

Expert-derived indicators for evaluating design thinking prototypes in teacher education

Mary Cris J. Go, Jovelyn G. Delosa, Christine C. Royo

Institute for Teacher Education, Northern Bukidnon State College, Manolo Fortich, Philippines

Article Info

Article history:

Received Jan 1, 2026

Revised Mar 12, 2026

Accepted May 14, 2026

Keywords:

Design thinking evaluation

Educational evaluation

Instrument development

Prototype assessment

Teacher education research

ABSTRACT

Evaluating design thinking prototypes in teacher education remains challenging due to the absence of standardized and theory-aligned evaluation criteria. Existing assessment practices are often inconsistent and insufficiently aligned with the process-oriented nature of design thinking-based research. This study aimed to develop expert-derived indicators to inform the construction of a prototype evaluation instrument for teacher education research. Using a qualitative instrument development approach, a focus group discussion (FGD) was conducted with research advisors and panel members experienced in evaluating undergraduate and graduate research outputs. Data were analyzed using Braun and Clarke's reflexive thematic analysis, resulting in six evaluation domains: problem alignment and functional relevance, novelty and intellectual property integrity, standardization of evaluation practices, alignment with design thinking processes, feasibility and sustainability, and demonstration of research and professional competencies. These domains were operationalized into a pool of observable indicators representing key dimensions of prototype quality. The resulting domains provide initial content validity evidence for the development of a standardized prototype evaluation instrument, supporting subsequent content validation and psychometric testing. The findings offer an expert-informed evaluation framework that contributes to more rigorous, transparent, and consistent assessment of design thinking prototypes in teacher education and higher education research.

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



Corresponding Author:

Mary Cris J. Go

Institute for Teacher Education, Northern Bukidnon State College

Manolo Fortich, Bukidnon, Philippines

Email: mjgo@nbsc.edu.ph

1. INTRODUCTION

The increasing integration of design thinking within teacher education has reshaped how research problems are conceptualized and addressed, particularly through the development of prototypes as tangible and context-responsive research outputs. Design thinking emphasizes human-centered problem identification, iterative ideation, prototyping, and testing, making it well aligned with educational contexts that seek practical solutions to authentic instructional and institutional challenges [1], [2]. As higher education institutions increasingly require research students to generate design-oriented outputs rather than purely theoretical studies, the need for systematic and defensible evaluation of such prototypes has become an important concern within educational evaluation and assessment practices.

Despite its growing adoption, the evaluation of design thinking prototypes in teacher education remains methodologically underdeveloped. Existing assessment practices are frequently inconsistent, reliant on individual evaluators' interpretations, and often adapted from tools originally designed for traditional research outputs. Educational assessment scholarship emphasizes the necessity of clear, transparent, and

standardized criteria to ensure fairness, validity, and reliability in evaluating student work [3], [4]. However, when design-oriented outputs are assessed using ad hoc or outcome-oriented measures that do not reflect their iterative and process-based nature, risks to construct validity and reliability increase. Despite the increasing use of design thinking in teacher education, no empirically grounded and expert-derived evaluation framework currently exists for assessing design thinking prototypes as research outputs.

The literature further indicates conceptual ambiguity in distinguishing design thinking prototypes from outputs generated through basic research or action research frameworks, particularly in higher education settings. While action research emphasizes implementation and measurable impact, design thinking prioritizes solution generation and iterative refinement prior to large-scale testing [5], [6]. When evaluation frameworks fail to account for these theoretical distinctions, design thinking outputs may be assessed using inappropriate effectiveness metrics, thereby undermining both pedagogical intent and evaluation accuracy. Recent studies highlight the importance of aligning assessment criteria with the theoretical foundations and processes underlying the research approach in order to produce meaningful and defensible evaluation outcomes [7], [8].

