

Exploring the perceptions of literacy in assessment for learning among high school English teachers

Sukarno¹; Nur Hidayanto Pancoro Setyo Putro¹; Ihtiar Fitrianiingsih^{1*}; Yasir Ayed Alsamiri²; Fadhil Mohammed Awadh Gharamah³; Ingrid A. Gavilan Tatin⁴

¹Universitas Negeri Yogyakarta, Indonesia

²Islamic University of Madinah, Saudi Arabia

³Yofea University College, Lahej University, Yemen

⁴Universidad Católica de Temuco, Chile

*Corresponding Author. E-mail: ihtiar.f@uny.ac.id

ARTICLE INFO

Article History

Submitted:

10 February 2024

Revised:

14 September 2024

Accepted:

18 September 2024

Keywords

assessment for learning;
English language teacher;
literacy

Scan Me:



ABSTRACT

In the field of English language education, the crucial role of assessment for learning (AFL) requires teachers to possess robust assessment literacy. This study explores AFL literacy among high school English teachers in Yogyakarta, Indonesia by utilizing Alonzo's validated AFL survey that establishes a comprehensive six-factor model, delineating teachers as assessors, pedagogists, student partners, motivators, learners, and stakeholder partners. Exploiting confirmatory factor analysis and examining demographic variations, this quantitative research invited 202 English teachers in the Special Region of Yogyakarta, Indonesia, selected purposively based on the geographical service area. Data were collected through an online questionnaire adopting Alonzo's 42-questions AFL and were analyzed quantitatively via Confirmatory Factor Analysis (CFA) with four indices, namely comparative fit index (CFI), Tucker-Lewis's index (TLI), root mean square error of approximation (RMSEA), and standard root mean square residual (SRMR). The findings substantiate the efficacy of the six-factor AFL model, underscoring educators' roles extending beyond traditional frameworks. The investigation also introduces a tool featuring detailed performance descriptors, addressing deficiencies, and harmonizing with AFL principles. It deduces that heightened foundational comprehension among English educators cultivates enhancements in AFL literacy and propels the refinement of professional evaluative competencies, thereby enriching the nuanced discourse surrounding AFL within language pedagogy. While the study's scope is confined to a specific geographical area and a limited cohort of participating instructors, it significantly enriches our comprehension of AFL literacy among English pedagogies. This research, therefore, provides a foundation for professional growth initiatives and facilitates enhancements in pedagogical approaches and academic achievement

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



To cite this article (in APA style):

Sukarno, S., Putro, N., Fitrianiingsih, I., Alsamiri, Y., Gharamah, F., & Tatin, I. (2024). Exploring the perceptions of literacy in assessment for learning among high school English teachers. *REID (Research and Evaluation in Education)*, 10(2), 143-154. doi:<https://doi.org/10.21831/reid.v10i2.71324>

INTRODUCTION

In contemporary educational discourse, an increasing body of literature acknowledges the pivotal role played by assessment for learning (AFL) as a determinant of efficacy and success in the English language learning process within the classroom. A growing number of researchers, exemplified by [Davison and Michell \(2014\)](#) and [Popham \(2009\)](#), have undertaken investigations focusing on the impact of positive assessment. Their collective findings emphasize the necessity for assessment literacy for learning to be integral to the professional knowledge, skills, and development of both in-service and pre-service teachers. In essence, a compelling imperative exists to cultivate a high level of expertise among teachers in utilizing assessment for learning, as their

proficiency in assessment literacy significantly influences their confidence in deploying diverse assessment strategies, whether teacher-developed or systemic (Davison & Michell, 2014). Similarly, the literacy levels of teachers in assessment for learning empower them to effectively employ assessment information for critical decision-making in the realms of learning and teaching, thereby contributing to enhanced student learning outcomes (Popham, 2009)

Recognizing the paramount significance of assessment for learning literacy, Alonzo (2016) was motivated to formulate and validate an instrument—the assessment for learning survey—for measuring teacher literacy in assessment. This survey, grounded in robust theoretical reasoning and empirical evidence, underwent validation through Rasch analysis at the item level and exploratory and confirmatory factor analyses to ensure alignment with analytical and factor-analytic requirements. The research outcomes yield a teacher literacy assessment instrument for learning assessment, featuring a novel framework for assessment for learning. This framework encompasses teachers as assessors, teachers as pedagogists, teachers as student partners, teachers as motivators, teachers as learners, and teachers as stakeholder partners.

Given the centrality of AFL, the comprehension of the six dimensions delineated by Alonzo (2016) regarding the teacher's role in implementing AfL is pivotal for in-service teachers. A failure to grasp these roles, encompassing teachers as assessors, pedagogists, student partners, motivators, learners, and stakeholder partners, is likely to significantly impact subsequent decision-making processes, including the interpretation of student ability levels and the subsequent pedagogical interventions post-assessment. Consequently, numerous efforts have been made to investigate teachers' perceptions and the perceived importance of assessment literacy in learning.

