

Formative assessment as a predictor of metacognitive skills: A structural equation modeling study

Roberto Bellido-García^{1*} , José Corro-Quispe² , Marco Rodríguez-Huamán³ ,
José Huaranga-Charapaqui⁴ , Roxana García Caveró¹ , Gerardo Rejas-Borjas² 

¹ Universidad San Ignacio de Loyola, Lima, PERU

² Universidad Nacional Mayor de San Marcos, Lima, PERU

³ Universidad de San Martín de Porres, Lima, PERU

⁴ Universidad Nacional de Educación Enrique Guzmán y Valle, Lurigancho-Chosica, PERU

Received 01 August 2025 ▪ Accepted 23 February 2026

Abstract

One of the central challenges in contemporary education lies in fostering autonomous and reflective thinking. Within this context, the present quantitative study investigates the role of formative assessment in promoting metacognitive skills at the secondary level. The sample comprised 156 students from Lima, Peru, who completed two validated instruments: the formative assessment practices questionnaire and the metacognitive awareness inventory, both of which demonstrated satisfactory reliability. Data were analyzed using structural equation modeling, with parameter estimation conducted under the diagonally weighted least squares method and supported by SPSS, AMOS, and R. The findings indicate a statistically significant association between formative assessment and metacognitive skills ($\beta = 1.090$; $p < 0.001$), particularly in the dimensions of metacognitive knowledge and regulation. Model fit indices (comparative fit index = 1.000; Tucker-Lewis index = 0.999; root mean square error of approximation = 0.026) confirmed the robustness of the proposed structural configuration. Overall, the results suggest that formative assessment is associated with stronger self-regulation and autonomous learning, underscoring its relevance for curricular integration as an effective approach to enhancing students' cognitive and metacognitive development.

Keywords: self-regulation, cognition, secondary education, formative assessment, metacognitive skills

INTRODUCTION

Over the past two decades, the Peruvian education system has undergone a paradigmatic shift, transitioning from a conventional model to the consolidation of a cognitive-oriented perspective. The former is associated with measurement, outcome achievement, and a behaviorist approach, whereas the latter seeks the holistic development of competencies.

Despite efforts led by the Ministry of Education (MINEDU) in recent years, alarming results persist in reading comprehension, mathematical proficiency, and scientific-technological competencies. Peruvian students continue to perform significantly below the expected

levels on international standardized assessments, such as PISA, where the country consistently ranks among the lowest-performing nations (MINEDU, 2022). This situation not only reflects structural deficiencies within the education system but also highlights the absence of strong political decisions aimed at achieving a genuine transformation aligned with the demands of 21st century education.

Structural adjustment policies in Peru provided the backdrop for the Peruvian educational reform, which was based on competency achievement. Thus, formative assessment, along with cognitive and metacognitive processes, gradually became an integral part of the basic education curriculum. However, they have not yet

Contribution to the literature

- In contrast to previous studies, this research employs a structural equation modeling (SEM) approach using diagonally weighted least squares (DWLS) estimation, which ensures statistical validity when analyzing ordinal data and enables a more accurate understanding of the phenomenon under study.
- The differentiated treatment of the two dimensions of metacognition has allowed for a more precise identification of the internal functioning of the processes that enhance self-regulated learning.
- Formative assessment is shown to be more than an evaluative technique. Based on the evidence generated in this study, it emerges as a transformative pedagogical strategy. As such, the findings provide a solid foundation for rethinking education policies that aim to foster self-regulated learning in mathematics and science.

achieved the concreteness that system managers would like to see. Therefore, there is reason to investigate the potential causes of this phenomenon. The current approach emphasizes its qualitative, reflexive nature and its focus on effective feedback for both teachers and students.

International publications have shown that formative assessment has a favorable impact on cognitive and metacognitive development (OECD, 2005; Shepard, 2006). In particular, the feedback strategy helps students reflect critically, self-regulate their performance, and make informed decisions based on previous successful outcomes (Ravela et al., 2008; Shepard, 2006). At the national level, there is a lack of systematic knowledge to support the evidence obtained beyond the country's borders (Mato et al., 2017; Mellado et al., 2024). Furthermore, there is poor coordination between cognitive and metacognitive processes, which makes it challenging to solve mathematical problems of a certain complexity in the classroom (Ullauri & Ullauri, 2018). Furthermore, failing to achieve optimal development of these skills reduces students' opportunities for academic success (Gamarra & Pujay, 2021). Therefore, it is necessary for educational research to provide feedback for the design of public policies and to serve as a tool for teacher development.

State of the Art

According to MINEDU (2024a), formative assessment is one of the primary indicators within the teaching-learning dimension. Following the analysis of 11,107 classroom sessions at the national level, findings revealed that only a small percentage of teachers implement formative assessment or promote constructive learning. Greater instructional complexity was associated with improved learning outcomes. However, most teachers only reached level 2, labeled "in progress." Similarly, 83.4% of the sessions scored at this same level in terms of monitoring and feedback, while 71.0% of the feedback provided on written work fell under level 1, "initiating." Only 1% of observed sessions achieved the highest levels of effectiveness (level 3 and level 4). In practice, teachers tend to offer feedback that does not address students' errors or difficulties in a

problematizing manner. Likewise, they rarely conduct thorough reviews of students' written tasks (MINEDU, 2024b).

Previous studies have demonstrated that formative assessment is an effective tool to promote metacognitive skills in mathematics and science education (Beekman et al., 2021; Fraile et al., 2021). When formative strategies—such as process-oriented feedback, self-assessment, and peer assessment—are integrated into instruction, students not only improve academic performance but also develop greater awareness of their learning processes (Wafubwa & Csikos, 2022). Feedback helps students identify where errors occur and serves as a premise for understanding how and why they should improve, which is critical for fostering self-regulation.

Despite growing evidence in favor of formative assessment, recent research has identified several nuances. For instance, Gedikli and Bulur (2022) reported improvements in academic achievement following active and experiential learning approaches, noting shifts in students' study behaviors. Their analysis indicates that when assessment prioritizes learning processes rather than outcomes alone, students tend to recognize not only the practical value of applying content but also the personal and social satisfaction derived from achievement.

