



Using a Needs-Based Professional Development Program to Enhance Pre-Service Teacher Assessment for Learning Literacy

Cherry Zin Oo

Yangon University of Education, Yangon, Myanmar, cherryzino@yuoe.edu.mm

Dennis Alonzo

University of New South Wales, Sydney, Australia, d.alonzo@unsw.edu.au

Chris Davison

University of New South Wales, Sydney, Australia, c.davison@unsw.edu.au

Across the globe, pre-service teachers' (PSTs) assessment for learning literacy remains relatively low despite the critical role it plays in increasing student outcomes. Research highlight that this is due to insufficient initial teacher preparation with most teacher education curricula are using a less effective approach. To address this gap, this study explores the utilisation of a needs-based professional development (PD) program to build PSTs' assessment for learning (AfL) literacy. Using design-based research methodology, a needs-based AfL literacy program was developed at one of the teacher training institutes in Myanmar. A total of 335 PSTs involving 30 PSTs in the intervention group and 305 PSTs in the cohort group were asked to self-assess their AfL literacy before and after the program by using survey instrument. Findings revealed that there was an effect of the PD program on PSTs' AfL literacy that PSTs increased their AfL literacy after the PD program. Implications of the results for the development and implementation of the needs-based PD program are discussed.

Keywords: assessment for learning, pre-service teacher, assessment literacy, professional development, needs-based

INTRODUCTION

This study highlights the use of design-based research (DBR) as an approach to increase pre-service teachers' (PSTs) assessment for learning (AfL) knowledge and skills, known as AfL literacy, which is critically important for ensuring effective learning and teaching. In addition, this is in response to the limited empirical evidence on the impact of the existing approaches, both stand-alone and integrated courses in assessment, on developing PSTs' AfL literacy (Brevik et al., 2017; DeLuca & Klinger, 2010; Ogan-Bekiroglu & Suzuk, 2014).

Citation: Oo, C. Z., Alonzo, D., & Davison, C. (2023). Using a needs-based professional development program to enhance pre-service teacher assessment for learning literacy. *International Journal of Instruction*, 16(1), 781-800. <https://doi.org/10.29333/iji.2023.16144a>

Teacher AfL literacy, which has been defined by Alonzo (2016) as ‘teacher AfL literacy comprises “the knowledge and skills to make highly contextualised, fair, consistent and trustworthy assessment decisions, to inform learning and teaching, and to effectively support both students and teachers’ professional learning” (p. 58). The work of Black and Wiliam (1998) and Hattie (2008) highlight that preparing teachers to be literate in assessment, particularly the use of AfL, has the highest potential to increase students outcomes. This is further supported by many studies that highlight the critical role of teacher AfL literacy to significantly increase student outcomes (BOSTES, 2016; Charteris & Dargusch, 2018).

However, the level of teacher assessment literacy remains relatively low in various teacher education institutions including in Australia (Davison & Michell, 2014), the United States (DeLuca, Chavez, Bellara, et al., 2013), Europe (Tsagari & Vogt, 2017), Hong Kong (Lam, 2015), Philippines (Hailaya, 2014), and Myanmar (Hardman et al., 2016). This has drawn much attention from stakeholders across various educational institutions (Abera et al., 2017; Davison, 2013; Eyers, 2014; Wyatt-Smith et al., 2014). Similarly, the same trend is seen in pre-service teachers’ (PST) education. It has also been found that they are not well prepared to use appropriate AfL strategies to support student learning. This has been documented in the study of McGee & Colby (2014) and Siegel & Wissehr (2011) in the United States, DeLuca & Klinger (2010) and Volante & Fazio (2007) in Canada, Izcı & Caliskan (2017) in Turkey, and Rwehera (2017) in Myanmar. Their low AfL literacy is due to their lack of confidence in applying their AfL knowledge and building their skills (Ogan-Bekiroglu & Suzuk, 2014).

There have been few empirical investigations into ways to enhance PSTs’ AfL literacy to bridge the gap between assessment knowledge and skills, including building effective assessment practices (McGee & Colby, 2014). The integration of AfL into initial teacher education (ITE) curricula is still limited (DeLuca & Volante, 2016). Some teacher education curricula still lack the necessary assessment components to fully prepare PSTs (James & Pedder, 2006; Mertler & Campbell, 2005; Siegel & Wissehr, 2011). Therefore, there is a need to ensure that teacher education focuses on AfL, demonstrating how to integrate assessment into teaching. A more detailed account of the assessment component in teacher education is presented in the next section.

Assessment component in teacher education

As argued above, the assessment component of ITE curricula is insufficient for PSTs to be AfL literate (Grainger & Adie, 2014; McNeilly et al., 2022; Oo et al., 2022; Siegel & Wissehr, 2011). Over the past decade, a number of studies (Brookhart, 1999; Greenberg & Walsh, 2012; Popham, 2011) indicated the issues regarding assessment component in ITE which need to be addressed. Firstly, the insufficient teacher preparation in AfL is a concern. Many literature highlight the need for mandatory stand-alone assessment course that covers the current assessment systems and the skills demand of teachers and stakeholders (Greenberg & Walsh, 2012; Popham, 2011). For example, in the United States, the teacher education in terms of AfL is undeveloped as the literature points out inadequate teacher preparation for AfL (Berry & Daughtrey, 2011; Chesley & Jordan, 2012). In addition, it was highlighted by Berry & Daughtrey (2011) that teacher

education and professional development system are needed to reflect the common core standards and the new assessments. In Vietnam, the basic guidelines for teacher education by the Ministry of Education and Training (MOET) do not include the assessment component as a compulsory subject. The training related to assessment is not sufficient as the mandatory subjects, which include foundation knowledge, subject-matter knowledge and professional knowledge for teaching methodology and professional experience in their teacher education (Nguyen, 2016). Professional experience is used in this study to mean student teaching, practicum or field studies in other educational bureaucracies. This is the in-school learning experience of pre-service teachers under the supervision of school teachers serving as their mentors.