While extant research has explored teacher assessment of learning across diverse contexts and nations (e.g., Chappuis & Stiggins, 2002; Maclellan, 2010; Swaffield, 2011), there remains a dearth of studies utilizing Alonzo (2016) framework to examine the perceptions of English teachers, particularly those affiliated with the High School English Subject Teacher Conference (STC) in the Special Region of Yogyakarta, Indonesia. Considering the acknowledged importance of assessment for learning literacy and the utility of Alonzo's framework, this research endeavors to illuminate and articulate the assessment for learning literacy of high school English teachers within the association. The primary focus of this inquiry centers on elucidating how English teachers comprehend the concept of assessment for learning across the six dimensions outlined by Alonzo (2016), namely teachers as assessors, teachers as pedagogists, teachers as student partners, teachers as motivators, teachers as learners, and teachers as stakeholder partners. It is anticipated that the findings of this study will serve to enhance English teachers' foundational understanding of key concepts and foster improvements in assessment for learning literacy, ultimately facilitating the development of their professional assessment skills. Additionally, the research aims to discern potential variations in the interpretation of assessment for learning dimensions based on demographic variables such as gender and teaching experience among English teachers.

In recent decades, educational systems worldwide have embarked on endeavors to delineate and articulate teaching competencies that ensure the provision of high-quality learning and teaching experiences. This commitment is notably evident in the widespread movement towards formulating and implementing professional standards for educators across numerous countries (Chappuis & Stiggins, 2002). These professional standards aim to enhance teacher performance and the overall quality of education by affording comprehensive support and training about teachers' roles and responsibilities. Within this spectrum of professional standards, a particular emphasis is placed on teachers' professional knowledge in the realm of assessing student learning, particularly within the framework of AFL. The ensuing paragraphs delve into an exploration of the significance of AFL and various factors that can impact teachers' engagement with assessment for learning.

Learning Assessment

Presently, various countries have established teacher professional standards, such as the Pedagogical and Professional Competency Standards for Teachers in Indonesia, Teaching and

School Leadership Standards (AITSL) in Australia, National Competency-Based Teacher Standards for the Philippines, British Teacher Standards, and Teaching Standards in the Asian Region-Pacific (Alonzo, 2016). However, existing teacher performance assessment tools often remain overly broad, encompassing the entirety of learning and teaching functions. While the importance of effective teaching is readily acknowledged (Ferguson & Brown, 2000), the specific attributes defining effective teaching have yet to be fully elucidated. The multifaceted nature of teacher competency, spanning dimensions like classroom management, pedagogical and content knowledge, competency assessment, student-teacher relationships, and other vital aspects, contributes to the challenge of formulating standards that strike an appropriate balance between generality and specificity (Samson & Collins, 2012).

Recent trends, however, reflect an increasing emphasis within some standards on the teacher's role in enhancing student learning and performance, grounded in evidence correlating student achievement with teacher competency (Ferguson & Brown, 2000; Sanders & Rivers, 1996). The research underscores that highly effective teachers create conducive learning environments, with teacher contributions accounting for approximately 30% of the observed variance in student performance (Hattie & Timperley, 2008). Consequently, teacher competence emerges as the second most influential determinant of student learning achievement, following students themselves, who account for 50% of the variance in learning outcomes.

Notably, not all instructional activities undertaken by teachers yield desired learning improvements. Hattie and Timperley (2008) assert that while around 95% of teachers' actions in the classroom influence student learning achievement, not all teachers succeed in implementing activities that positively impact student learning. To optimize student learning and achievement, governmental efforts must target specific components of teaching competency that demonstrably contribute to positive learning outcomes. Furthermore, a commitment to supporting teachers in their primary function of facilitating effective student learning necessitates a focus on performance evaluation and continuous professional development aligned with the teacher's central role in positively impacting student learning and achievement.

Research in the realm of assessment for learning (AFL) emphasizes the significance of social interactions, cultural context, and students' belief systems, shaping students' identities and the dynamics of power and control in the classroom (Black et al., 2006; Cowie, 2005; Keppell & Carless, 2006; Marshall & Drummond, 2006; Marshall & William, 2006; Munns & Woodward, 2006). Although AFL aims to foster students' independence and self-directed learning, the self-regulation process heavily relies on interactions with teachers and peers to facilitate and support learning.

Building upon the studies and concepts, Alonzo (2016) conducted research to develop and validate an assessment for learning measuring tool in the form of a questionnaire. This effort resulted in a framework delineating the concepts and dimensions of assessment for learning (AFL), validated in several countries. The AFL construct comprises six dimensions: teacher as assessor, teacher as pedagogue, teacher as student partner, teacher as motivator, teacher as learner, and teacher as stakeholder partner. This framework serves as a foundation for the current research, aiming to provide insights into and understanding of assessment for learning literacy among English teachers affiliated with the High School English Subject Teacher Conference (MGMP) in Special Region of Yogyakarta.