Although several studies have explored innovation assessment, rubric design, and competency-based evaluation in education, existing instruments primarily focus on traditional research quality indicators or emphasize implementation outcomes. These approaches often fail to capture critical dimensions of design thinking-based outputs, including process fidelity, originality safeguards, feasibility considerations, and the demonstration of research and professional competencies. This limitation underscores the need for a domain-specific evaluation framework grounded in expert consensus and aligned with the distinctive characteristics of design thinking-based inquiry [9], [10]. To address this gap, this study seeks to answer the following research question: what expert-derived evaluation domains and indicators are necessary for assessing design thinking prototypes in teacher education?

Using a qualitative instrument development approach, the study analyzes focus group discussion (FGD) data from research advisors and panel members through Braun and Clarke's reflexive thematic analysis [11]. The study contributes by systematically translating expert judgment into structured evaluation domains and observable indicators, thereby providing an empirically grounded foundation for a standardized prototype evaluation framework. By strengthening alignment between evaluation practices and the iterative, problem-driven nature of design thinking, this research advances methodological rigor, transparency, and consistency in the evaluation of design-oriented research outputs in teacher education and higher education.

2. METHOD

2.1. Research design

This study employed a qualitative instrument development design, constituting the initial phase of a broader evaluation tool construction process. Qualitative approaches are appropriate for generating evaluative indicators when constructs are complex, emergent, or insufficiently operationalized [12], [13]. Consistent with best practices in educational measurement, expert judgment was used as the primary source of evidence for identifying evaluation-relevant dimensions prior to quantitative validation [14]. The qualitative phase focused on eliciting expert perspectives to inform the development of indicators for evaluating design thinking prototypes produced by research students in teacher education. This phase aimed to establish conceptual clarity, define construct boundaries, and ensure content relevance—key prerequisites for subsequent validation and reliability testing [15].

2.2. Participants

A total of 12 research advisors and panel members participated in the FGD. All participants were affiliated with teacher education programs in higher education institutions and were actively involved in supervising and evaluating undergraduate and graduate research outputs. A purposive expert sampling strategy was employed to ensure the inclusion of participants with relevant domain expertise. Selection criteria required that participants: i) had experience supervising undergraduate or graduate research; ii) were directly involved in evaluating research proposals, manuscripts, or prototype-based outputs; and iii) demonstrated familiarity with design thinking, action research, or related educational research approaches.

The inclusion of expert participants is consistent with instrument development scholarship, which emphasizes that the generation of evaluative indicators should be grounded in the informed judgments of individuals with substantive content and assessment expertise [16], [17]. The number of participants was considered sufficient for the purposes of qualitative instrument development, as thematic saturation was achieved during analysis, with no substantively new evaluative dimensions emerging in the later stages of discussion. All participants were informed of the study's objectives and procedures, and written informed consent was obtained prior to data collection.

2.3. Data collection

Data were collected through a FGD designed to elicit shared evaluative standards and criteria used in assessing design thinking prototypes. FGDs are particularly suitable for instrument development studies because they facilitate interaction among experts, allowing for the clarification, negotiation, and refinement of evaluative constructs [18]. The FGD followed a semi-structured protocol that prompted participants to discuss current evaluation practices, challenges in assessing design thinking outputs, and criteria considered essential in judging prototype quality. The session was audio-recorded with permission and transcribed verbatim to ensure accuracy and completeness of the data [19].

2.4. Data analysis

The transcribed FGD data were analyzed using Braun and Clarke's six-phase thematic analysis [20], which provides a systematic and flexible approach for identifying patterns relevant to instrument development and evaluation research. The analysis followed a reflexive thematic analysis approach, emphasizing interpretive analytic development rather than code-reliability or inter-coder quantification, consistent with the study's objective of translating expert judgments into evaluative constructs for instrument development. The analysis followed Braun and Clarke's six-phase reflexive thematic analysis procedure. First, the researchers familiarized themselves with the data through repeated reading of the transcripts. Second, initial codes were generated to capture evaluation-related ideas emerging from the discussions. Third, patterns among the codes were examined to identify potential evaluative dimensions. Fourth, candidate themes were reviewed and refined to ensure conceptual coherence and clear distinction between domains. Fifth, themes were defined and named as evaluation domains relevant to prototype assessment. Finally, an analytic narrative was developed linking the identified domains to prototype evaluation criteria.