Secondly, the competing focuses of assessment component in ITE is another issue that needs to be addressed. Xu and Brown (2016) scoping review of 100 studies on teacher assessment literacy found that current assessment courses have different emphasis on assessment content. It was argued that the general description and the basic concept of the assessment in the course cannot provide adequate support to be literate in assessment. Greenberg & Walsh's (2012) analysis of 455 assessment courses from over 180 initial teacher education qualifications across 30 states in the United States illustrated that only 3% of these assessment programs showed adequate assessment preparation for student teachers in initial teacher education. In Myanmar, the assessment course in teacher education universities is offered as part of the Educational Test and Measurement subject which is mandatory. The contents are largely related to the construction of the tests, and the main functions of the tests and item analysis. Even though the different forms of assessment are covered, for example, summative assessment, formative assessment, performance assessment, alternative assessment, and portfolio, these are not highly emphasised and as consequence, teacher practices remain to be examination-driven.

Finally, the quality of the assessment course should be emphasised in teacher preparation. For example, the assessment curriculum in teacher education should be clearly stated and aligned with the philosophy of effective assessment practices with consideration of contextual factors (Hill et al., 2014). The findings of Brookhart (2011) suggests that the assessment course regarding communicating assessment results or using assessment data is needed to attend the quality of assessment courses given to PSTs. Therefore, Ogan-Bekiroglu and Suzuk (2014) point out that “content and context of an assessment course in teacher education programmes are prominent issues” (p. 362).

Given the need to fully prepare PSTs to take on the assessment demands in schools, there is a need to address their A_fL literacy needs. The current approaches, both stand-alone and integrated courses in assessment, have limited empirical evidence in terms of their impact on increasing PSTs' A_fL literacy (Brevik et al., 2017; DeLuca & Klinger, 2010; Ogan-Bekiroglu & Suzuk, 2014). Whilst these strategies are reported to raise pre-service teachers (PSTs) assessment literacy, research evidence demands for a more strategic approach with a greater focus on extended practice to provide opportunities for

PSTs to acquire practical AfL knowledge and skills (DeLuca & Volante, 2016; Deneen & Brown, 2016; Heck et al., 2020).

As the revision of ITE curricula requires a stronger rationale and policy support, and a very costly intervention, there is a need to explore other potential modalities to ensure that PSTs are equipped with the basic level of AfL literacy after completing their degree. One of the least explored means is the use of a professional development (PD) program that targets the specific needs of PSTs, which explain in detail in next section.

Use of a needs-based PD program for PSTs

Using a PD program in teacher education may improve PSTs' assessment literacy. Although this form of professional learning is mostly used for in-service teachers (Koh, 2011; Mertler, 2009), there is also evidence that it can be used for PSTs as well (Dayal & Lingam, 2015; DeLuca & Klinger, 2010; Grainger & Adie, 2014; McGee & Colby, 2014; Ogan-Bekiroglu & Suzuk, 2014).

Recently, there are few studies that highlight the use of PD program for PSTs (DeLuca et al., 2013; Giraldo & Murcia, 2019; Izci & Caliskan, 2017; McGee & Colby, 2014). However, the findings are inconclusive. For example, McGee and Colby (2014) found that PSTs increased their AfL literacy after engagement in the program. Similar findings were found by Siegel and Wissehr (2011), that is, a greater impact of the program on PSTs' assessment literacy. Contrary to these findings, DeLuca and Klinger (2010) found that PSTs have not developed their AfL literacy. Only 33% of PSTs have used alternative assessments in their actual practice. This is similar to the findings of Buck et al. (2010) that after their engagement in PD, there were still PSTs who have not developed a clear understanding of AfL principles and practices. A large number of PSTs still view assessment from a narrower perspective aligned with the summative role of assessments (Dayal & Lingam, 2015).

In addition, the program needs to conduct needs-analysis in order to inform the design of the program. Especially for AfL, identifying the current assessment knowledge and skills of PSTs is critical to inform the content of the PD program (Davison, 2015; Timperley, 2014). This also addresses the issue raised by Tsagari and Vogt (2017) about the difficulties of teachers in identifying their needs for professional development. Through problem analysis, the demands of dealing with the reality of AfL in school classrooms can be addressed. For example, in the study of Giraldo and Murcia (2019), their language assessment course was developed based on the findings of diagnostic stage in order to meet the needs of PSTs. They found that the context of the program is a pivotal for PSTs' assessment practices.

The timing of the assessment program is also needed to be considered in ITE. This is important for PSTs to get the experience of assessment into teaching (Hill et al., 2014; Ismail et al., 2019; Izci & Caliskan, 2017; Ogan-Bekiroglu & Suzuk, 2014; Oo et al., 2021). An intervention is needed to ensure that they can perform the assessment demands in the classroom during their professional experience. PSTs cannot be sufficiently prepared in AfL without opportunities for them to transfer their knowledge in the actual classroom practice (Hill et al., 2014; Izci & Caliskan, 2017). Therefore,

practical experiences to use their AfL knowledge in teaching should follow after acquisition of theoretical knowledge (Hill et al., 2014; Hill & Eyers, 2016).

In previous studies, there are some evidence that the improvement of teacher assessment literacy was moderated by gender and teaching experience. For example, Çambay and Kazanç (2021) found that gender is a significant variable for teachers in applying assessment literacy approaches. Moreover, female teachers have higher assessment literacy than male teachers (Çambay & Kazanç, 2021; Muhammad et al., 2020) especially in formative assessment (Guo & Xu, 2020). Similarly, teaching experience is found as the crucial role in improving teacher assessment literacy. In the study of Onalan and Gursoy (2020), teachers who have more experienced can support students to involve in assessment process. Zolfaghari and Ashraf (2015) also found that the positive relationship between teachers' assessment literacy and their teaching experience. Therefore, this study also considered gender and teaching experience in the analysis in order to examine if the inclusion of them may influence the improvement of PSTs assessment literacy.