A contrasting pattern emerged in the work of Ozan and Kincal (2018). Although their findings in the social sciences indicated a positive association between formative assessment and self-regulation, the difference did not reach statistical significance. This outcome reinforces the view that self-regulation develops progressively and requires sustained engagement. As such, expecting its full consolidation within a single academic term or short course may be unrealistic.

Further evidence was provided by Vinogradova and Skornyakova (2022), who observed that assessment modality influenced student outcomes. Performance tended to be higher in digitally mediated environments, possibly reflecting more favorable contextual conditions. However, traditional formats—such as paper-and-pencil assessments—showed a stronger relationship with the development of self-regulatory competencies. These findings suggest that structured, teacher-guided settings

that demand focused attention may serve as important contextual supports for fostering self-regulation.

Similarly, van der Linden et al. (2023) identified discrepancies between teachers' pedagogical beliefs and their enacted classroom practices. While many educators endorsed the value of self-regulated learning, feedback strategies designed to cultivate it were inconsistently implemented. Instruction frequently emphasized task completion or procedural accomplishment, operating under the implicit assumption that students had already acquired self-regulatory capacities. Such misalignment between belief and practice can limit the transformative potential of formative assessment.

Theoretical Framework

This study adopts the principles of *formative assessment* as conceptualized by Sanmartí (2020), viewing it not merely as an evaluative technique but as a pedagogical strategy that provides actionable information on the teaching and learning process. Such information enables educators to make informed decisions to help students improve their performance and progress toward autonomous learning through self-regulatory practices (Martínez, 2019). The core functions of formative assessment include reporting on student progress, identifying learning gaps, and delivering feedback to monitor learning goals.

Sanmartí (2020) proposes several foundational principles for formative assessment:

- (1) assessment drives learning processes, influencing both content and pedagogy,
- (2) its primary purpose is to support students' capacity to regulate their errors and limitations;
- (3) errors should be seen as opportunities for understanding and regulation,
- (4) if assessment does not foster student self-evaluation, it fails in its purpose—students must appropriate evaluative strategies and criteria,
- (5) assessment is a shared responsibility, but student-led evaluation is the most meaningful,
- (6) assessment outcomes are contingent on the quality of evaluative and regulatory practices in teaching and learning,
- (7) assessment motivates only when it equips students to use errors as levers for success,
- (8) a variety of instruments should be used to analyze data, fostering self-regulation and learner autonomy,
- (9) genuine classroom innovation arises not from standardized tests but from contextualized, original tools and methods, and
- (10) ultimately, assessment does not measure what students know, but the quality of the curriculum implemented.

Regarding *metacognitive skills*, this study emphasizes their importance as essential psychological mechanisms in the development of autonomous learning. Metacognition and self-regulation are interrelated constructs with overlapping yet distinct characteristics (Martínez, 2019). This research draws upon the metacognitive theory of Flavell (1979), Zimmerman's (2002) model of self-regulated learning (2002), and contributions from Boekaerts et al. (1999).

Zimmerman's (2002) model remains the most robust framework for understanding the metacognition-self-regulation nexus. He conceptualizes self-regulation as a cyclical process, with phases that recur in every learning experience. Each phase involves monitoring and control actions, informed by previous experiences and adjusted to the current context (Zimmerman, 2002). Affect plays a key role in this model, with self-perception and confidence differentiating self-regulation from "pure" metacognitive activity (Zimmerman, 2002).

In the first phase—forethought—students engage in task analysis (e.g., defining relevant goals) and develop motivational beliefs (e.g., self-efficacy). The second phase—performance—includes self-control (e.g., maintaining focus, managing effort, and deploying strategies) and self-observation (e.g., monitoring specific aspects of performance). Ultimately, self-reflection entails personal judgments that assess whether the strategies employed achieved the desired outcomes (Zimmerman, 2002).

Boekaerts et al. (1999) further distinguish between the first three phases (forethought, monitoring, control) and the final phase (self-reflection), noting that the former fall within the individual's capacity for self-regulation, whereas the latter is more context-dependent. Although reflection is not always fully controllable, it can and should be continuously monitored, as it governs attributions and influences earlier stages of regulation.

Operationalization of Variables

To operationalize formative assessment, this study employed the questionnaire developed by Cerón-Urzuá et al. (2020), designed to assess teacher performance related to formative evaluative practices. It comprises six dimensions:

- (1) **grading dimension:** feedback provided post-hetero-evaluation, using interim results to inform learning,
- (2) **proactive dimension:** design of upcoming activities with clear and precise guidelines,
- (3) **interactive dimension:** fostering student awareness of errors as learning opportunities,
- (4) **metacognitive dimension:** strategies directly related to self-regulation,
- (5) **retroactive dimension:** student engagement and intrinsic motivation, and

- (6) **adjusted dimension:** integration of improvements into session planning based on ongoing pedagogical reflection.

For metacognitive skills, the study employed the metacognitive awareness inventory (MAI) by Schraw and Dennison (1994), which was validated in the regional context by Huertas-Bustos et al. (2014). The MAI comprises two core dimensions:

- (1) **knowledge of cognition** refers to individuals' awareness of their cognitive processes (declarative, procedural, and conditional knowledge) and
- (2) **regulation of cognition** refers to the metacognitive activities involved in controlling learning and thinking processes, including planning, monitoring, and evaluation.

Research Gap

Although both formative assessment and metacognitive skills are part of the broader cognitive paradigm for 21st century learning, the intersection between them remains underexplored. A comprehensive review of the literature reveals a notable lack of empirical studies, particularly in Latin American contexts of secondary education. This study thus aims to fill a significant knowledge gap by providing rigorous empirical evidence on this relationship.

Hypotheses

Based on the theoretical and empirical foundations reviewed, the following research hypotheses are proposed:

- H1.** Formative assessment has a significant effect on metacognitive skills.
- H2.** Formative assessment has a significant effect on knowledge of cognition.
- H3.** Formative assessment has a significant effect on the regulation of cognition.