Thematic analysis was selected for its capacity to support both inductive coding and analytic abstraction, making it suitable for transforming expert discourse into structured evaluation indicators [21]. To enhance analytic rigor, emerging themes were iteratively reviewed and compared against established evaluative constructs in educational assessment literature to ensure conceptual alignment with recognized evaluation frameworks. Coding decisions and theme development processes were documented throughout the analysis to maintain an audit trail and support transparency in the analytic procedure.

2.5. Trustworthiness of the study

Methodological rigor was ensured through strategies consistent with qualitative trustworthiness [22]. Credibility was supported through the use of expert participants and sustained engagement with the data. Dependability was enhanced by applying a systematic analytic procedure and maintaining detailed documentation of coding and theme development. Confirmability was addressed by grounding interpretations directly in the data and minimizing researcher bias through reflexive memoing. Transferability was supported by providing sufficient methodological detail to enable application of findings in similar teacher education and higher education contexts [23].

2.6. Ethical considerations

The study complied with established ethical standards for research involving human participants. Participation was voluntary, informed consent was obtained prior to data collection, and confidentiality was maintained throughout the research process. Ethical clearance was secured from an institutional research ethics committee, consistent with national regulations and international ethical guidelines governing educational research [24].

3. RESULTS AND DISCUSSION

The thematic analysis of expert focus group data yielded six interrelated evaluation domains that collectively define quality and rigor in the assessment of design thinking prototypes in teacher education. Rather than functioning as isolated criteria, these domains operate as a coherent evaluative framework that integrates problem relevance, originality, process fidelity, feasibility, and competency demonstration. Together, they respond to persistent challenges in evaluating design-oriented research outputs using criteria originally designed for traditional academic products. The domains are presented with emphasis on their distinct evaluative contributions and implications for prototype assessment.

3.1. Problem alignment and functional relevance of prototypes

Experts consistently emphasized that effective prototype evaluation begins with assessing the extent to which a prototype is aligned with a clearly articulated educational problem and whether it functions as a viable solution within its intended context. Participants underscored that, prototypes should not merely demonstrate creativity or technical sophistication but must respond to authentic instructional or institutional

needs and exhibit coherence between the identified problem, research objectives, and proposed solution. This emphasis aligns with evaluation literature asserting that relevance and functional alignment are central criteria for judging the quality of educational innovations [25], [26]. In design-oriented research, the value of a prototype lies primarily in its problem–solution fit rather than in aesthetic or conceptual appeal alone. Accordingly, indicators derived from this domain emphasize clarity of problem definition, appropriateness of the solution, and contextual applicability, reinforcing construct relevance and interpretive validity in prototype evaluation [27].

3.2. Novelty, originality, and intellectual property integrity

A second domain focused on the evaluation of originality while ensuring adherence to intellectual property and ethical standards. Experts expressed concern that prototypes are often judged as “innovative” based solely on localized or contextual novelty, without sufficient engagement with existing literature or comparable solutions beyond the immediate institutional setting. Such practices increase the risk of unintentional duplication and weaken the scholarly credibility of research outputs.

This domain reflects broader concerns in educational evaluation regarding academic integrity and responsible innovation [28], [29]. Indicators associated with this domain extend beyond surface-level novelty by incorporating criteria related to literature grounding, differentiation from prior work, and ethical attribution. By integrating originality with intellectual property integrity, prototype evaluation aligns innovation with scholarly responsibility and ethical research practice [30].

3.3. Need for a standardized and systematic prototype evaluation framework

Experts highlighted the absence of a unified and systematic framework for evaluating design thinking prototypes, noting that current practices are often improvised and dependent on individual evaluators’ preferences. Participants reported adapting rubrics originally designed for traditional research outputs, resulting in inconsistent judgments and unclear expectations for students. This finding aligns with prior studies emphasizing that standardized evaluation instruments enhance reliability, fairness, and transparency in educational assessment [3], [31]. Without systematic criteria, prototype evaluation becomes susceptible to subjectivity and limited comparability across programs and institutions. The emergence of this domain therefore underscores the need for a structured evaluation framework capable of supporting consistent and defensible assessment practices within teacher education contexts [2].