Due to the competing results to support the usefulness of a PD program to build PSTs' AfL literacy, there is a need to find more empirical evidence to support its utilisation in the context of ITE. Thus, this paper is guided by the following research questions:

- (1) To what extent does a needs-based professional learning program enhance PSTs AfL literacy?
- (2) How do factors (gender and teaching experience) affect PSTs acquisition of assessment knowledge and skills?

METHOD

This paper reports the analysis of the surveys (the pre-survey and the post-survey) to explore the utilisation of a PD program in increasing PSTs' perception of their assessment literacy. This is part of a larger study aimed to investigate the development of PSTs AfL literacy in Myanmar using a design-based research (DBR) approach. Based on the qualities of DBR, a needs-based PD program was developed through the analysis of the pre-survey. Following the program, PSTs were encouraged to implement the new AfL strategies in their professional experience. At the end of the professional experience, the post-survey was conducted to examine the changes of PSTs AfL literacy.

Participants

The participants in this study are PSTs from one of the universities of Education in Myanmar. Non-probability population sampling method was used due to voluntary nature of participation. Fourth-year PSTs were chosen who already had the experience of practice teaching (professional experience) in the third year. In third year, all PSTs had practice teaching at schools for a month. A total of 335 PSTs: 30 PSTs in the intervention group and 305 PSTs in the cohort group were involved. Before the data collection process, the ethics approval was gained from the Institutional Review Board (IRB) of Ethics Committee and the written permission was gained from the Head of the participating university, Myanmar.

Of the 30 PSTs involved in the intervention group, 10 PSTs (33%) were male, and 20 PSTs (67%) were female. In terms of teaching experiences, 11 PSTs (37%) only had practice teaching experiences, 13 PSTs (43%) had two years and less than two years previous teaching experiences, and 6 PSTs (20%) had more than two years previous teaching experiences. Turning now to the cohort group, of the 305 PSTs involved, 54 PSTs (18%) were male and 251 PSTs (82%) were female. In terms of teaching experiences, 184 PSTs (60%) only had practice teaching experiences, 86 PSTs (28%) had two years and less than two years previous teaching experiences, and 35 PSTs (11%) had more than two years previous teaching experiences.

Instrument and development of PD program

A tailor-made PD program was developed for the purpose of building teacher professional learning. First, the needs analysis, using a survey instrument and follow-up semi-structured interviews, identified the current status of PST AfL literacy to inform the PD program. A survey instrument developed by Nguyen (2016) was adapted specifically for targeted audience. Nguyen's questionnaire was chosen for its good reliability with Cronbach's alpha high ($\alpha = .85$) using a five-point Likert scale. Five main sub-constructs of teacher perception of their AfL literacy were included: (i) beliefs and understanding of assessment; (ii) confidence in planning assessment; (iii) confidence in conducting assessment; (iv) confidence in using assessment; and (v) adherence to ethical issues in assessment. A total of 335 PSTs participated in the pre-survey. Second, the follow-up semi-structured individual interviews with 10 PSTs were conducted concerning PSTs' AfL experiences, their practices, and their perceptions. These interviews' results clarified PSTs' understanding of AfL.

The analysis results of pre-survey and the follow-up semi-structured individual interviews were used to inform the PD program. For example, the pre-survey results indicated that 30% of PST did not agree with using different assessment methods in their teaching while 70% agreed. However, the findings of the follow-up interview highlighted that they did not know most AfL strategies used in improving students' learning. They could not give a detailed clarification of the characteristics of formative assessment. These findings highlight that PSTs need to have a clear understanding of the concept of AfL and types of AfL. Therefore, PD program pertaining to AfL needs to be constructed with an understanding of the interrelationship between assessment, teaching and learning, and the characteristics of AfL. Based on these results, the content of the program was adjusted.

The program was grounded in the view of AfL literacy by Alonzo (2016) that reflects the principles of AfL. Each workshop session was constructed based on the main sources of material from the assessment project in Australia and Hong Kong. These materials were more suitable for designing PD programs in this study than other sources. First, the sources from Australia were chosen because the assessment system in Australia is more progressive in terms of AfL culture than in most other countries, and there is a wide range of resources available to support teacher formative assessment (Centre for Education Statistics and Evaluation, 2015), as opposed to the USA, where most material for teachers supports only summative assessment due to the influence of

high stakes standardised tests (Flaitz, 2011; Stiggins, 2005), with resources related to how to develop tests rather than formative assessment. Second, this program was also developed based on the sources from Hong Kong school-based assessment teacher support material which complemented the Australian resources. The education system in Hong Kong is still very exam-oriented (Hamp-Lyons, 2007), so the experience of Hong Kong is well-suited to apply to Myanmar, particularly on how to embed AfL in an examination-driven system.

The PD AfL literacy program includes four main parts, (i) AfL strategies, (ii) applying AfL to practice, (iii) developing teacher AfL literacy, and (iv) microteaching or peer-group practice teaching. Part 1 is designed to improve the beliefs and understanding of AfL, Part 2 includes how to apply AfL into practice, Part 3 focuses on developing teacher AfL literacy, and Part 4 includes peer-group practice teaching. The program was conducted over two months (a total of 36 hours), locate each session taking two hours. After the program, PSTs had the professional experience of using these assessments in the school classroom. At the end, the post-survey was conducted with the same participants who completed the pre-survey. The same questionnaire for the pre-survey and the post-survey developed by Nguyen (2016) were used. In addition, demographic data were collected for the analyses of the interactions of gender and teaching experiences (TE). These demographic data were also considered as the independent variables in order to investigate factors that affect PSTs acquisition of assessment knowledge and skills.