Finally, after reviewing the empirical background and theoretical foundations informing the study, three hypotheses were established to guide the analysis and subsequent discussion:

- H1.** Formative assessment is associated with metacognitive skills.
- H2.** Formative assessment is associated with knowledge of cognition.
- H3.** Formative assessment is associated with regulation of cognition.

METHOD

The study adopted a quantitative approach and a basic research orientation, aimed at extending the theoretical knowledge base and deepening

understanding of the underlying scientific and technical aspects of the phenomenon. A non-experimental, cross-sectional design was employed, as no variables were manipulated and data were collected at a single point in time (Creswell & Creswell, 2018). The research was informed by existing scientific literature, including research reports, specialized texts, peer-reviewed journal articles, and relevant statistical sources.

Given the cross-sectional and non-experimental nature of the study, the estimated structural relationships should not be interpreted as evidence of strict causal directionality. The findings are appropriately understood in terms of statistical associations among latent variables rather than causal effects.

Furthermore, the formative assessment practices questionnaire captured students' perceptions of teachers' assessment practices. Accordingly, the results reflect subjective representations of the evaluative process rather than direct observations of pedagogical implementation. This distinction should be carefully considered when interpreting the findings.

The formative assessment questionnaire comprised six dimensions—grade-related, proactive, interactive, retroactive, metacognitive, and adaptive—spanning 24 items, with a Cronbach's alpha of 0.913. The metacognitive skills questionnaire consisted of two dimensions—knowledge of cognition and regulation of cognition—encompassing 52 items and demonstrating a Cronbach's alpha of 0.954. Both instruments were administered using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The sample included 156 fifth-year secondary students from a school in Lima, Peru. Parental authorization was secured through informed consent prior to data collection. According to Kline (2016), for relatively small samples (100-200 cases), simple models with high factor loadings are acceptable.

It is important to note that the sample consisted exclusively of students from a single public secondary school in Metropolitan Lima. Accordingly, the findings should be interpreted within this specific institutional context, as the school's organizational, pedagogical, and sociocultural characteristics may have shaped the observed patterns.

Data were analyzed using SEM, a multivariate statistical technique that enables the examination of complex relationships between latent and observed variables and is widely applied in social science research (Kline, 2016). In this study, the structural model followed a parsimonious hierarchical configuration, with levels of analysis organized coherently and without redundancy. This arrangement allowed the phenomenon to be explained with the minimum necessary dimensions while preserving interpretative depth. Two primary latent variables were specified—formative assessment

and metacognitive skills—both with high, statistically significant factor loadings, thereby reducing the model's effective complexity. In addition, the standardized coefficients were substantial (exceeding 0.70 for most indicators), providing empirical support for model stability and the adequacy of the sample size. Formative assessment was conceptualized as a multidimensional construct composed of six specific dimensions (grade-related, proactive, interactive, metacognitive, retroactive, and adaptive), which enabled a more precise delineation of classroom assessment practices. Metacognitive skills, in turn, were examined through two central components: knowledge of cognition and regulation of cognition.

The structural analyses were conducted using AMOS v.24.0, applying the maximum likelihood estimation method, which is widely accepted in quantitative research. In parallel, R (version 4.2.2) was employed, specifically the DWLS estimator, which is well-suited for ordinal variables derived from Likert-type scales. Prior to model estimation, SPSS v.25.0 was used for data screening and to assess the basic assumptions underlying multivariate analysis.

To complement the covariance-based SEM approach, SmartPLS 4 was used to estimate models under the partial least squares framework (PLS-SEM), enabling the simultaneous evaluation of the measurement and structural components. This approach is especially appropriate for relatively small samples. Convergent validity was examined through the average variance extracted (AVE), which reflects the proportion of variance in the indicators explained by their respective constructs. The AVE values were as follows: knowledge of cognition (0.584), regulation of cognition (0.548), grade-related (0.582), proactive (0.564), interactive (0.528), retroactive (0.659), and metacognitive and adaptive (0.585). As recommended by Fornell and Larcker (1981), convergent validity is considered adequate when the AVE is ≥ 0.50 ; this criterion was met for all constructs.

In addition, the variance inflation factor (VIF) was calculated to assess potential collinearity among the model's predictor variables. The resulting VIF values were 3.01 for knowledge of cognition and 3.01 for regulation of cognition. Since values below 5.0 indicate the absence of significant structural collinearity concerns (Hair et al., 2021), the results support the stability of the specified model.

Regarding overall model fit, the obtained indices were highly satisfactory and within recommended thresholds. Notably, the comparative fit index (CFI) reached 1.000, the Tucker-Lewis index (TLI) 0.999, and the root mean square error of approximation (RMSEA) 0.026. The Chi-square (χ^2)-to-degrees-of-freedom (df) ratio was $\chi^2/df = 1.105$, with $p = .339$. Collectively, these

statistics indicate that the proposed model adequately captures the data's underlying structure.

Such values may be observed in parsimonious models estimated using DWLS, particularly in sub-models with limited degrees of freedom, where incremental fit indices can appear inflated and residuals may approach 0.000 due to rounding. To provide a more comprehensive evaluation, additional indices were reported, including goodness-of-fit index (GFI) (0.998), adjusted goodness-of-fit index (AGFI) (0.987), normed fit index (NFI) (0.996), and RMR (0.037), all of which indicated good overall model fit. These results remained stable across both AMOS and R, reinforcing the robustness of the dual-software analytical strategy.

The proper convergence of the model and the absence of inadmissible solutions—such as negative variances or non-significant loadings—further suggest that the sample size was adequate for the intended analyses. Likewise, in the specific models corresponding to the dimensions of knowledge and regulation of cognition, fit indices met established criteria, thereby strengthening the validity of the adopted modeling approach (Kline, 2016).

In addition, key statistical assumptions underlying the validity of the estimated structural model were examined. Multivariate normality was assessed using skewness (ranging from -2 to +2) and kurtosis (ranging from -7 to +7), both of which remained within acceptable thresholds reported in the methodological literature (Mardia, 1970). Linearity among the latent variables was evaluated using scatterplots and bivariate correlations, confirming direct, consistent, and theoretically coherent associations among the model's factors. Potential multicollinearity was also assessed through the VIF, with all values falling below the critical cut-off ($VIF < 5.0$), indicating the absence of structural redundancy among predictors.

