<https://doi.org/10.1186/1471-2288-8-45>
- Title, C. K. (1994). Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments. *Educational Psychologist*, 29(3), 149-162. [https://doi.org/10.1207/s15326985ep2903\\_4](https://doi.org/10.1207/s15326985ep2903_4)
- Tsagari, D. & Vogt, K. (2017). Assessment literacy of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6(1), 41-63.  
[https://arts.unimelb.edu.au/\\_\\_data/assets/pdf\\_file/0010/2349928/6\\_1\\_SI3TsagariVogt.pdf](https://arts.unimelb.edu.au/__data/assets/pdf_file/0010/2349928/6_1_SI3TsagariVogt.pdf)
- Verberg, C. P. M., Tigelaar, D. E. H., van Veen, K. & Verloop, N. (2016). Teacher agency within the context of formative teacher assessment: An in-depth analysis. *Educational Studies*, 42(5), 534-552. <https://doi.org/10.1080/03055698.2016.1231060>
- Volante, L. & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, 30(3), 749-770. <https://eric.ed.gov/?id=EJ780818>
- Walker, R., Oliver, R. & Mackenzie, R. (2021). Group differences in secondary school students' perceptions of feedback. *Issues in Educational Research*, 31(1), 309-326.  
<http://www.iier.org.au/iier31/walker.pdf>
- Wiliam, D. (2013). Assessment: The bridge between teaching and learning. *Voices from the Middle*, 21(2), 15-20. <https://library.ncte.org/journals/vm/issues/v21-2/24461>
- Willis, J. & Klenowski, V. (2018). Classroom assessment practices and teacher learning: An Australian perspective. In H. Jiang & M. F. Hill (Eds.), *Teacher learning with classroom assessment: Perspectives from Asia Pacific*. pp.19-37. Springer Singapore.  
<http://link.springer.com/10.1007/978-981-10-9053-0>
- Wurf, G. & Povey, R. (2020). "They show how smart you are": A mixed methods study of primary students' perceptions of assessment tasks. *Issues in Educational Research*, 30(3), 1162-1182. <http://www.iier.org.au/iier30/wurf.pdf>
- Xu, Y. & Brown, G. T. L. (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>

- Yurdabakan, I. (2012). The effect of co- and peer assessment training on self-assessment skills of teacher trainees. *Education and Science*, 37(163), 190-202.  
[https://www.researchgate.net/publication/232733893\\_The\\_Effect\\_of\\_Co-\\_and\\_Peer\\_Assessment\\_Training\\_on\\_Self-Assessment\\_Skills\\_of\\_Teacher\\_Trainees\\_Ortak\\_ve\\_Akran\\_Degerlendirme\\_Egitimin\\_in\\_Ogretmen\\_Adaylarinin\\_Oz-degerlendirme\\_Becerileri\\_Uzerine\\_Etkisi](https://www.researchgate.net/publication/232733893_The_Effect_of_Co-_and_Peer_Assessment_Training_on_Self-Assessment_Skills_of_Teacher_Trainees_Ortak_ve_Akran_Degerlendirme_Egitimin_in_Ogretmen_Adaylarinin_Oz-degerlendirme_Becerileri_Uzerine_Etkisi)
- Zulaiha, S., Mulyono, H. & Ambarsari, L. (2020). An investigation into EFL teachers' assessment literacy: Indonesian teachers' perceptions and classroom practice. *European Journal of Contemporary Education*, 9(1).  
<https://doi.org/10.13187/ejced.2020.1.189>

**Dr Cherry Zin Oo** (corresponding author) is currently a lecturer in the Department of Educational Psychology at Yangon University of Education, Myanmar. She is also a recent PhD graduate from the School of Education, University of New South Wales (UNSW), Sydney, Australia.  
ORCID: <https://orcid.org/0000-0002-3227-8010>  
Email: [cherryzinn@gmail.com](mailto:cherryzinn@gmail.com)

**Dr Dennis Alonzo** is a lecturer at the School of Education, Faculty of Arts, Design and Architecture, University of New South Wales (UNSW), Sydney, Australia.  
ORCID: <https://orcid.org/0000-0001-8900-497X>  
Email: [d.alonzo@unsw.edu.au](mailto:d.alonzo@unsw.edu.au)

**Dr Ria A. Asih** is a lecturer at the Directorate of Postgraduate Program, Universitas Muhammadiyah Malang in Indonesia. She is a recent graduate from the School of Education, University of New South Wales (UNSW), Sydney, Australia.  
Email: [aristaria86@umm.ac.id](mailto:aristaria86@umm.ac.id)

**Please cite as:** Oo, C. Z., Alonzo, D. & Asih, R. A. (2022). Acquisition of teacher assessment literacy by pre-service teachers: A review of practices and program designs. *Issues in Educational Research*, 32(1), 352-373. <http://www.iier.org.au/iier32/oo.pdf>