Data analysis

Analyses of the pre-survey and the post-survey was conducted to address the research questions of this study. The responses of the PSTs were encoded in Excel Spreadsheet. Data analyses were carried out using the IBM statistical package for social science (SPSS) v. 24. Repeated measures MANOVA was used to analyse whether there were any significant variances of dependent variables due to the program (the between-subject factors) and the time (the within-subject factors). Preliminary assumptions testing was conducted for the suitability of the data for repeated measures MANOVA. The analysis of repeated measures MANOVA was conducted using the multivariate test (the combined effects of the sub-constructs) and the univariate tests (the effects of each sub-construct). The analyses including the interactions of gender and the PD program, and the interactions of teaching experience and the PD program were conducted to examine if gender and teaching experience were factors in the findings.

FINDINGS

This section informs the PSTs changes in the overall perception of their AfL literacy and each of the five sub-constructs over time. This analysis was conducted to explore the utilisation of the PD program to build their perceived AfL literacy before their second professional experience. The empirical evidence associated with the impact of the PD program are presented.

Changes in the overall PSTs' Perceived AfL literacy

To explore the changes in the overall PSTs' perceived AfL literacy, multivariate tests were used to analyse whether there are significant effects of the time, intervention (PD program), gender and teaching experiences (TE) on PSTs overall AfL literacy. Results are shown in Table 1, time means the changes between the pre-survey and the post-survey, program means the differences between the intervention group and the cohort group. The interaction effects, Time \times Program; Time \times Program \times Gender; and Time \times Program \times TE, means the effect of one factor which depends on the conditions in the other factor.

Table 1
Results of the multivariate tests for PSTs overall AfL literacy

Effect	Value	<i>F</i>	<i>df</i> ₁	<i>df</i> ₂	<i>p</i>
Time	.94	2.52*	5	211	.031
Time \times Program	.93	3.41**	5	211	.006
Time \times Program \times Gender	.99	.11	5	211	.991
Time \times Program \times TE	.95	1.21	10	422	.284

Note. TE = teaching experience.

* $p < .05$. ** $p < .01$

There are noteworthy findings from the results of multivariate tests for PSTs overall perceived AfL literacy. The results indicate that the significant effect of the time and the program (intervention) and non-significant interaction effects including gender and teaching experience on the PSTs overall AfL literacy. First, the results show that time is affecting PSTs AfL literacy given by $F(5, 211) = 2.52, p < .05$, Wilks' Lambda = .94, $\eta^2 = .06$, indicating that PSTs overall AfL literacy changed over time. This means that over a period of time, PSTs can increase or decrease their overall AfL literacy. Second, there was a significant two-way interaction effect observed between the time and the program, $F(5, 211) = 3.41, p < .01$, Wilks' Lambda = .93, $\eta^2 = .08$, indicating the interaction effect of time and PD in increasing PSTs overall AfL literacy. This suggests that the PD program has an effect on the changes of PSTs overall AfL literacy over time.

In terms of gender and TE, there were no significant three-way interaction effects: between Time \times Program \times Gender, $F(5, 211) = .11, p > .05$, Wilks' Lambda = .99, $\eta^2 = .00$; and between Time \times Program \times TE, $F(10, 422) = 1.21, p > .05$, Wilks' Lambda = .95, $\eta^2 = .03$. This suggests that PSTs did not change their overall AfL literacy differently across gender or TE. Hence the program has an effect on the changes of PSTs AfL literacy non-dependent to gender and teaching experience.

Although the multivariate tests presented provide significant evidence to the effect of the time and the program to overall AfL literacy of PSTs, it is unclear from the results if the mean difference is significant in each sub-construct of AfL literacy. This is because the non-significant overall effect of the multivariate tests does not mean that the observed mean difference is not significant in each sub-construct. To answer this

limitation, a series of univariate ANOVAs was conducted to explore which particular sub-constructs are significantly influenced by time and other factors. The results of univariate analysis are presented in the succeeding sections.

Changes in PSTs' Perceived AfL literacy in the five sub-constructs

This section describes the changes in PSTs' perceived AfL literacy in each of the five sub-constructs over time, belief and understanding of assessment (dv1), the confidence in planning assessment (dv2), the confidence in conducting assessment (dv3), the confidence in using assessment data (dv4), and the adherence to ethical issues in assessment (dv5). The analysis results of the univariate tests are presented in terms of the effect of time, the interaction effect of time and program, and the interaction effect of time and program including gender and TE shown in Table 2.

Table 2
Univariate follow-up tests in each sub-construct of AfL literacy

Effect	DVs	MS	F	p	Adjusted p value by G-G	Adjusted p value by H-F
Time	dv1	22.38	3.43	.07	>.05	>.05
	dv2	7.14	2.80	.10	>.05	>.05
	dv3	8.38	1.87	.17	>.05	>.05
	dv4	44.83	9.99**	.002	<.05	<.05
	dv5	.22	.05	.82	>.05	>.05
Time × Program	dv1	26.18	4.01*	.047	<.05	<.05
	dv2	17.52	6.87**	.009	<.05	<.05
	dv3	23.80	5.31*	.022	<.05	<.05
	dv4	47.32	10.55*	.001	<.05	<.05
	dv5	.39	.10	.758	>.05	>.05
Time × Program × Gender	dv1	.06	.01	.93	>.05	>.05
	dv2	.76	.30	.59	>.05	>.05
	dv3	.41	.09	.76	>.05	>.05
	dv4	.22	.05	.83	>.05	>.05
	dv5	.04	.01	.92	>.05	>.05
Time × Program × TE	dv1	8.06	1.23	.29	>.05	>.05
	dv2	7.27	2.85	.06	>.05	>.05
	dv3	.67	.15	.86	>.05	>.05
	dv4	15.04	3.35*	.037	<.05	<.05
	dv5	1.44	.35	.71	>.05	>.05

Note. DVs = Dependent variables; dv1 = Belief and understanding of AfL; dv 2 = Confidence in planning assessment; dv 3 = Confidence in conducting assessment; dv 4 = Confidence in using assessment data; dv 5 = Adherence to ethical issues in assessment; G-G = Greenhouse-Geisser; H-F = Huynh-Feldt

* $p < .05$, ** $p < .01$.