3.4. Conceptual alignment between design thinking processes and evaluation criteria

Another domain emphasized the need to align evaluation criteria with the core processes of design thinking. Experts observed that prototypes are frequently assessed using outcome-based measures typically associated with action research or experimental studies, which overlook the iterative and exploratory nature of design thinking. Such misalignment can lead to construct underrepresentation and inaccurate judgments of prototype quality.

Evaluation literature stresses that assessment criteria must reflect the theoretical foundations of the constructs being evaluated [14], [32]. In design thinking contexts, this includes recognizing empathy, ideation, prototyping, and iterative refinement as key indicators of quality. Indicators derived from this domain therefore capture process fidelity and developmental progression, ensuring that evaluation practices remain theoretically coherent and methodologically aligned with design thinking–based inquiry [33].

3.5. Feasibility, sustainability, and contextual adaptability of prototypes

Experts further emphasized that prototype evaluation should extend beyond immediate functionality to consider feasibility, sustainability, and contextual adaptability. Prototypes that require excessive resources, specialized expertise, or narrowly defined conditions were viewed as limited in evaluative value, particularly in teacher education contexts where scalability and long-term implementation are critical considerations. This emphasis aligns with evaluation frameworks that prioritize practicality, sustainability, and usability as indicators of quality educational innovations [34], [35]. From an evaluation perspective, feasibility and adaptability enhance both the utility and external validity of prototypes. Consequently, indicators associated with this domain assess resource requirements, implementation constraints, and the potential for adaptation across diverse educational contexts [36].

3.6. Demonstration of research and professional competencies through prototypes

Finally, experts viewed prototypes as integrative evidence of students’ research and professional competencies. Beyond assessing the artifact itself, evaluation should capture the extent to which prototypes reflect analytical reasoning, theoretical grounding, methodological rigor, and professional judgment.

Participants emphasized that prototypes serve as tangible manifestations of students' ability to translate research knowledge into practice.

This domain aligns with competency-based evaluation approaches that conceptualize student outputs as indicators of underlying cognitive and professional skills [37], [38]. By incorporating competency-focused indicators, prototype evaluation extends beyond product assessment to encompass learning outcomes central to teacher education. This strengthens the interpretive value of evaluation results and supports alignment with program-level learning outcomes [39].

3.7. Operationalization of evaluation domains into indicators

Following thematic analysis, the six evaluation domains were systematically operationalized into observable and assessable indicators reflecting core dimensions of design thinking prototype quality. Indicator development was directly grounded in expert discourse and refined through iterative review to ensure conceptual clarity, non-redundancy, and alignment with established principles of educational evaluation. By translating qualitative expert insights into structured evaluative criteria, this process provides initial evidence of content validity, as the resulting indicators represent consensus-based judgments regarding appropriate standards for assessing design thinking prototypes.

Beyond content relevance, the explicit alignment of indicators with theoretically defined evaluation domains strengthens construct coherence by ensuring consistency between the underlying constructs and their operational representation. Unlike conventional prototype assessments that emphasize outcome effectiveness alone, the proposed framework integrates problem alignment, originality safeguards, process fidelity, feasibility considerations, and competency demonstration within a unified evaluative structure. This integrated operationalization addresses the methodological gap identified in the literature, where design thinking outputs are frequently assessed using criteria misaligned with their iterative and problem-driven nature.

To synthesize the progression from expert judgment to measurable evaluation criteria, each domain was mapped to its core evaluative focus and representative indicators. Table 1 presents this synthesis, illustrating how expert-derived constructs were transformed into an initial indicator pool designed to support subsequent content validation, pilot testing, and psychometric analysis. As summarized in Table 1, the domains capture complementary yet distinct dimensions of prototype quality, collectively forming a coherent and theoretically grounded framework for the systematic assessment of design thinking outputs in teacher education. This structured operationalization strengthens transparency, defensibility, and consistency in prototype evaluation while establishing a robust foundation for the development of a standardized assessment instrument.