Factors Related to Learning Assessment

Theoretical considerations and findings from various studies posit that learning assessment is correlated with several variables, both demographic and educational (Alonzo, 2016; Birenbaum & Rosenau, 2006; Willingham & Cole, 2013). Notably, demographic variables such as gender, academic year level, and academic achievement have garnered attention in the literature (Gallik, 1999; Gilbert & Fister, 2011; Gob, 2007; Kirby & DiPaola, 2011; Kuh, 2007; McDougal et al., 2011; Willingham & Cole, 2013). Consequently, this research incorporates various external variables related to demographics, including gender, GPA, and expectations for further education.

METHOD

This research is quantitative research that uses surveys as a primary data collection tool. This research was carried out by distributing online surveys or questionnaires collected via Google Forms. This research was conducted at the High School English Subject Teacher Conference (STC) in the Special Region of Yogyakarta. This research was conducted for seven months in 2023.

The respondents targeted to participate in this research were 202 English teachers who were members of the High School English Subject Teacher Conference (STC) in the Special Region of Yogyakarta, Indonesia. Those respondents were anonymized to maintain confidentiality. Furthermore, this research guarantees that the data obtained during this study were used only for research purposes.

The assessment for learning teacher literacy survey developed by [Alonzo \(2016\)](#) was used in this research. This survey includes 42 questions that reflect the six dimensions that will be looked at in this research. The 42 survey points are summarized in [Appendix 1](#).

Confirmatory factor analysis (CFA) is used as the main analysis in this research to see the extent to which the six dimensions proposed by [Alonzo \(2016\)](#) appear in this research. Confirmatory factor analysis was conducted with Mplus version 7.2 ([Muthén & Muthén, 2012](#)). Maximum likelihood estimation with robust standard errors (MLR) was chosen as the estimator in the analysis. Four evaluative fit indices were used to assess the model's fit: the standard root mean square residual (SRMR), the root mean square error of approximation (RMSEA), the Tucker-Lewis index (TLI), and the comparative fit index (CFI). A good and acceptable fit is indicated by CFI and TLI values greater than 0.90 ([Wang & Wang, 2020](#)). Although values up to 0.08 are acceptable, RMSEA and SRMR values of less than 0.05 are advised ([Lira et al., 2020](#)). In addition to these four indices, the chi-square statistic (χ^2) is also considered: a ratio of 1/3 or less between df and χ^2 indicates that the model is acceptable ([Millsap, 2019](#)). Since the χ^2 statistic is very sensitive to sample size, the significance of χ^2 is not used as the main criterion for rejecting the model. Next, the confirmatory factor analysis (CFA) results were used to see the extent to which the dimensions of learning assessment differed in terms of gender and teacher tenure using MANOVA.

FINDINGS AND DISCUSSION

Findings

A dataset derived from a self-assessment survey conducted among 202 secondary school English teachers in the Special Region of Yogyakarta (DIY) was utilized to examine the six-factor assessment model for literacy learning proposed by [Alonzo \(2016\)](#). The results of the CFA are detailed as follows.

Model Test Statistics and Fit Index of the Proposed Model

Evaluation of model fit commences with the examination of model fit statistics, crucial for determining the proximity of covariance matrices between the CFA-derived model and the sample ([Alonzo, 2016](#)). The Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR) constitute the key statistics. Acceptable and good fit criteria were established with CFI and TLI values exceeding 0.90 ([Sathyanarayana & Mohanasundaram, 2024; Wang & Wang, 2012](#)). RMSEA values below 0.05, and up to 0.08, and SRMR values less than 0.05 were deemed acceptable.

Multiple rounds of CFA indicated satisfactory fit statistics for the six-factor model within the defined thresholds. The RMSEA and SRMR indices yielded values of 0.047 and 0.055, respectively, below the threshold of 0.08 ([Wang & Wang, 2012](#)). CFI (0.910) and TLI (0.902) demonstrated a good and acceptable model fit. The four indices collectively affirm the acceptability of the 6-factor teacher AfL literacy model, per [Alonzo's \(2016\)](#) conceptualization.

In addition to the fit indices, the chi-square statistic (χ^2) was examined, with a ratio of 1/3 or less between degrees of freedom (df) and χ^2 indicating model acceptability (Byrne, 2010; Bentler 1990). The chi-square statistic for the model ($\chi^2= 887.91$, $df= 614$, $p=0.000$) suggested acceptability, given the sensitivity of χ^2 to sample size. Kline's (2005), criterion, which considers the ratio of χ^2 to degrees of freedom, indicated an acceptable model, with a χ^2/df ratio of 1.40.

Consequently, the results substantiate the support for the 6-factor literacy learning assessment model proposed by Alonzo (2016). In summary, the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), and chi-square (χ^2) collectively validate the acceptability of the six-factor model. However, 12 items out of the 42 items adopted from the original AfL scale by Alonzo (2016) were excluded due to overlap with other items.