Effect of Time

With regards to the effect of time, although there was a significant multivariate effect of time for the changes of PSTs' perceived overall AfL literacy, this was not the case for all sub-constructs. As shown in Table 2, results of the analysis provided a significant

result only for the sub-construct, *dv4*, $F(1,215) = 9.99$, $p < .01$, $\eta^2 = .04$, where a significant result is observed. This result suggests that the time impacted PSTs AfL literacy in the area of confidence in using assessment data which means that PSTs AfL literacy for this sub-construct will change over time even without the program.

Interaction Effect of Time and Program

In terms of the interaction effect between the time and the program (Time \times Program) on PSTs AfL literacy, the univariate analysis revealed that there were significant effects to the four sub-constructs: *dv1*; *dv2*; *dv3*; and *dv4* (see Table 2). However, there was no significant interaction effect revealed for the sub-construct, *dv5*. The interaction effects for each sub-construct are presented in detail in the following.

With regard to sub-construct *belief and understanding of assessment* (*dv1*), the results showed that building this sub-construct over time needed the PD program (intervention) which showed the significant interaction effect between time and program, $F(1,215) = 4.01$, $p < .05$, $\eta^2 = .02$, as seen in Table 2. According to Figure 1, the pre-post change for the intervention group was greater whereas the pre-post change for the cohort group was smaller as the mean difference was just 0.14. This finding implies that the PD program is needed as time could not change PSTs AfL literacy in this sub-construct.

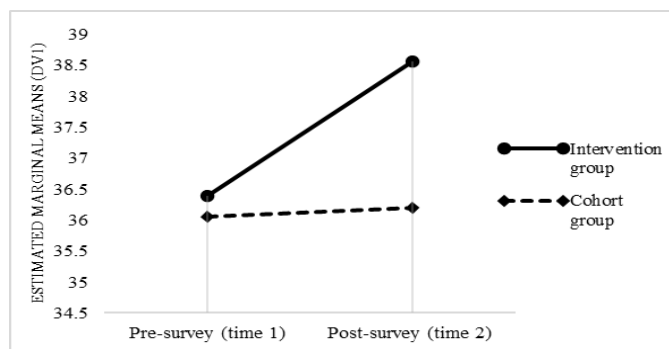


Figure 1
Interaction effect between time and program for *dv1*

In terms of the sub-construct, *confidence in planning assessment* (*dv2*), building PSTs confidence in this area over time needed the program which showed the significant interaction effect between the time and the program, $F(1,215) = 6.87$, $p < .01$, $\eta^2 = .03$. Figure 2 illustrates that the pre-post change for the intervention group was greater and positive in direction whereas the pre-post change for the cohort group was smaller and negative in direction. This finding implies that the PD program is needed as time alone could not change PSTs AfL literacy in this sub-construct.

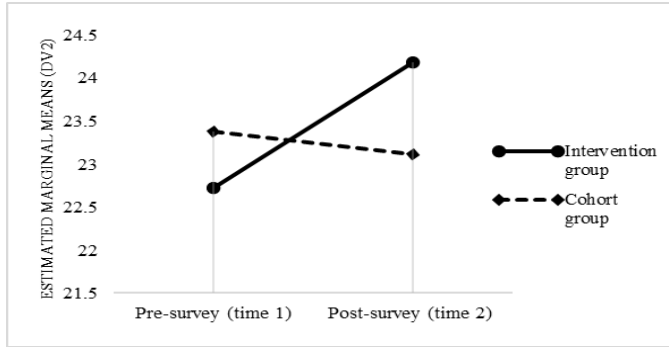


Figure 2
Interaction effect between time and program for dv2

For the sub-construct, *confidence in conducting assessment* (dv 3), similar to the two findings above, building PSTs confidence in this area over time needed the program which showed the significant interaction effect between the time and the program, $F(1,215) = 5.31, p < .05, \eta^2 = .02$ as seen in Table 2. As can be seen in Figure 3, the pre-post change for the intervention group was greater and positive in direction, whereas the pre-post change for the cohort group was smaller and negative in direction. This finding implies that the PD program is needed as time alone could not change PSTs AfL literacy in this sub-construct.

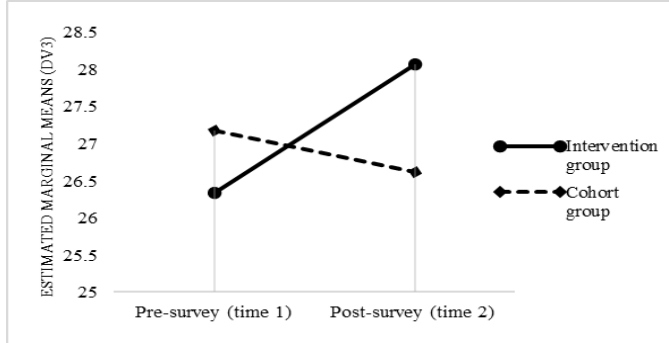


Figure 3
Interaction effect between time and program for dv3

In terms of the sub-construct, *confidence in using assessment data* (dv 4), building PSTs confidence in this area over time needed the program which showed the significant interaction effect between time and program, $F(1,215) = 10.55, p < .01, \eta^2 = .05$ as seen in Table 2. This result reveals a similar case to the other three sub-constructs. From the data in Figure 4, the pre-post change for the intervention group was greater and positive in direction whereas the pre-post change for the cohort group was smaller and negative in direction. This finding implies that the PD program is needed, however, univariate findings revealed that time could change PSTs AfL literacy in this sub-

construct. Like the other sub-constructs: dv2, and dv3, PSTs from the cohort group decreased this sub-construct over time. Therefore, it implies that, without the program, their confidence would become lower.