Table 1. Expert-derived evaluation domains, core evaluative focus, and sample indicators for design thinking prototype assessment

Evaluation domain	Core evaluative focus	Sample evaluation indicators
Problem alignment and functional relevance	Alignment between identified educational problem and prototype solution; responsiveness to authentic instructional or institutional needs	The prototype addresses a clearly articulated and research-grounded educational problem.
Novelty, originality, and intellectual property integrity	Evidence of innovation while ensuring ethical differentiation from existing tools or solutions	The prototype demonstrates originality beyond existing or commonly used solutions.
Standardization of evaluation practices	Need for consistent, transparent, and defensible evaluation criteria across evaluators and contexts	The prototype can be assessed using a shared rubric or framework.
Alignment with design thinking processes	Consistency between evaluation criteria and the iterative, process-oriented nature of design thinking	The prototype reflects key design thinking stages (e.g., empathy, ideation, prototyping, testing).
Feasibility, sustainability, and contextual adaptability	Practicality of implementation, long-term use, and adaptability to different educational contexts	The prototype can be implemented using available institutional resources.
Demonstration of research and professional competencies	Extent to which the prototype reflects research rigor and professional judgment	The prototype demonstrates analytical reasoning and theoretical grounding.

4. CONCLUSION

This study developed an expert-informed evaluation framework for assessing design thinking prototypes in teacher education, addressing the absence of systematic and theory-aligned criteria for evaluating design-oriented research outputs. Through rigorous thematic analysis of FGD with research advisors and panel members, six interrelated evaluation domains were identified and translated into observable indicators representing key dimensions of prototype quality, including relevance, originality, process fidelity, feasibility, sustainability, and competency demonstration. The systematic progression from expert judgment to operationalized criteria provides initial evidence of content validity and establishes

a coherent structure for aligning evaluation practices with the problem-driven and iterative nature of design thinking. By offering a transparent and defensible basis for prototype assessment, the proposed framework contributes to reducing subjectivity in evaluation and strengthening consistency across teacher education programs, while supporting the assessment of prototypes as both research outputs and professional learning artifacts. Methodologically, this study contributes to evaluation science by demonstrating how expert qualitative judgment can be systematically translated into defensible evaluation criteria for design-oriented research outputs. Future research should empirically validate the proposed indicators through formal content validation, pilot implementation, and psychometric analysis across diverse institutional contexts to further establish reliability, construct validity, and broader applicability.

ACKNOWLEDGMENTS

The authors sincerely thank the research advisors and panel members who participated in the FGD and shared their expertise and evaluative insights, which were essential to the development of the evaluation domains and indicators. The authors also acknowledge Northern Bukidnon State College for its institutional support in the conduct of this research.

FUNDING INFORMATION

This research was funded by Northern Bukidnon State College (NBSC) as part of its institutional research and development initiatives. No external grant funding was involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mary Cris J. Go	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Jovelyn G. Delosa	✓									✓				✓
Christine C. Royo	✓									✓			✓	✓

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare no competing financial or personal interests that could have influenced this study. The funding institution had no role in the study design, data collection, analysis, interpretation, manuscript preparation, or decision to publish.

INFORMED CONSENT

Informed consent was obtained from all participants prior to data collection. Participants were informed of the study's purpose, procedures, voluntary nature, and their right to withdraw without penalty, including consent for audio recording and the use of anonymized data for research and publication.

ETHICAL APPROVAL

The study was conducted in accordance with national regulations and institutional policies and in compliance with the Declaration of Helsinki. Ethical approval was obtained from the Research Ethics Committee of Davao De Oro State College prior to data collection.

DATA AVAILABILITY

The data supporting the findings of this study are available from the corresponding author, [MCJG], upon reasonable request. Due to confidentiality and privacy considerations, the datasets are not publicly available; however, derived data are included within the article.