Taken together, these findings provide empirical support for the adequacy of the specified model and the stability of the parameter estimates. The instruments demonstrated satisfactory reliability and validity, thereby reinforcing the measurement model's internal consistency. Accordingly, the results support the conclusion that formative assessment significantly affects the development of metacognitive skills. Moreover, the methodological approach adopted—grounded in SEM—offers a robust framework for examining complex relationships among latent constructs within educational research (Kline, 2016).

As shown in **Table 1**, a significant proportion of students (51.9%) demonstrated a high level of formative assessment, 45.5% exhibited a medium level, and only 2.6% showed a low level, indicating that a considerable number of students effectively monitor their learning process. Regarding the *grading-linked* dimension, 62.2% reported a medium level, highlighting the need to foster

Table 1. Description of formative assessment characteristics at the general level and by dimensions

Variables	Low		Medium		High		Total	
	F	P (%)	F	P (%)	F	P (%)	F	P (%)
Formative assessment	4	2.60	71	45.50	81	51.90	156	100
Linked to grading	15	9.60	97	62.20	44	28.20	156	100
Proactive	6	3.80	71	45.50	79	50.60	156	100
Interactive	9	5.80	71	45.50	76	48.70	156	100
Metacognitive	7	4.50	86	55.10	63	40.40	156	100
Feedback-based	10	6.40	65	41.70	81	51.90	156	100
Aligned (adjusted)	12	7.70	87	55.80	57	36.50	156	100

Note. F: Frequency & P: Percentage

Table 2. Description of metacognitive skills at the general level and by dimensions

Variables	Low		Medium		High		Total	
	F	P (%)	F	P (%)	F	P (%)	F	P (%)
Metacognitive skills	3	1.90	68	43.60	85	54.50	156	100
Knowledge of cognition	1	0.60	48	30.80	107	68.60	156	100
Regulation of cognition	5	3.20	72	46.20	79	50.60	156	100

Note. F: Frequency & P: Percentage

motivation beyond grades by reinforcing a genuine interest in learning. In the *proactive* dimension, 50.6% of students reached a high level, revealing their ability to anticipate academic demands and plan study strategies accordingly. Similarly, 48.7% scored high in the *interactive* dimension, suggesting the presence of effective collaboration and communication practices.

In the *metacognitive* dimension, 55.1% of students were at a medium level, as were 55.8% in the *Aligned* dimension—both representing solid baselines that could be further enhanced through strategies aimed at promoting reflection and the continuous adjustment of learning techniques. Finally, 51.9% attained a high level in the *feedback-based* dimension, indicating that many students are willing to reflect upon and revise their study behaviors based on evaluative input.

As shown in **Table 2**, a substantial proportion of students (54.5%) exhibited a high level of metacognitive skills, 43.6% demonstrated a medium level, and only 1.9% scored at a low level. This suggests that the majority of students possess adequate control over their learning processes; however, there remain opportunities to strengthen these skills among those at intermediate levels. Regarding the *knowledge of cognition* dimension, 68.6% of students achieved a high level, indicating that most have a solid understanding of their cognitive processes, which enables them to plan and monitor their learning effectively. Nonetheless, 30.8% of students were at a medium level, and 0.6% at a low level, indicating the need for instructional strategies that foster a deeper awareness of how learning occurs.

In the *regulation of cognition* dimension, 50.6% of students attained a high level, while 46.2% were at a medium level. This suggests that a significant portion of the student population has not yet fully developed the ability to manage and regulate their learning autonomously. Additionally, 3.2% demonstrated a low

level of regulation, underscoring the importance of targeted interventions aimed at enhancing students' executive control over their cognitive processes. The estimation method employed was DWLS, which is appropriate for categorical or ordinal data. The sample comprised 156 participants, all of whom were students. The model converged successfully, and the analysis was conducted using R version 4.2.2.

As shown in **Table 3**, the structural model results indicate a positive and statistically significant association between formative assessment (V1) and metacognitive skills (V2) ($\beta = 1.090$, $p < 0.001$). Although this standardized coefficient exceeds unity, it does not necessarily imply model misspecification. When estimation is conducted using DWLS, which relies on polychoric correlations and is particularly appropriate for ordinal data, standardized coefficients greater than 1 may occur when the relationship between latent variables is especially strong and residual variances are comparatively small (Kline, 2016).

In the present study, both constructs were operationalized as conceptually distinct and empirically independent variables using validated instruments that encompassed specific dimensions for each. This theoretical and methodological delimitation reduces the likelihood of construct redundancy or conceptual overlap, thereby strengthening the model's discriminant validity. Accordingly, the estimated coefficient should be interpreted as evidence of a high-magnitude structural relationship between students' perceptions of formative assessment practices and the development of metacognitive skills in secondary education.

The six dimensions constituting the formative assessment construct displayed a robust unifactorial structure, with positive and substantial standardized loadings: $\lambda = 0.765$ (grade-related), $\lambda = 0.706$ (proactive), $\lambda = 0.799$ (interactive), $\lambda = 0.786$ (metacognitive), $\lambda = 0.803$

Table 3. Structural and measurement model coefficients for the relationship between formative assessment and metacognitive skills

Relationship	Unstandardized coefficient	Standard error	Standardized coefficient	Significance
V2 ← V1	1.231	0.0614	1.090	< 0.001
Z1 ← V1	1.000	0.0000	0.765	< 0.001
Z2 ← V1	0.922	0.0425	0.706	< 0.001
Z3 ← V1	1.043	0.0448	0.799	< 0.001
Z4 ← V1	1.027	0.0464	0.786	< 0.001
Z5 ← V1	1.049	0.0441	0.803	< 0.001
Z6 ← V1	0.908	0.0430	0.695	< 0.001
DCC ← V2	1.000	0.0000	0.872	< 0.001
DRC ← V2	1.079	0.0489	0.941	< 0.001

Note. V1: Formative assessment; V2: Metacognitive skills; Z1: Linked to grading; Z2: Proactive; Z3: Interactive; Z4: Metacognitive; Z5: Feedback-based; Z6: Aligned; DCC: Knowledge of cognition; & DRC: Regulation of cognition

(retroactive), and $\lambda = 0.695$ (adaptive). These values support the internal coherence of the measurement model and the empirical representation of the latent construct. Similarly, the two dimensions underlying metacognitive skills—knowledge of cognition and regulation of cognition—showed high standardized loadings ($\lambda = 0.872$ and 0.941 , respectively), confirming their structural consistency within the analyzed sample.