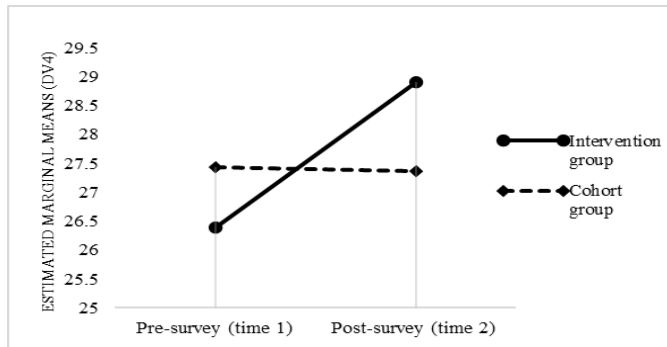


Figure 4
Interaction effect between time and program for dv4

For the sub-construct, *adherence to ethical issues in assessment* (dv 5), there were no statistically significant interaction effects between time and program, $F(1, 215) = .10$, $p > .05$, $\eta^2 = 0$ as seen in Table 2. What stands out in Figure 5 is that the two groups changed differently between the pre-survey and the post-survey in this sub-construct. The intervention group showed increased perception of their AfL literacy for this sub-construct while the cohort group showed smaller changes and negative in the direction. However, the program did not appear to have an effect on the changes of PSTs' perceived AfL literacy in terms of this construct.

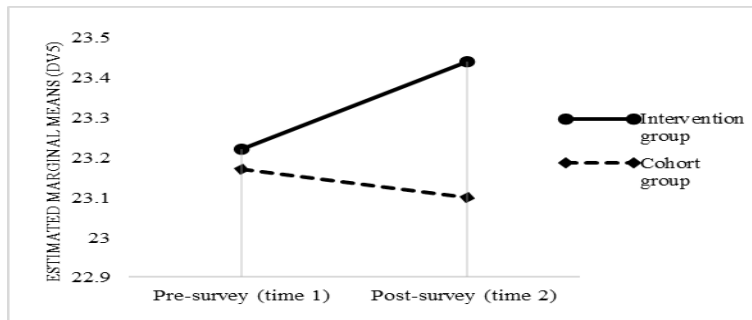


Figure 5
Interaction effect between time and program for dv5

In general, these results demonstrate the changes of their perceived AfL literacy especially in the four sub-constructs which were significantly increased because of the PD program. These suggest that the needs-based professional learning program was effective to change PSTs AfL literacy, particularly in these four sub-constructs: beliefs

and understanding of assessment, confidence in planning assessment, confidence in conducting assessment, and confidence in using assessment.

Interaction Effect of the Time and the Program including Gender and TE

The analysis including gender and TE were conducted to explore if these demographic factors affect PSTs' acquisition of assessment knowledge and skills. With regard to the interaction effects between Time \times Program \times Gender, there were no significant three-way multivariate interaction effects presented in Table 1. This was the case for all sub-constructs, shown in Table 2, as revealed by the non-significant interaction effects between Time \times Program \times Gender in each sub-construct. These results suggest that PSTs' AfL literacy which was affected by the program did not change differently across gender.

In terms of the interaction effects between Time \times Program \times TE, although there were no significant three-way multivariate interaction effects described in Table 1, this was not the case for all sub-constructs, shown in Table 2, as revealed by a significant interaction effect for the sub-construct, $F(2, 215) = 3.35, p < .05, \eta^2 = .03$. Meanwhile, these findings indicate non-significant results for the other sub-constructs. This suggests that PSTs' AfL literacy, especially in the areas of confidence in using assessment data, changed differently for PSTs from TE. Therefore, TE influence the changes in PSTs' confidence in using assessment.

DISCUSSION

The results of this paper add to the growing body of empirical evidence for the utilisation of a needs-based PD program to increase PSTs' perceived AfL literacy. The positive significant changes in PSTs' perception of their AfL literacy after their engagement in PD program confirms the usefulness of PD program as shown by other studies (DeLuca et al., 2013; DeLuca & Klinger, 2010; Gu, 2014; Levy-Vered & Alhija, 2015; McGee & Colby, 2014; Ogan-Bekiroglu & Suzuk, 2014). The approach we have used in designing and implementing the PD program might address the issue raised by Izci and Caliskan (2017) who found that after PD improvement of PSTs' conceptions of assessment did not change significantly. PD program used in this study was constructed based on the PSTs' assessment needs where the results of needs assessment are used as inputs to both PD design and content. Therefore, this study found that PSTs in the cohort group who had no PD showed smaller changes and were negative in the direction of their AfL literacy.

The key contribution of our study is the exploration of what sub-constructs of PSTs' AfL literacy can be improved by a PD program. Earlier studies highlight only the effect of PD program on increasing PSTs' understanding of key aspects of assessment (Hill et al., 2014; Ogan-Bekiroglu & Suzuk, 2014) and assessment beliefs (Eyers, 2014; Nguyen, 2016). The results of our study expanded this understanding to cover dimensions including (i) belief and understanding of assessment; (ii) confidence in planning assessment; (iii) confidence in conducting assessment; and (iv) confidence in using assessment data.

In terms of the belief and understanding of AfL, this result provides strong support for earlier studies (Hill et al., 2014; Ogan-Bekiroglu & Suzuk, 2014) which showed that the PD program had a greater impact on PST understanding of key aspects of assessment. Moreover, this study confirms previous studies (Eyers, 2014; Nguyen, 2016) that showed changes in PST assessment beliefs because of the program. These students in the comparison group who did not do the PD program had smaller changes in their AfL literacy, with a mean difference of just 0.14. Similar to the findings of Nguyen (2016), this shows is difficult to change beliefs in PST assessment literacy without a tailor-made intervention program.