REFERENCES

- [1] D. Vlachopoulos and A. Makri, "A systematic literature review on authentic assessment in higher education: best practices for the development of 21st century skills, and policy considerations," *Studies in Educational Evaluation*, vol. 83, p. 101425, Dec. 2024, doi: 10.1016/j.stueduc.2024.101425.
- [2] E. Huber *et al.*, "Towards a framework for designing and evaluating online assessments in business education," *Assessment & Evaluation in Higher Education*, vol. 49, no. 1, pp. 102–116, Jan. 2024, doi: 10.1080/02602938.2023.2183487.
- [3] G. T. L. Brown, "The past, present and future of educational assessment: a transdisciplinary perspective," *Frontiers in Education*, vol. 7, p. 1060633, Nov. 2022, doi: 10.3389/educ.2022.1060633.
- [4] S. McKenney and T. Reeves, *Conducting educational design research*, London: Routledge, 2020, doi: 10.4324/9780203818183.
- [5] M. Z. Kamarudin and M. S. A. M. Noor, "What do we know about the selection of action research methodologies in primary science education? – A systematic literature review," *Educational Action Research*, vol. 32, no. 5, pp. 825–847, Oct. 2024, doi: 10.1080/09650792.2023.2261502.
- [6] J. Liedtka, "Putting technology in its place: design thinking's social technology at work," *California Management Review*, vol. 62, no. 2, pp. 53–83, Feb. 2020, doi: 10.1177/0008125619897391.
- [7] Z. Yan and D. Carless, "Self-assessment is about more than self: the enabling role of feedback literacy," *Assessment & Evaluation in Higher Education*, vol. 47, no. 7, pp. 1116–1128, Oct. 2022, doi: 10.1080/02602938.2021.2001431.
- [8] M. Wakid, H. Sofyan, A. Widowati, and A. Z. Ilma, "Learning-oriented assessment: a systematic literature network analysis," *Cogent Education*, vol. 11, no. 1, pp. 1–12, Dec. 2024, doi: 10.1080/2331186X.2024.2366075.
- [9] G. O. Boateng, T. B. Neilands, E. A. Frongillo, H. R. Melgar-Quinonez, and S. L. Young, "Best practices for developing and validating scales for health, social, and behavioral research: a primer," *Frontiers in Public Health*, vol. 6, p. 149, Jun. 2018, doi: 10.3389/fpubh.2018.00149.
- [10] M. S. B. Yusoff, "ABC of content validation and content validity index calculation," *Education in Medicine Journal*, vol. 11, no. 2, pp. 49–54, Jun. 2019, doi: 10.21315/eimj2019.11.2.6.
- [11] T. O. Nyumba, K. Wilson, C. J. Derrick, and N. Mukherjee, "The use of focus group discussion methodology: insights from two decades of application in conservation," *Methods in Ecology and Evolution*, vol. 9, no. 1, pp. 20–32, Jan. 2018, doi: 10.1111/2041-210X.12860.
- [12] J. W. Creswell and V. L. P. Clark, *Designing and conducting mixed methods research*, 3rd ed. Thousand Oaks, CA: SAGE Publications, Inc., 2018.
- [13] A. Shorten and J. Smith, "Mixed methods research: expanding the evidence base," *Evidence Based Nursing*, vol. 20, no. 3, pp. 74–75, Jul. 2017, doi: 10.1136/eb-2017-102699.
- [14] M. Kane, "Validating score interpretations and uses," *Language Testing*, vol. 29, no. 1, pp. 3–17, Jan. 2012, doi: 10.1177/0265532211417210.
- [15] T. A. Kyriazos, "Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general," *Psychology*, vol. 9, no. 8, pp. 2207–2230, 2018, doi: 10.4236/psych.2018.98126.
- [16] E. Fokides, "Development and testing of a scale for examining factors affecting the learning experience in the Metaverse," *Computers & Education: X Reality*, vol. 2, p. 100025, 2023, doi: 10.1016/j.cexr.2023.100025.
- [17] A. Pechard *et al.*, "Development and validation of a scale to assess students' perceptions of team feedback in collaborative tasks," *Assessment & Evaluation in Higher Education*, vol. 51, no. 2, pp. 379–397, Feb. 2026, doi: 10.1080/02602938.2025.2536566.
- [18] V. Braun and V. Clarke, "One size fits all? What counts as quality practice in (reflexive) thematic analysis?" *Qualitative Research in Psychology*, vol. 18, no. 3, pp. 328–352, Jul. 2021, doi: 10.1080/14780887.2020.1769238.
- [19] L. S. Nowell, J. M. Norris, D. E. White, and N. J. Moules, "Thematic analysis: striving to meet the trustworthiness criteria," *International Journal of Qualitative Methods*, vol. 16, no. 1, pp. 1–13, Dec. 2017, doi: 10.1177/1609406917733847.
- [20] V. Braun and V. Clarke, "Reflecting on reflexive thematic analysis," *Qualitative Research in Sport, Exercise and Health*, vol. 11, no. 4, pp. 589–597, Aug. 2019, doi: 10.1080/2159676X.2019.1628806.
- [21] G. Konstantinos, "Thematic analysis: a practical guide by Virginia Braun and Victoria Clarke, Los Angeles, USA, Sage Publications, 2022," *European Journal of Psychotherapy & Counselling*, vol. 26, no. 3–4, pp. 461–464, Oct. 2024, doi: 10.1080/13642537.2024.2391666.
- [22] L. Birt, S. Scott, D. Cavers, C. Campbell, and F. Walter, "Member checking," *Qualitative Health Research*, vol. 26, no. 13, pp. 1802–1811, Nov. 2016, doi: 10.1177/1049732316654870.
- [23] C. O'Connor and H. Joffe, "Intercoder reliability in qualitative research: debates and practical guidelines," *International Journal of Qualitative Methods*, vol. 19, pp. 1–13, Jan. 2020, doi: 10.1177/1609406919899220.
- [24] World Medical Association, "World medical association Declaration of Helsinki," *JAMA*, vol. 310, no. 20, pp. 2191–2194, Nov. 2013, doi: 10.1001/jama.2013.281053.
- [25] C. Rapanta, L. Botturi, P. Goodyear, L. Guàrdia, and M. Koole, "Balancing technology, pedagogy and the new normal: post-pandemic challenges for higher education," *Postdigital Science and Education*, vol. 3, no. 3, pp. 715–742, Oct. 2021, doi: 10.1007/s42438-021-00249-1.
- [26] S. McKenney and T. C. Reeves, "Educational design research: portraying, conducting, and enhancing productive scholarship," *Medical Education*, vol. 55, no. 1, pp. 82–92, Jan. 2021, doi: 10.1111/medu.14280.
- [27] H. S. You, S. Park, J. A. Marshall, and C. Delgado, "Interdisciplinary science assessment of carbon cycling: construct validity evidence based on internal structure," *Research in Science Education*, vol. 52, no. 2, pp. 473–492, Apr. 2022, doi: 10.1007/s11165-020-09943-9.
- [28] T. Bretag *et al.*, "Contract cheating and assessment design: exploring the relationship," *Assessment & Evaluation in Higher Education*, vol. 44, no. 5, pp. 676–691, Jul. 2019, doi: 10.1080/02602938.2018.1527892.
- [29] O. L. Holden, M. E. Norris, and V. A. Kuhlmeier, "Academic integrity in online assessment: a research review," *Frontiers in Education*, vol. 6, p. 639814, Jul. 2021, doi: 10.3389/educ.2021.639814.
- [30] R. Mulenga and H. Shilongo, "Academic integrity in higher education: understanding and addressing plagiarism," *Acta Pedagogica Asiana*, vol. 3, no. 1, pp. 30–43, Jan. 2024, doi: 10.53623/apga.v3i1.337.