Beyond supporting the psychometric adequacy of the instruments, these findings reinforce the theoretical relevance of the proposed model. The consistent association between students’ perceptions of formative assessment practices and higher levels of metacognitive skills suggests that assessment experiences function as a critical pedagogical component in fostering self-regulated learning. In this regard, the mere technical implementation of formative strategies appears insufficient; their effectiveness depends on being explicit, comprehensible, and pedagogically meaningful to students, thereby facilitating formative feedback processes, reflective engagement, and strategic adjustment.

From an applied standpoint, the results underscore the need for teacher education programs to move beyond a purely instrumental view of assessment techniques and toward the systematic integration of communicative and metacognitive dimensions. In particular, teachers must be prepared to clarify assessment criteria, processes, and purposes so that students not only receive feedback but also understand its formative rationale and incorporate it as a resource for self-regulation.

Within this framework, formative assessment should not be conceived merely as a technical mechanism for monitoring performance. Rather, it can be understood as a structural pedagogical device that shapes cognitive autonomy. Its transformative potential lies in encouraging students to take an active role in interpreting evidence, analyzing errors, and making strategic decisions about their own learning—especially at the secondary level, where abstract reasoning and

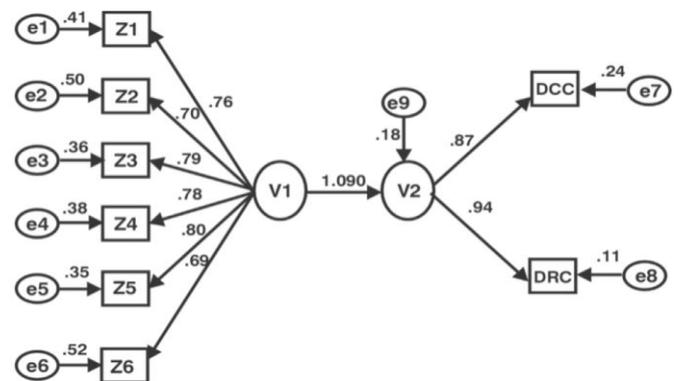


Figure 1. Estimated structural model between formative assessment and metacognitive skills (N = 156) (V1: Formative assessment; V2: Metacognitive skills; Z1: Linked to grading; Z2: Proactive; Z3: Interactive; Z4: Metacognitive; Z5: Feedback-based; Z6: Aligned; DCC: Knowledge of cognition; & DRC: Regulation of cognition) (Source: Authors’ own elaboration)

academic self-regulation become increasingly consolidated.

As illustrated in **Figure 1**, formative assessment (V1) emerges as a well-defined construct, reflected by its six indicators (Z1-Z6), with standardized factor loadings ranging from 0.69 to 0.80. The structural path from formative assessment (V1) to metacognitive skills (V2) is substantial ($\beta = 1.090$). Although this coefficient exceeds unity, it does not necessarily indicate model misspecification; rather, it may be due to the use of the DWLS estimator, which can yield standardized estimates above 1 under certain conditions.

Metacognitive skills (V2), in turn, function as a strong predictor of both knowledge of cognition (DCC) and regulation of cognition (DRC), as evidenced by their high standardized loadings (0.87 and 0.94, respectively). Each observed indicator includes an associated error term (e1-e9), representing the portion of variance not accounted for by the latent construct.

As shown in **Table 4**, the overall fit indices indicate that the estimated structural model provides an adequate fit to the observed data. The χ^2 test yielded a non-significant result: $CMIN = \chi^2 (19) = 21.0, p = 0.339$,

Table 4. GFIs for the estimated structural model between formative assessment and metacognitive skills

Name	Fit index	Value	Acceptable threshold
Normed fit index	NFI	0.996	≥ 0.90
Goodness-of-fit index	GFI	0.998	≥ 0.90
Adjusted goodness-of-fit index	AGFI	0.987	≥ 0.90
Relative fit index	RFI	0.993	≥ 0.90
Root mean square residual	RMR	0.037	≤ 0.05
Comparative fit index	CFI	1.000	≥ 0.90
Tucker-Lewis index	TLI	0.999	≥ 0.90
Root mean square error of approximation	RMSEA	0.026	< 0.06
Standardized root mean square residual	SRMR	0.037	< 0.05

Table 5. Structural model coefficients for the relationship between formative assessment and the knowledge of cognition dimension

Relationship	Coefficient	Standard error	Standardized coefficient	Significance	Structural and measurement model
DCG ← V1	1.470	.152	1.033	< 0.05	DCG = 1.03 x V1 + e10
Z1 ← V1	1.079	.119	.760	< 0.05	z1 = 0.76 x V1 + e6
Z2 ← V1	.903	.106	.717	< 0.05	z2 = 0.72 x V1 + e5
Z3 ← V1	1.064	.117	.769	< 0.05	z3 = 0.77 x V1 + e4
Z4 ← V1	1.044	.119	.737	< 0.05	z4 = 0.74 x V1 + e3
Z5 ← V1	1.208	.126	.797	< 0.05	z5 = 0.80 x V1 + e2
Z6 ← V1	1.000		.707	< 0.05	z6 = 0.71 x V1 + e1
m11 ← DCG	1.000		.784	< 0.05	m11 = 0.78 x DCG + e7
m12 ← DCG	.719	.065	.797	< 0.05	m12 = 0.80 x DCG + e8
m13 ← DCG	.820	.072	.812	< 0.05	m13 = 0.81 x DCG + e9

Note. V1: Formative assessment; DCG: Knowledge of cognition dimension of metacognitive skills; Z1: Linked to grading; Z2: Proactive; Z3: Interactive; Z4: Metacognitive; Z5: Feedback-based; Z6: Aligned; m11: Declarative knowledge; m12: Procedural knowledge; m13: Conditional knowledge; & e1 to e10: Error terms

Table 6. GFIs for the estimated structural model of the knowledge of cognition dimension

Name	Fit index	Value	Acceptable threshold*
Normed fit index	NFI	0.930	≥ 0.90
Goodness-of-fit index	GFI	0.916	≥ 0.90
Adjusted goodness-of-fit index	AGFI	0.855	≥ 0.85
Relative fit index	RFI	0.904	≥ 0.90
Root mean square residual	RMR	0.000	≤ 0.10

Note. *Acceptable thresholds based on standard SEM guidelines

exceeding the conventional threshold of 0.05. This finding suggests that there are no statistically significant discrepancies between the observed covariance matrix and the one reproduced by the model.