Regarding confidence in designing assessment including confidence in planning assessment, confidence in conducting assessment and confidence in using assessment, our findings contradict with the reported results of Smith et al. (2009), and Volante and Fazio (2007) studies that PD program do not build PSTs confidence in this area of assessment. One possible explanation for this finding is the influence of extended practice that provided opportunities for PSTs to apply their theoretical knowledge in AfL in the actual classroom setting to acquire practical skills. This has been demonstrated in other studies that practical experience helps PSTs to develop critical skills in teaching (DeLuca & Klinger, 2010; Ogan-Bekiroglu & Suzuk, 2014). Comparing the design of the PD programs used in these studies with our study, we used a needs-based PD program. This implies that the design of the PD program is a critical element in building PSTs AfL literacy. In addition, our study highlights the potential use of DBR as a pedagogical tool to address PSTs AfL literacy needs. Using DBR as the pedagogical approach helps extend the theoretical knowledge of PSTs AfL literacy.

Another interesting finding of our study is that not all assessment skills can be developed using one approach only. Even with the overall increase in PSTs' AfL literacy, they have not increased their confidence level in terms of adherence to ethical standards in assessment. This contradicts McGee and Colby (2014) who argue that PSTs have higher mean scores on this sub-construct after PD. The relatively short length of this PD program, only one semester for a total of 36 hours, may not have provided enough exposure for PSTs to recognise unethical assessment practices. Hence, this suggests that it might be useful for PSTs to provide more opportunities to explore the problem of unethical assessment practices and improper uses of assessment. This is because PSTs need to know how to deal with some of the ethical issues relating to assessment and be able to recognise inappropriate uses of assessment (Green et al., 2007).

This study also found that the sub-construct "confidence in using assessment data" was affected by the time and length of teaching experiences, including the PD program. These results reflect those of Nguyen (2016) who found a correlation between teaching experience and confidence levels in assessment literacy. Therefore, the current study points out the need for differentiation between those with no teaching experiences and those with teaching experience in the PD program, especially for the sub-construct of confidence in using assessment data. In other words, PSTs should be provided with more opportunities for assessment practice.

CONCLUSION

We have demonstrated that a PD program designed using DBR is one of the possible approaches to enhance PSTs AfL literacy. The PD program in this study was designed based on PSTs needs and implemented both with PSTs microteaching and extended practical experience. As PSTs have a chance to extend their practical experience, they will understand the role of students being involved in assessment-based activities. The program has benefited the PSTs in terms of increasing their confidence in implementing assessment. This program was tailored for teachers from an exam-dominated culture and would appear to be suitable for both in-service teachers and PSTs, not only in the university where this study was undertaken, but also in other teacher-training institutes or colleges, to improve their professional learning of AfL literacy. In addition, this study demonstrates the value of DBR as a pedagogical approach to professional development in pre-service teacher education the contribution of explicit teaching of AfL literacy to PSTs as one way to change the assessment culture in schools to focus more on improving student learning.

Although this paper reports a strong empirical evidence for the utilisation of the PD program for PSTs, it was not exhaustive to include other factors that influence the development of their AfL literacy aside from PSTs' demographic profile. An in-depth interview with the PSTs in the intervention group about their implementation of AfL strategies during professional experience is underway to explore other potentially significant factors including the role of the supervising teachers, the role of the students, the role of the physical context and PSTs' personal effort. This qualitative data analysis will allow us to understand how PSTs develop or implement their AfL literacy in classroom practice.

REFERENCES

- Abera, G., Kedir, M., & Beyabeyin, M. (2017). The implementations and challenges of continuous assessment in public universities of Eastern Ethiopia. *International Journal of Instruction*, 10(4), 109–128. <https://doi.org/10.12973/iji.2017.1047a>
- Alonzo, D. (2016). *Development and application of a teacher assessment for learning (AfL) literacy tool* [PhD thesis, The University of New South Wales (UNSW)]. <http://unsworks.unsw.edu.au/fapi/datastream/unsworks:38345/SOURCE02?view=true>
- Berry, B., & Daughtrey, A. (2011). *New student assessments and advancing teaching as a results-oriented profession*. Center for Teaching Quality, Inc.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- BOSTES. (2016). *Learning assessment: A report on teaching assessment in initial teacher education in NSW*. NSW Government. <https://educationstandards.nsw.edu.au>
- 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 and 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 Spring*, 30, 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>

Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010). Making formative assessment discernable to pre-service teachers of science. *Journal of Research in Science Teaching*, 47(4), 402–421. <https://doi.org/10.1002/tea.20344>

Çambay, Ö., & Kazanç, S. (2021). Self-efficacy perceptions of science teachers for assessment literacy. *International Journal of Education and Literacy Studies*, 9(2), 142–150. <https://doi.org/10.7575/AIAC.IJELS.V.9N.2P.142>

Centre for Education Statistics and Evaluation. (2015). *Re-assessing assessment*. <http://www.cese.nsw.gov.au/publications-filter/report-re-assessing-assessment-2015>

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>

Chesley, G. M., & Jordan, J. (2012). What’s missing from teacher prep. *Educational Leadership*, 69(8), 41–45.

Davison, C. (2013). Innovation in assessment: Common misconceptions and problems. In K. Hyland & L. L. Wong (Eds.), *Innovation and change in English language education* (pp. 263–275). Routledge.

Davison, C. (2015). Enhancing teacher assessment literacy: Practising what we preach. *Paper Presented at 2015 Assessment in Schools Conference*.