Factor Loadings and Estimates for the Six-Factor Model

In addition to the goodness-of-fit indices, unstandardized and standardized estimates, along with their standard errors, were scrutinized to support the six-factor model. Table 2 summarizes the estimates, revealing that the ratios between standard estimates and associated standard errors are all equal to or greater than 1.96. This signifies the significance of all estimates of unconstrained loadings at the 0.05 level. Figure 1 visually depicts the factor structure of the six-factor model.

The CFA results, encompassing model fit indices and factor loadings, unequivocally endorse the proposed six-factor model derived from Alonzo's (2016) initial work. These six factors align precisely with Alonzo's (2016) model, positing that all dimensions of teacher assessment for literacy learning are discernible among pre-service English teachers in Indonesia. The factors are labelled as follows.

First, the role of the teacher as an assessor encompasses teachers' proficiency in assessment literacy, involving the design of assessment tasks, utilization of these tasks, measurement of student learning, and consideration of factors influencing student performance. Notably, the four items related to this factor demonstrated substantial loadings, ranging from 0.561 to 0.670.

Second, teachers as pedagogical experts constitute a factor comprising six items, reflecting teachers' assessment for learning (AfL) literacy. This literacy involves the identification of suitable teaching methods, and consideration of students' prior knowledge, current ability levels, and interests to inform the design of teaching and learning activities. The six items associated with this factor exhibited high loadings, ranging from 0.575 to 0.758.

Third, the factor of teachers as students' partners is characterized by the teacher's intention to collaborate with students and involve them in both the assessment and learning processes. This is evidenced by eight items, with factor loadings ranging from 0.546 to 0.786.

Fourth, the teacher as a motivator signifies the AfL literacy dimension of teachers in utilizing data collected from assessments to respond to the individual learning needs of students. This factor is represented by five items, each exhibiting factor loadings ranging from 0.611 to 0.780.

Fifth, teachers as learning teachers encapsulate the dimension where teachers reflect on their assessment experiences and employ assessment data to identify and address their professional development needs. Comprising four items, this factor demonstrated factor loadings ranging from 0.680 to 0.849. Lastly, teachers as stakeholder partners delineate teachers' AfL literacy in collaborating with stakeholders to respond to and enhance their assessment literacy. This factor is represented by three items, with factor loadings ranging from 0.601 to 0.809.

Figure 1 illustrates the impact of six factors on the teacher's role in the AfL. The configuration of these six factors aligns with the conceptual framework delineated by Alonzo (2016), particularly concerning teachers' assessment for learning (AfL) literacy at the construct level. This observation underscores the recognition among secondary school English teachers in the Special Region of Yogyakarta (DIY) that assessment transcends mere grade acquisition. Instead, teachers must assume multifaceted roles and responsibilities as proficient assessors, pedagogical experts, collaborative partners with students, motivators, continuous learners, and stakeholders.