The normed χ^2 value (CMIN/df = 1.105) remained below the recommended cut-off of 2.0, indicating minimal divergence between the empirical data and the specified structural configuration.

Taken together, the NFI, GFI, AGFI, CFI, TLI, RMSEA, and SRMR indices fell within accepted ranges reported in the methodological literature, thereby supporting the adequacy and internal consistency of the proposed structural model within the adopted theoretical framework.

Table 5 presents the structural estimates for the cognition dimension. The standardized coefficient for this dimension ($\beta = 1.03$; $p < 0.05$) indicates a positive, statistically significant association with formative assessment within the specified model.

In addition, the subdimensions—declarative, procedural, and conditional knowledge—exhibited high standardized loadings, ranging from 0.78 to 0.81. These values support the construct's internal consistency and its adequate representation within the measurement model. Overall, the findings suggest that students' perceptions of formative assessment practices are associated with higher levels of metacognitive knowledge in the analyzed context.

As shown in **Table 6**, the goodness-of-fit indices for the estimated structural model fall within the ranges recommended in the methodological literature. NFI = 0.930 and the GFI = 0.916 both exceed the 0.90 threshold, indicating an adequate correspondence between the specified structure and the observed data. Likewise, the adjusted goodness-of-fit index (AGFI = 0.855) remains within the generally acceptable range (≥ 0.85).

As shown in **Table 7**, the structural estimates reveal statistically significant associations between formative assessment and the regulation of cognition dimension.

Table 7. Structural estimates for the relationship between formative assessment and the regulation of cognition dimension

Relationship	Coefficient	Standard error	Standardized coefficient	Significance	Structural and measurement model
DRC ← V1	2.092	.183	1.070	< 0.05	DRC = 1.07 x V1 + e12
Z1 ← V1	1.069	.111	.743	< 0.05	z1 = 0.74 x V1 + e1
Z2 ← V1	.908	.098	.713	< 0.05	z2 = 0.71 x V1 + e2
Z3 ← V1	1.077	.108	.769	< 0.05	z3 = 0.77 x V1 + e3
Z4 ← V1	1.153	.111	.804	< 0.05	z4 = 0.80 x V1 + e4
Z5 ← V1	1.162	.120	.757	< 0.05	z5 = 0.76 x V1 + e5
Z6 ← V1	1.000		.698	< 0.05	z6 = 0.70 x V1 + e6
m21 ← DRC	1.000		.862	< 0.05	m21 = 0.86 x DRC + e7
m22 ← DRC	1.396	.084	.904	< 0.05	m22 = 0.90 x DRC + e8
m23 ← DRC	.943	.061	.869	< 0.05	m23 = 0.87 x DRC + e9
m24 ← DRC	.624	.049	.783	< 0.05	m24 = 0.78 x DRC + e10
m25 ← DRC	.864	.060	.837	< 0.05	m25 = 0.84 x DRC + e11

Note. V1: Formative assessment; DRC: Regulation of cognition dimension; Z1: Linked to grading; Z2: Proactive; Z3: Interactive; Z4: Metacognitive; Z5: Feedback-based; Z6: Aligned; m21: Planning; m22: Organization; m23: Monitoring; m24: Debugging; m25: Evaluation; & e1 to e12: Error terms

Table 8. GFIs for the estimated structural model of the regulation of cognition dimension

Name	Fit index	Value	Acceptable threshold*
Normed fit index	NFI	0.952	≥ 0.90
Goodness-of-fit index	GFI	0.922	≥ 0.90
Adjusted goodness-of-fit index	AGFI	0.880	≥ 0.85
Relative fit index	RFI	0.939	≥ 0.90
Root mean square residual	RMR	0.000	≤ 0.10

Note. *Acceptable thresholds based on standard SEM guidelines

The corresponding standardized coefficient ($\beta = 1.070$; $p < 0.05$) indicates a positive relationship within the specified model.

The indicators associated with this dimension—planning, organization, monitoring, debugging, and evaluation—exhibited high standardized loadings, ranging from 0.70 to 0.90, supporting the internal consistency of the construct within the measurement model. Overall, these findings suggest that students’ perceptions of formative assessment practices are linked to higher levels of metacognitive regulation in the analyzed context.

As shown in **Table 8**, the fit indices for the estimated structural model corresponding to the regulation of cognition dimension fall within the ranges recommended in the methodological literature. In particular, the root mean square residual (RMR = 0.000) indicates an adequate correspondence between the specified structural configuration and the observed data.

DISCUSSION

Formative Assessment and Its Association with Metacognitive Skills

H1 received empirical support, indicating that students’ perceptions of formative assessment practices are statistically associated with higher levels of metacognitive skills. This pattern was confirmed in the present study (see **Table 3** and **Table 4**) and is consistent

with prior research (Gedikli & Buldur, 2022; Wafubwa & Csikos, 2022).

These findings suggest that an assessment grounded in a formative perspective encourages students to reflect on the learning processes they are constructing, with the teacher’s guidance and mediation. As a result, learners may act more effectively upon their own learning, developing greater awareness of the strengths and limitations inherent in the strategies they employ to achieve improved academic performance. Such experiences can foster deeper critical reflection on both the process and outcomes of learning, contributing to the consolidation of increasingly autonomous and reflective learners. Nevertheless, this relationship should be interpreted strictly in associative terms. Given the study’s cross-sectional design, no definitive causal inferences can be drawn between the examined variables.