Dayal, H. C., & Lingam, G. I. (2015). Fijian teachers’ conceptions of assessment. *Australian Journal of Teacher Education*, 40(8). <https://doi.org/10.14221/ajte.2015v40n8.3>

DeLuca, C., Chavez, T., Bellara, A., & Cao, C. (2013). Pedagogies for preservice assessment education: Supporting teacher candidates’ assessment literacy development. *Teacher Educator*, 48(2), 128–142. <https://doi.org/10.1080/08878730.2012.760024>

DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. *Assessment in Education: Principles, Policy & Practice*, 20(1), 107–126. <https://doi.org/10.1080/0969594X.2012.668870>

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., & 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).
- 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), 1–14. <https://doi.org/10.1080/2331186X.2016.1225380>
- Eyers, G. (2014). *Preservice teachers' assessment learning: Change, development and growth* [PhD thesis, The University of Auckland]. <http://hdl.handle.net/2292/23757>
- Flaitz, J. (2011). Assessment for learning: US perspectives. In R. Berry & B. Adamson (Eds.), *Assessment reform in education: Policy and practice*. Springer.
- 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., & Adie, L. (2014). How do preservice teacher education students move from novice to expert assessors? *Australian Journal of Teacher Education*, 39(7). <https://doi.org/10.14221/ajte.2014v39n7.9>
- Green, S. K., Johnson, R. L., Kim, D.-H., & Pope, N. S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23(7), 999–1011. <https://doi.org/10.1016/j.tate.2006.04.042>
- Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review*. http://www.nctq.org/p/edschools/docs/assessment_publication.pdf
- Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school? *Assessment in Education: Principles, Policy and Practice*, 21(3), 286–305. <https://doi.org/10.1080/0969594X.2013.836076>
- Guo, Q., & Xu, Y. (2020). Formative assessment use in university EFL writing instruction: A survey report from China. *Asia Pacific Journal of Education*, 41(2), 221–237. <https://doi.org/10.1080/02188791.2020.1798737>
- Hailaya, W. M. (2014). *Teacher assessment literacy and student outcomes in the Province of Tawi-Tawi, Philippines* [PhD Thesis, University of Adelaide]. <https://hdl.handle.net/2440/99098>
- Hamp-Lyons, L. (2007). The impact of testing practices on teaching: Ideologies and alternatives. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 487–504). Springer.
- 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, 98–118. <https://doi.org/10.1080/02188791.2014.906387>

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analysis relating to achievement*. Routledge.

Heck, D., Willis, A., Simon, S., Grainger, P., & Smith, K. (2020). Becoming a teacher: Scaffolding post-practicum reflection. In S. Billett, J. Orrell, D. Jackson, & F. Valencise-Forrester (Eds.), *Enriching Higher Education Students' Learning through Post-work Placement Interventions, Professional and Practice-based Learning* (pp. 173–188). Springer.

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.

Ismail, S. N., Busthami Nur, A. H., Raman, A., & Purnomo, Y. W. (2019). A mixed-method study of the epistemological teacher-beliefs towards educational research in classroom teaching practices. *International Journal of Instruction*, 12(4), 393–406. <https://doi.org/10.29333/iji.2019.12426a>

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 (IJRES)*, 3(2), 464–474. <https://doi.org/10.21890/ijres.327906>

James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *Curriculum Journal*, 17(2), 109–138. <https://doi.org/10.1080/09585170600792712>

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

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. (2015). Modelling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21(5–6), 378–406. <https://doi.org/10.1080/13803611.2015.1117980>

McGee, J., & Colby, S. (2014). Impact of an assessment course on teacher candidates' assessment literacy. *Action in Teacher Education*, 36, 522–532. <https://doi.org/10.1080/01626620.2014.977753>

McNeilly, E., Nickel, J., Burns, A., Danyluk, P., Kapoyannis, T., & Kendrick, A. H. (2022). The Canadian preservice teacher education practicum: An examination of fostering university and school collaboration, mentor teacher guidance, and re-centring

- the practicum. *Interchange*, 53(1), 65–74. <https://doi.org/10.1007/s10780-022-09455-7>
- 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). Measuring teachers' knowledge & application of classroom assessment concepts: Development of the assessment literacy inventory. *Paper Presented at the Annual Meeting of the American Educational Research Association*. <https://files.eric.ed.gov/fulltext/ED490355.pdf>
- Muhammad, N. B., Ali, N. B., Zamani, S. B., Yamin, N., & Ismail, N. N. B. (2020). Examining assessment literacy: A study of technical teacher. *European Journal of Molecular & Clinical Medicine*, 7(8), 705–717. https://ejmcm.com/article_3191.html
- Nguyen, L. T. H. (2016). *Development of Vietnamese pre-service EFL teachers' assessment literacy*. (Doctoral thesis). Victoria University of Wellington, New Zealand.
- 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>
- Onalan, O., & Gursoy, E. (2020). Primary and secondary level EFL teachers' use of assessment and assessment results in Turkey. *Journal on English Language Teaching*, 10(3), 1–11.
- Oo, C. Z., Alonzo, D., & Asih, R. (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>
- Oo, C. Z., Alonzo, D., & Davison, C. (2021). Pre-service teachers' decision-making and classroom assessment practices. *Frontiers in Education*, 6. <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/10.1080/00405840802577536>
- 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>
- Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22, 371–391. <https://doi.org/10.1007/s10972-011-9231-6>
- Smith, L. F., Corkery, G., & Buckley, J. (2009). How university students see assessment. *Paper Presented at the American Educational Research Association Annual Meeting*.
- Stiggins, R. J. (2005). From formative assessment to assessment for learning: A path to

success in standards-based schools. *The Phi Delta Kappan*, 87(4), 324–328.

Timperley, H. (2008). Teacher professional learning and development. *The International Academy of Education*, 1(18), 1–30.

Timperley, H. (2014). Using assessment information for professional learning. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning*. Springer.

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://doi.org/10.19126/suje.621319>

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.

Wyatt-Smith, C., Klenowski, V., & Colbert, P. (2014). *Designing assessment for quality learning*. Springer.

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>

Zolfaghari, S., & Ashraf, H. (2015). The relationship between EFL teachers' assessment literacy, their teaching experience, and their age: A case of Iranian EFL teachers. *Theory and Practice in Language Studies*, 5(12), 2556. <https://doi.org/10.17507/TPLS.0512.16>