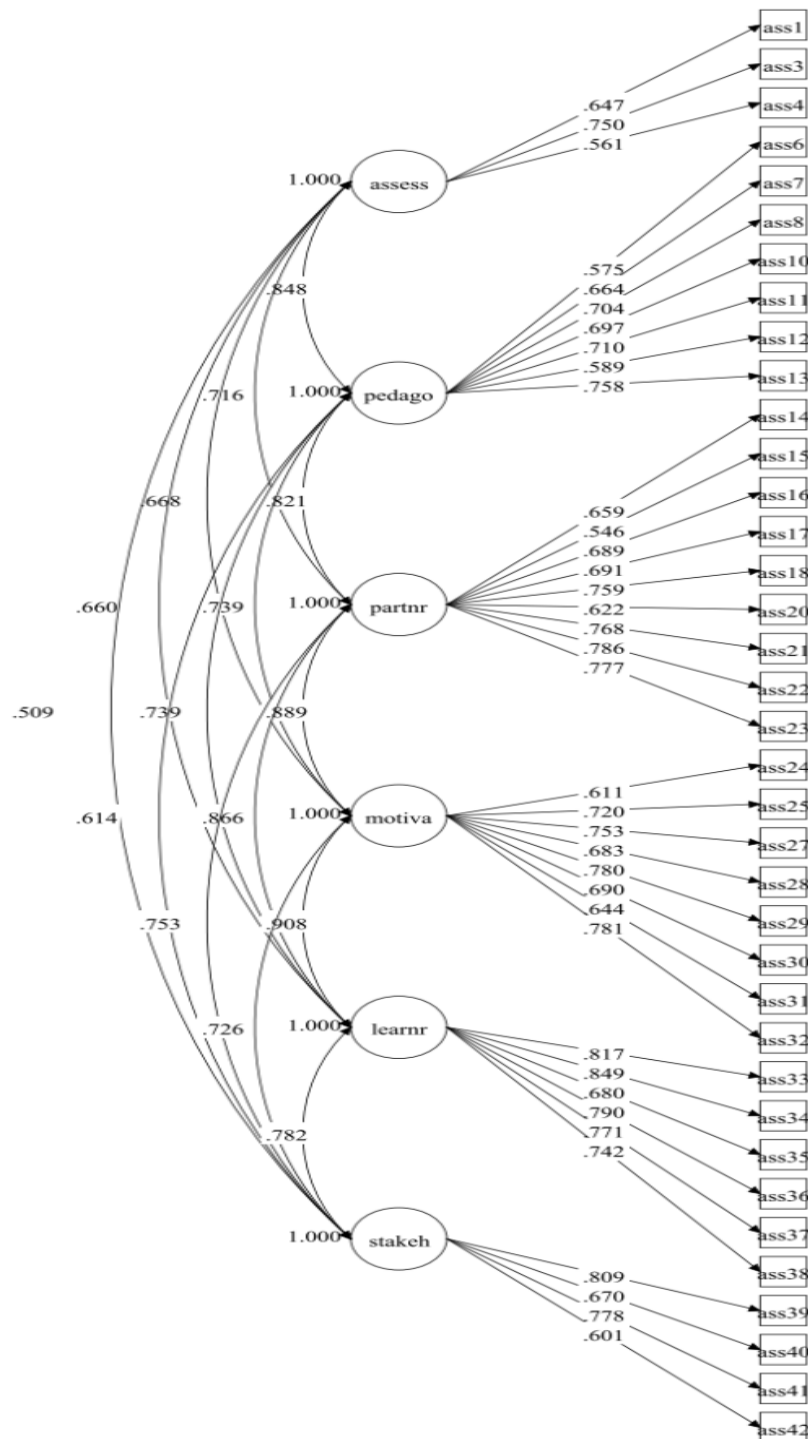


Figure 1. The Influence of Six Factors on the Teacher's Role in Assessment for Learning

Results of Multivariate Analysis of Variance (MANOVA)

Multivariate analysis of variance (MANOVA) was conducted to determine whether there were statistically significant differences in the six dimensions of assessment for learning (AfL) by independent groups (i.e. teacher gender and expected education). Interpretation of the analysis is taken from the coefficient value (F) or Sig. mark. If the F value is greater than the F table or Sig. the value is less than 0.05 ($p \leq 0.05$), then the alternative hypothesis is accepted, and the null hypothesis is rejected.

Table 1. MANOVA Assessment Test Results for Learning Dimensions Based on Gender

		Multivariate Tests				
Effect		Value	F	Hypothesis df	Error df	Sig.
Gender	Pillai's Trace	.029	.966b	6.000	195.000	.450
	Wilks' Lambda	.971	.966b	6.000	195.000	.451
	Hotelling's Trace	.030	.966b	6.000	195.000	.452
	Roy's Largest Root	.030	.966b	6.000	195.000	.453

a. Design: Intercept + Gender

b. Exact statistic

Based on Table 1, four types of tests—Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root—are employed. Utilizing Wilks' statistics, the outcomes indicate that gender does not exert a significant effect on the six dimensions of literacy learning assessment, as evidenced by $\Lambda = 0.971$, $F(6, 195) = 0.966$, $p = 0.450$. These findings suggest that male and female teachers share a common perspective regarding the roles of teachers as assessors, pedagogical experts, student partners, motivators, learners, and stakeholder partners.

Discussion

The primary objective of this research is to assess prospective teachers' comprehension of assessment for learning teacher (AfL) literacy. Empirical evidence derived from confirmatory factor analysis reveals that each item serves as a construct indicator referencing the AfL literacy model proposed by Alonzo (2016). These items are systematically arranged into six dimensions of teacher AfL literacy. The emergence of these six dimensions underscores a conceptualization of teacher AfL literacy emphasizing the pivotal role of teachers in utilizing assessment effectively to support student learning and ensure the development of stakeholder assessment literacy.

While the findings of this study align with previous conceptualizations of teacher assessment literacy as a construct with overarching dimensions, there are notable differences from previous research (Fulcher, 2012; Jin, 2021). These differences stem from a misalignment with AfL principles in the definition of teacher assessment literacy. Many existing frameworks and models, as highlighted by Alonzo (2016), primarily focus on teacher knowledge and skills related to measurement principles, neglecting the teacher's role in classroom assessment to foster student responsibility. This emphasis on measurement principles overlooks the context-dependent nature of assessment (Brookhart & Chen, 2020) and fails to consider other crucial teacher assessment skills essential for enhancing student learning (Sadler, 2010; Andrade & Brookhart, 2016; Separd, 2019).