Formative Assessment and Its Association with Knowledge of Cognition

H2 was supported, indicating that formative assessment is significantly associated with students’ immediate learning processes and contributes to the foundations of metacognitive development. This relationship was evidenced in the present study (see **Table 5** and **Table 6**) and aligns, at least partially, with previous research (Estrada et al., 2023; Vinogradova & Skorniyakova, 2022).

The findings suggest that formative assessment is linked to a deeper understanding of one's own cognitive processes, facilitating progress in declarative knowledge (what one knows), procedural knowledge (how tasks are performed), and conditional knowledge (when and why to apply specific knowledge). Through continuous and specific feedback, students cannot only recognize what they know but also evaluate how they apply that knowledge and in which contexts it is most effective.

Such conditions foster an adaptive learning environment in which students are assessed not solely on final outcomes but also on the design and regulation of their learning processes. Within this context, learners are encouraged to reflect on both their achievements and areas for improvement, thereby fostering a more comprehensive and functional understanding of their cognitive capacities.

Formative Assessment and Its Association with Regulation of Cognition

H3 was supported, indicating that formative assessment is meaningfully associated with how students plan, monitor, and evaluate the learning processes they undertake to achieve their academic goals. This relationship was evidenced in the present study (see **Table 7** and **Table 8**) and is partially consistent with prior research (Beekman et al., 2021; Xiao & Yang, 2019).

The preceding analysis suggests that when teachers adopt formative assessment practices—for instance, through the use of analytic rubrics—students are encouraged to reflect on the steps they have followed, assess their effectiveness, and make informed decisions for improvement. In doing so, learners are more likely to establish clear objectives prior to engaging in a task and to continuously monitor their progress toward those goals.

Several limitations should be acknowledged when interpreting the findings. First, the results must be understood as theoretically grounded statistical associations rather than evidence of causal directionality. Longitudinal or experimental designs in future research could provide stronger evidence regarding the directional nature of these relationships.

Second, formative assessment was measured through students' perceptions of teachers' practices, which may not fully correspond to actual classroom implementation. Although student perceptions constitute a valuable source of information, they may be shaped by contextual, affective, or motivational factors. Subsequent studies could complement self-report data with direct classroom observations or analyses of instructional performance.

A further limitation concerns the sample's scope. All participants were drawn from a single public secondary school in Lima, which constrains the generalizability of

the findings to other educational settings, whether public or private, urban or rural.

Although the estimated structural model demonstrated statistical consistency within the adopted theoretical framework, the findings should be interpreted as context-specific rather than representative of the Peruvian or broader Latin American educational system. Future research employing multicenter samples and comparative designs would help assess the stability of the proposed model across diverse institutional contexts.

CONCLUSIONS

Formative assessment was found to be positively and significantly associated with the development of metacognitive skills, thereby fulfilling the study's overall objective. The validated structural model indicated that the subdimensions linked to formative assessment (e.g., interactive, proactive, and retroactive) exhibited substantial factor loadings. Goodness-of-fit measures further confirmed the robustness of the proposed model (CFI = 1.000; RMSEA = 0.026). These findings suggest that formative assessment not only supports more effective monitoring of student progress but also reinforces higher-order cognitive processes, including metacognition.

In addition, formative assessment showed a statistically significant association with cognitive knowledge. The structural model yielded a coefficient of 1.033 ($p < 0.05$), indicating a strong relationship between these variables. The model demonstrated acceptable fit indices (NFI = 0.930; GFI = 0.916; RMR = 0.000), indicating acceptable internal consistency. This pattern suggests that formative assessment may actively foster the advancement of metacognitive knowledge, enabling students to better understand and regulate their own learning processes.

Similarly, formative assessment was significantly associated with cognitive regulation. The standardized coefficient of 1.070 ($p < 0.05$) in the structural model underscores the independent variable's explanatory power for self-regulatory processes. Model fit indices (NFI = 0.952; GFI = 0.922; RMR = 0.000) met accepted standards, confirming the adequacy of the specified model. Taken together, these results indicate that formative assessment may strengthen cognitive regulation competencies, positioning it as an effective pedagogical strategy for cultivating students who are more aware of and engaged in their own learning processes.

Author contributions: **RB-G:** conceptualization, investigation, resources, writing – review & editing; **JC-Q:** methodology, formal analysis, software, data curation (statistical dataset), validation (statistical procedures); **MR-H:** investigation, data collection, resources, writing – review & editing; **JH-C:** investigation, data collection, validation, writing – review & editing; **RGC:** literature review, investigation, data curation (non-statistical components),

writing - review & editing; **GR-B:** writing - language editing and translation, writing - review & editing. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study was reviewed and approved by the Institutional Ethics Committee of Educational Institution 3056, Great Britain, Lima, Peru (Approval Code: Directorial Resolution No. 059-2025-3056-DIE-GB-UGEL-02). Written informed consent was obtained from the parents or legal guardians of all participating students prior to data collection.

AI statement: The authors stated that an explicit statement regarding the use of Generative AI tools has been included. The declaration clarifies that AI-based tools were used exclusively for language editing and did not contribute to data generation, analysis, interpretation, or authorship responsibilities.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