- [31] J. Tai, R. Ajjawi, D. Boud, P. Dawson, and E. Panadero, “Developing evaluative judgement: enabling students to make decisions about the quality of work,” *Higher Education*, vol. 76, no. 3, pp. 467–481, Sep. 2018, doi: 10.1007/s10734-017-0220-3.
- [32] D. Carless and D. Boud, “The development of student feedback literacy: enabling uptake of feedback,” *Assessment & Evaluation in Higher Education*, vol. 43, no. 8, pp. 1315–1325, Nov. 2018, doi: 10.1080/02602938.2018.1463354.
- [33] Y. Gao, “Deep learning-based strategies for evaluating and enhancing university teaching quality,” *Computers and Education: Artificial Intelligence*, vol. 8, p. 100362, Jun. 2025, doi: 10.1016/j.caeai.2025.100362.
- [34] J. van der Linden, C. van der Vleuten, L. Nieuwenhuis, and T. van Schilt-Mol, “Formative use of assessment to foster self-regulated learning: the alignment of teachers’ conceptions and classroom assessment practices,” *Journal of Formative Design in Learning*, vol. 7, no. 2, pp. 195–207, Nov. 2023, doi: 10.1007/s41686-023-00082-8.
- [35] A. Hargreaves, “What the COVID-19 pandemic has taught us about teachers and teaching,” *Facets*, vol. 6, no. 1, pp. 1835–1863, Jan. 2021, doi: 10.1139/facets-2021-0084.
- [36] C. Dell’Era, S. Magistretti, C. Cautela, R. Verganti, and F. Zurlo, “Four kinds of design thinking: from ideating to making, engaging, and criticizing,” *Creativity and Innovation Management*, vol. 29, no. 2, pp. 324–344, Jun. 2020, doi: 10.1111/caim.12353.
- [37] M. Griffith and C. Lechuga-Jimenez, “Design thinking in higher education case studies: disciplinary contrasts between cultural heritage and language and technology,” *Education Sciences*, vol. 14, no. 1, p. 90, Jan. 2024, doi: 10.3390/educsci14010090.
- [38] C. S. Roßnagel, K. L. Baido, and N. Fitzallen, “Revisiting the relationship between constructive alignment and learning approaches: a perceived alignment perspective,” *PLoS ONE*, vol. 16, no. 8, p. e0253949, Aug. 2021, doi: 10.1371/journal.pone.0253949.
- [39] L. Darling-Hammond *et al.*, *Preparing teachers for deeper learning*. Cambridge, MA: Harvard Education Press, 2019.