The dimensions identified in this research reinforce a broader conceptualization of AfL literacy among teachers. Specifically, teacher abilities related to measurement principles are consolidated into one factor (Factor 1: teacher as assessor), while five additional dimensions contribute to measuring teacher AfL literacy. These dimensions encompass the teacher's role in utilizing assessment information to plan teaching and learning activities (teacher as pedagogist), using assessment to ensure high student motivation (teacher as motivator), involving students in learning and teaching (teachers as student partners), reflecting on assessment experiences to identify professional development needs (teachers as learning teachers), and ensuring literacy assessments of parents/guardians and the community in general (teachers as stakeholder partners). These six dimensions establish a clear link between teachers' ability to assess student learning and other capabilities requiring the use of assessment information to support student learning, including the acquisition of assessment experience by teachers and stakeholder assessment literacy.

Apart from providing robust empirical evidence supporting the six dimensions of teacher AfL literacy identified in this research, its multidimensional and multifunctional features amalgamate various theoretical models of teacher assessment literacy, culminating in a more

comprehensive model of teacher AfL literacy. As asserted by [Alonzo \(2016\)](#), existing theoretical models of teacher assessment literacy serve instructional and evaluative purposes, guiding teachers' classroom practices or assessing teacher assessment practices. In contrast, the AfL teacher literacy tool combines these functions, serving both as a learning guide and for evaluation purposes.

Teacher assessment for learning (AfL) literacy not only contributes to the augmentation of prevailing theoretical models but also encompasses the delineation of dimensions that constitute teacher AfL literacy. Additionally, it incorporates performance descriptions for each indicator, elucidating the five levels of performance. These dimension descriptions delineate the fundamental roles of teachers engaged in utilizing assessment to effectively bolster student learning. Simultaneously, the criteria and standards embedded within these descriptions serve as instruments for self-reflection, evaluation of assessment practices, and avenues for professional development. The inclusion of performance descriptions and standards in these tools addresses a deficiency in existing teacher assessment tools, specifically regarding the lack of detailed descriptions outlining what teachers can practically undertake. This identified gap aligns with observations made by [Brown and Bailey \(2008\)](#) and [Inbar-Lourie \(2008\)](#). Furthermore, the tool adheres to assessment for learning principles.

The novelty of this research lies in its development of a sophisticated framework for teacher assessment for learning (AfL) literacy, particularly through the inclusion of detailed performance descriptors. Unlike previous studies, this research provides five distinct performance levels, enabling a structured approach to teacher self-reflection and professional development. The elaboration of these performance descriptors is a key contribution, offering clarity on how teachers can practically implement AfL principles within the classroom—an area that has been under-explored in prior models ([Panadero et al., 2018](#); [Schildkamp et al., 2020](#)).

This study enhances earlier frameworks, such as those proposed by [Black and Wiliam \(2009\)](#), by extending the role of the teacher beyond the traditional assessor. It emphasizes multifaceted teacher responsibilities, including pedagogical guidance, student partnership, motivation, and collaboration with stakeholders. [Earl \(2013\)](#) has underscored the importance of integrating these roles to facilitate deeper student engagement, a dimension that this research addresses comprehensively. By including actionable performance descriptors, this model provides teachers with an essential tool to reflect on and elevate their AfL practices.

Additionally, the study aligns with [OECD's \(2020\)](#) emphasis on the need for performance metrics in teacher evaluation, addressing gaps in existing AfL frameworks where practical, detailed guidance was lacking. In this sense, the inclusion of performance descriptors responds to [Hattie and Timperley's \(2018\)](#) call for enhanced feedback mechanisms that serve both students and teachers. Thus, this research not only broadens the theoretical scope of AfL but also offers a tangible, evidence-based tool for improving teaching effectiveness and student outcomes.

CONCLUSION

Ultimately, this study has yielded significant knowledge regarding the assessment for learning (AfL) literacy of high school English teachers in DIY, illuminating their perspectives and methods. The use of Alonzo's verified AfL survey and confirmatory factor analysis has enhanced the empirical basis of the six-factor model, which includes teachers as assessors, pedagogists, student partners, motivators, learners, and stakeholder partners. The findings confirm the thoroughness of this approach, highlighting the diverse responsibilities that go beyond traditional frameworks. The use of a tool that includes performance descriptors fills the gaps that currently exist in teacher evaluation tools and is in line with the principles of evaluation for learning (AfL). Moreover, the study's emphasis on demographic variables provides detailed insights into how gender and teaching experience can impact teachers' viewpoints on AfL. While limited only in terms of the participating teachers in certain regions, this research ultimately enriches our comprehension of AfL literacy among English teachers, serving as a foundation for professional development activities and promoting advancements in teaching methodologies and student learning achievements. Further

investigation on the use of AfL literacy across various educational settings needs to be carried out to enhance the ability to assess AfL literacy in a way that reflects the unique demands of various disciplines. Incorporating AfL in teacher training programs would also offer a strong basis for teachers to enhance their teaching methodologies and student learning outcomes.

DISCLOSURE STATEMENT

The authors declare that they have no conflict of interest to disclose.