- Beekman, K., Joosten-Ten Brinke, D., & Boshuizen, E. (2021). Sustainability of developed self-regulation by means of formative assessment among young adolescents: A longitudinal study. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.746819>
- Boekaerts, M., Zeidner, M., & Pintrich, R. (1999). *Handbook of self-regulation*. Elsevier. <https://doi.org/10.1016/B978-012109890-2/50030-5>
- Cerón-Urzuá, C., Cossio, M., Pezoa, P., & Gomez, R. (2020). Diseño y validación de un cuestionario para evaluar desempeño docente asociado a las prácticas evaluativas formativas [Design and validation of a questionnaire to evaluate teaching performance associated with formative assessment practices]. *Revista Complutense de Educación*, 31(4), 463-472. <https://doi.org/10.5209/rced.65512>
- Creswell, J., & Creswell, D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- Estrada-Araoz, G., Sayed, T., Niyazova, G., & Lami, D. (2023). Comparing the effects of computerized formative assessment vs. computerized dynamic assessment on developing EFL learners' reading motivation, reading self-concept, autonomy, and self-regulation. *Language Testing in Asia*, 13, Article 39. <https://doi.org/10.1186/s40468-023-00253-1>
- Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911. <https://doi.org/10.1037//0003-066x.34.10.906>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- Fraile, J., Ruiz-Bravo, P., Zamorano-Sande, D., & Orgaz-Rincón, D. (2021). Evaluación formativa, autorregulación, feedback y herramientas digitales: Uso de Socrative en educación superior [Formative assessment, self-regulation, feedback and digital tools: Use of Socrative in higher education]. *Retos*, 42, 724-734. <https://doi.org/10.47197/retos.v42i0.87067>
- Gamarra, G., & Pujay, O. (2021). Resolución de problemas, habilidades y rendimiento académico en la enseñanza de la matemática [Problem-solving, skills, and academic performance in mathematics teaching]. *Educación*, 45(1), 1-12. <https://doi.org/10.15517/revedu.v45i1.41237>
- Gedikli, H., & Buldur, S. (2022). The effects of formative assessment practices in science education on students' metacognitive knowledge and regulation skills. *Hacettepe University Journal of Education*, 37(4), 1393-1415. <https://doi.org/10.16986/HUJE.2022.454>
- Hair, F., Hult, M., Ringle, M., Sarstedt, M., Danks, P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer international publishing. <https://link.springer.com/content/pdf/10.1007/978-3-030-80519-7.pdf>
- Huertas-Bustos, P., Vesga, J., & Galindo, M. (2014). Validación del instrumento 'inventario de habilidades metacognitivas (MAI)' con estudiantes Colombianos [Validation of the 'metacognitive skills inventory (MAI)' instrument with Colombian students]. *Praxis & Saber*, 5(10), Article 55. <https://doi.org/10.19053/22160159.3022>
- Kline, B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.
- Mardia, V. (1970) Medidas de asimetría y curtosis multivariadas con aplicaciones [Multivariate measures of skewness and kurtosis with applications]. *Biometrika*, 57, 519-530. <https://doi.org/10.1093/biomet/57.3.519>
- Martínez, J. (2019). Autorregulación y metacognición en el aprendizaje: Un modelo de integración funcional [Self-regulation and metacognition in learning: A functional integration model]. *Desarrollo del Profesorado Universitario*, 7(1), 17-36. <https://doi.org/10.26852/2357593X.188>
- Mato, D., Espiñeira, E., y López, V. (2017). Impacto del uso de estrategias metacognitivas en la enseñanza de las matemáticas [Impact of the use of metacognitive strategies in mathematics teaching]. *Perfiles Educativos*, 39(158), 91-111. <https://doi.org/10.22201/iisue.24486167e.2017.158.58759>
- Mellado, B., Andías, R., Nogués, P., González, A., & Maldonado, D. (2024). Metacognición en el aprendizaje de las matemáticas en estudiantes de secundaria [Metacognition in mathematics learning

- in secondary school students]. *UISRAEL*, 11(3), 71-87. <https://doi.org/10.35290/rcui.v11n3.2024.1159>
- MINEDU. (2022). Evaluación formativa: Preguntas frecuentes [Formative assessment: Frequently asked questions]. MINEDU. <https://repositorio.minedu.gob.pe/handle/20.500.12799/8310>
- MINEDU. (2024a). Monitoreo de prácticas escolares 2024: Reporte de resultados [Monitoring of school practices 2024: Results report]. MINEDU. <https://repositorio.minedu.gob.pe/handle/20.500.12799/11564>
- MINEDU. (2024b). Resolución Viceministerial N° 048-2024-MINEDU [Vice-ministerial resolution no. 048-2024-MINEDU]. MINEDU. <https://www.gob.pe/institucion/minedu/normas-legales/5518274-048-2024-minedu>
- OECD. (2005). *Formative assessment: Improving learning in secondary classrooms*. OECD Publishing. <https://doi.org/10.1787/9789264007413-en>
- Ozan, C., & Kincal, R. Y. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences: Theory and Practice*, 18(1), 85-118. <https://files.eric.ed.gov/fulltext/EJ1179831.pdf>
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Rizo, M., & Wolff, L. (2008). Las evaluaciones educativas que América Latina necesita [The educational assessments that Latin America needs]. *Revista Iberoamericana de Evaluación Educativa*, 1(1), 51-63. <https://doi.org/10.15366/riee2008.1.1.004>
- Sanmartí, N. (2020). *10 ideas clave. Evaluar para aprender* [10 key ideas. Assessment for learning]. Graó Ediciones. https://www.academia.edu/34839425/ideas_clave_Evaluar_para_aprender
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460-475. <https://doi.org/10.1006/ceps.1994.1033>
- Shepard, L. (2006). Formative assessment: Caveat emptor. In C. A. Dwyer (Ed.), *The future of assessment*. Routledge. <https://doi.org/10.4324/9781315086545-12>
- Ullauri, J., & Ullauri, M. (2018). Metacognición: Razonamiento hipotético y resolución de problemas matemáticos [Metacognition: Hypothetical reasoning and mathematical problem solving]. *Revista Científica*, 3(8), 121-137. <https://doi.org/10.29394/Scientific.issn.2542-2987.2018.3.8.6.121-137>
- Vinogradova, E., & Skornyakova, R. (2022). Self-regulation and formative assessment format interrelation in mining engineering ESP course. *European Journal of Contemporary Education*, 11(4), 1283-1297. <https://doi.org/10.13187/ejced.2022.4.1283>
- Wafubwa, N., & Csíkos, C. (2022). Impact of formative assessment instructional approach on students' mathematics achievement and their metacognitive awareness. *International Journal of Instruction*, 15(2), 119-138. <https://doi.org/10.29333/iji.2022.1527a>
- Xiao, Y., & Yang, M. (2019). Formative assessment and self-regulated learning: How formative assessment supports students' self-regulation in English language learning. *System*, 81, 39-49. <https://doi.org/10.1016/j.system.2019.01.004>
- Zimmerman, J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

<https://www.ejmste.com>