BIOGRAPHIES OF AUTHORS







Mary Cris J. Go     is an instructor at Northern Bukidnon State College, Philippines. She holds a Ph.D. in Mathematics Education and has taught mathematics and professional education courses at the secondary and tertiary levels. Her research interests include mathematics education, teacher education, educational technology, artificial intelligence in education, assessment and evaluation, and qualitative and mixed-methods research. She can be contacted at email: mjgo@nbsc.edu.ph.



Jovelyn G. Delosa     is an associate professor and the vice president for Academic Affairs of Northern Bukidnon State College, Philippines. She holds a Ph.D. in Education from Xavier University–Ateneo de Cagayan and a Master of Arts in Education from the University of Adelaide, Australia, with specialization in measurement, assessment, and research methods. She has extensive experience in teacher education, educational leadership, assessment and evaluation, curriculum development, and research supervision at both undergraduate and graduate levels. Her scholarly work focuses on educational measurement, research self-efficacy, assessment literacy, and evaluation practices in teacher education. She can be contacted at email: jgdelosa@nbsc.edu.ph.



Christine C. Royo     is a senior faculty member and academic administrator at Northern Bukidnon State College, Philippines. She holds a Ph.D. in Science Education with specialization in Chemistry from the University of Science and Technology in the Southern Philippines. She has extensive experience in teacher education, research supervision, academic leadership, and quality assurance, having served in various instructional and administrative roles in higher education. Her research interests include science education, teacher education, educational evaluation, institutional effectiveness, and assessment practices in higher education. She has actively participated in national and international research conferences and institutional research initiatives. She can be contacted at email: ccroyo@nbsc.edu.ph.