REFERENCES

- Alonzo, D. (2016). *Development and application of a teacher assessment for learning (AfL) literacy tool*. Doctoral dissertation, University of New South Wales. <https://doi.org/10.26190/unsworks/18809>
- Andrade, H. L., & Brookhart, S. M. (2016). The role of classroom assessment in supporting self-regulated learning. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation. The enabling power of assessment*, vol 4. Springer, https://doi.org/10.1007/978-3-319-39211-0_17
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588-606. <https://psycnet.apa.org/doi/10.1037/0033-2909.88.3.588>
- Birenbaum, M., & Rosenau, S. (2006). Assessment preferences, learning orientations, and learning strategies of pre-service and in-service teachers. *Journal of Education for Teaching: International Research and Pedagogy*, 32(2), 213-225. <https://doi.org/10.1080/02607470600655300>
- Black, P., McCormick, R., James, M., & Pedder, D. (2006). Learning how to learn and assessment for learning: A theoretical inquiry. *Research Papers in Education*, 21(2), 119–132. <https://doi.org/10.1080/02671520600615612>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Brookhart, S. M., & Chen, F. (2020). Formative assessment revisited. In *Handbook of educational assessment and evaluation* (pp. 123-136). Springer.
- Brown, D. J., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing*, 25(3). <https://doi.org/10.1177/0265532208090157>
- Byrne, B. M. (2010). *Structural equation modeling with Amos: Basic concepts, applications, and programming (2nd ed.)*. Taylor and Francis Group.
- Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership*, 60(1), 40–44.
- Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education*, 27(2), 199–214. <https://doi.org/10.1080/0950069042000276721>
- Davison, C., & Michell, M. (2014). EAL assessment: What do Australian teachers want? *TESOL in Context*, 24(2), 51–72. <https://search.informit.org/doi/pdf/10.3316/ielapa.949508765515560>
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Corwin, A Sage Company. <https://us.sagepub.com/en-us/nam/book/assessment-learning-1>

- Sukarno, Nur Hidayanto Pancoro Setyo Putro, Ihtiar Fitrianiingsih, Yasir Ayed Alsamiri, Fadhl Mohammed Awadh Gharamah, & Ingrid A. Gavilan Tatin
- Ferguson, R. F., & Brown, J. (2000). Certification test scores, teacher quality, and student achievement. In D. W. Grissmer & J. M. Ross (Eds.), *Analytic issues in the assessment of student achievement*, 133–156. <https://nces.ed.gov/pubs2000/2000050a.pdf>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Gallik, J. D. (1999). Do they read for pleasure? Recreational reading habits of college students. *Journal of Adolescent & Adult Literacy*, 42(6), 480–488.
- Gilbert, J., & Fister, B. (2011). Reading, risk, and reality: College students and reading for pleasure. *College and Research Libraries*, 72(5), 474–495. <https://doi.org/10.5860/crl-148>
- Gob, C. C. M. (2007). Teaching speaking in the language classroom. *SEAMEO Regional Language Centre*, 1–23.
- Hattie, J., & Timperley, H. (2018). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25(3), 385–402. <https://doi.org/10.1177/0265532208090158>
- Keppell, M., & Carless, D. (2006). Learning-oriented assessment: A technology-based case study. *Assessment in Education: Principles, Policy and Practice*, 13(2), 179–191. <https://doi.org/10.1080/09695940600703944>
- Kline, R. B. (2005). *Principles and practice of structural equation modelling*. Guilford Press. <https://www.guilford.com/books/Principles-and-Practice-of-Structural-Equation-Modeling/Rex-Kline/9781462551910>
- Kirby, M. M., & DiPaola, M. F. (2011). Academic optimism and community engagement in urban schools. *Journal of Educational Administration*, 49(5), 542–562. <https://doi.org/10.1108/09578231111159539>
- Kuh, G. D. (2007). What student engagement data tell us about college readiness. *PeerReview*, 9(1), 4–8. http://www.aacu.org/peerreview/pr-wi07/pr-wi07_analysis1.cfm
- Lira, M. R. d., Oliveira, A. S. d., França, R. A., Pereira, A. C. H., Godfrey, E., & Chaves, T. C. (2020). The Brazilian portuguese version of the exercise adherence rating scale (ears-br) showed acceptable reliability, validity and responsiveness in chronic low back pain. *BMC Musculoskeletal Disorders*, 21(1). <https://doi.org/10.1186/s12891-020-03308-z>
- Maclellan, E. (2010). Assessment for Learning: The differing perceptions of tutors and students. *Evaluation*, 903268926, 37–41. <https://doi.org/10.1080/0260293012006346>
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, 21(2), 133–149. <https://doi.org/10.1080/02671520600615638>
- Marshall, B., & William, D. (2006). *English inside the black box: Assessment for learning in the English classroom*. Granada Learning.
- McDougal, J. L., Bardos, A., & Meier, S. T. (2011). *Behavior intervention monitoring assessment system: Technical manual*. Multi-Health Systems Incorporated.
- Jin, J. (2021). Measuring teachers' knowledge and application of classroom assessment concepts: Development of assessment literacy inventory for Korean teachers. In *Classroom assessment*

SIG posters: On assessing collaborative learning and oral participation, designing assessments, formative assessment. <https://doi.org/10.3102/1686516>

- Millsap, R. E. (2019). *Statistical approaches to measurement invariance*. Routledge.
- Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: The REAL framework. *Assessment in Education: Principles, Policy and Practice*, 13(2), 193–213. <https://doi.org/10.1080/09695940600703969>
- Muthén, L., & Muthén, B. (2012). *Mplus Version 7 user's guide*. Muthén & Muthén. <https://doi.org/10.1111/j.1532-5415.2004.52225.x>
- OECD. (2020). *Education at a Glance 2020: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/69096873-en>
- Panadero, E., Andrade, H., & Brookhart, S. M. (2018). Fusing self-regulation and formative assessment: A road map of where we are, where we need to go, and how to get there. *Educational Psychology Review*, 30(3), 681–711. <https://doi.org/10.1007/s10648-018-9411-8>
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Samson, J. F., & Collins, B. A. (2012). *Preparing all teachers to meet the needs of English language learners applying research to policy and practice for teacher effectiveness*. Center for American Progress. <https://www.americanprogress.org/article/preparing-all-teachers-to-meet-the-needs-of-english-language-learners/>
- Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. University of Tennessee Value-Added Research and Assessment Center. <https://www.beteronderwijsnederland.nl/files/cumulative%20and%20residual%20effects%20of%20teachers.pdf>
- Schildkamp, K., van der Kleij, F. M., Heitink, M. C., & Veldhuis, M. (2020). Developing data use expertise in schools: Refining and validating a self-assessment instrument for teachers. *Studies in Educational Evaluation*, 65, 100835. <https://doi.org/10.1016/j.stueduc.2020.100835>
- Shepard, L. A. (2019). Classroom assessment to support teaching and learning. *The ANNALS of the American Academy of Political and Social Science*, 683(1), 183-200. <https://doi.org/10.1177/0002716219843818>
- Sathyanarayana, S., & Mohanasundaram, T. (2024). Fit indices in structural equation modeling and confirmatory factor analysis: Reporting guidelines. *Asian Journal of Economics, Business and Accounting*, 24(7), 561-77. <https://doi.org/10.9734/ajeba/2024/v24i71430>
- Swaffield, S. (2011). Getting to the heart of authentic Assessment for Learning. *Assessment in Education: Principles, Policy and Practice*, 18(4), 433–449. <https://doi.org/10.1080/0969594X.2011.582838>
- Wang, J., & Wang, X. (2020). *Structural equation modeling: Applications using Mplus*. Wiley. <https://doi.org/10.1002/9781119422730>
- Willingham, W. W., & Cole, N. S. (2013). *Gender and fair assessment*. Routledge.

Appendix 1. Research Instrument

Assessor	V3 Designs assessment tasks V7 Uses rubrics to assess students' learning V9 Considers factors that affect students' performance V10 Avoids interference in task completion V12 Engages in social moderation
Pedagogy Expert	V25 Translates learning standards to learning outcomes V28 Identifies appropriate teaching methods V29 Considers students' prior knowledge in lesson planning V30 Considers students' current level of abilities V31 Considers students' interest V32 Plans lessons according to students' learning needs V33 Tailors lessons to available resources V34 Develops teaching and learning resources
Student Partner	V17 Gives feedback on students' strengths and weaknesses V19 Assists students in using feedback to feed forward V27 Involves students in the development of learning outcomes V35 Makes students understand the learning outcomes V37 Involves students in the development of criteria and standards V39 Explains the criteria and standards V40 Develops students' capabilities in self and peer assessment V41 Engages students in self-assessment V42 Engages students in peer-assessment V43 Moderates feedback and results of self and peer assessment V46 Uses flexible teaching activities
Motivator	V46 Uses flexible teaching activities V47 Conducts assessment with consideration of student background V48 Develops an environment of trust V49 Ensures openness in the class V50 Builds students' interest to learn V51 Demonstrates belief in the ability of every student to improve V53 Affirms students' good performance V54 Clarifies students' misconceptions V55 Reinforces positive learning attitude of students
Teacher Learner	V58 Participates in professional development related to assessment V59 Engages in self- assessment/reflection V60 Engages in peer-review of teaching performance V62 Identifies subject- content knowledge needs V63 Searches new and relevant subject-content information V64 Undertakes further education/ training
Stakeholder Partner	V66 Collaborates with family to establish support activities V67 Informs community of the assessment practices V68 Reports to community about students' performance V70 Identifies key assessment and teaching issues for review
