INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 9

Authentic Task Assessment for Grade 10 Biology: A Mixed-Methods Analysis of Formative and Summative Performance on Feedback Mechanisms and Homeostasis

DOI: 10.5281/zenodo.22078097

SEGUNDINO D. BALBUENA
segundino.balbuena@deped.gov.ph
Cebu Technological University - Main Campus, R. Palma St., Cebu City
Department of Education – Mandaue City Division

Abstract:

This study determined the significance of assessment in evaluating Grade 10 Biology performance on Organisms as Having Feedback Mechanisms at Cabancalan National High School, Mandaue City, for the third quarter SY 2024-2025. Using a mixed-methods descriptive-comparative design, it analyzed formative and summative scores of 175 students across three MELCs: (1) explaining the role of hormones in female and male reproductive systems, (2) describing feedback mechanisms in the menstrual cycle, and (3) describing nervous system coordination for homeostasis; and explored teachers' lived experiences through thematic analysis of four biology teachers. Formative results showed 26.86% mastery for hormones, 11.43% for menstrual feedback (lowest, M=2.92, SD=1.47), and 31.43% for nervous system/homeostasis (highest formative mean M=3.78). Summative performance improved significantly: mastery rose to 44.57% (+17.71%), 42.86% (+31.43%), and 66.29% (+34.86%) respectively, with means of 3.99, 3.72, and 4.79. Paired t-tests revealed significant gains in all competencies ($p<.001$): hormones MD=0.611, d=0.379; menstrual feedback MD=0.800, d=0.418; nervous system/homeostasis MD=1.011, d=0.654, the largest and only medium-to-large effect. Qualitative analysis yielded three themes: (a) Barriers to Science Learning—abstract terminology, sociocultural sensitivity on reproductive health, and prior knowledge gaps; (b) Effective Strategies—visual scaffolds, flowcharts, comic strips, storytelling, real-life contextualization, and collaborative discourse; and (c) Time Constraints and Learning Pace—limited instructional time hindering deep formative assessment. Findings support integrating authentic, student-centered tasks to bridge abstract hormonal feedback to lived experience. Proposed output includes three authentic assessment tasks with rubrics.

Keywords: authentic task assessment, formative assessment, summative assessment, hormonal regulation, feedback mechanisms, feedback loops, nervous system, thematic analysis

Introduction:

Biology education in the 21st century transcends mere factual recall; it demands that learners apply concepts to authentic real-world problems, demonstrate scientific reasoning, and cultivate employability skills such as communication, collaboration, and adaptability. Authentic assessment, defined as tasks that mirror the complexity of real-life challenges, offers a promising avenue to achieve this transformation (Sambell et al., 2023; Villaroel et al., 2023).

Unlike traditional multiple-choice tests that continue to dominate Philippine public schools due to large class sizes, limited materials, and insufficient teacher training (Sarmiento et al., 2020), authentic tasks encourage creativity, critical thinking, and personal responsibility as students set their own targets, draft, revise, and manage resources (Brown & Sambell, 2023; McArthur, 2021). This shift from assessment of learning to assessment for and as learning is essential for developing scientific literacy and workforce readiness.

The COVID-19 pandemic, however, exposed critical vulnerabilities in the implementation of such student-centered science education. Laboratory work was severely curtailed, student engagement declined, and unequal access to technology widened existing learning gaps (Bird, 2021; Khajeloo et al., 2021). In the Philippines, the Department of Education responded through DepEd Order No. 31, s. 2020, which provided interim guidelines for flexible and holistic assessment under the Basic Education Learning Continuity Plan.

While the policy emphasized varied formative and summative strategies, Petalla (2021) observed that many performance tasks in Grade 10 Biology remained disconnected from real-world contexts and still emphasized memorization over application. This disconnection is particularly acute in the third-quarter topic Organisms as Having Feedback Mechanisms, which requires understanding abstract, multisystem interactions: endocrine signaling across age and gender, positive and negative feedback loops regulating the menstrual cycle, and nervous system coordination to maintain homeostasis.

These three Most Essential Learning Competencies (MELCs) are conceptually demanding and culturally sensitive. Hormonal regulation involves invisible chemical messengers — FSH, LH, estrogen, progesterone, and testosterone — whose fluctuating levels are difficult to visualize, leading students to perceive the topic as abstract and overwhelming (Almarghani & Mijatović, 2021; Lee & Salter, 2020). The feedback mechanisms of the menstrual cycle constitute a threshold concept that is troublesome due to its dynamic, non-linear causality and is often underrepresented in teaching materials because of sociocultural and religious taboos, resulting in low self-efficacy and persistent misconceptions (Navarro & Torres, 2022; Yilmaz & Kara, 2021).

As Juma and Ellis (2021) argue, unless reproductive health is taught through student-centered, trust-building, and gender-responsive discussion, misconceptions are likely to persist. In contrast, nervous system integration and homeostasis can be linked to everyday experiences such as thermoregulation and stress responses, making it relatively more accessible when taught through systems-based approaches that activate prior knowledge (Singh & Kaur, 2023; Brown & Harris, 2023).

To address these challenges, this study is anchored on three complementary theoretical lenses. Dewey's Experiential Learning Theory posits that knowledge is constructed through "learning by doing" — active engagement, experimentation, and reflection on real-life experiences (Pappas, 2023). Vygotsky's Social Constructivism emphasizes that learning occurs within the Zone of Proximal Development through scaffolding by teachers and peers, making collaborative laboratory work, visual scaffolds, and peer discourse essential. Black and Wiliam's (1998) Assessment for Learning framework reframes assessment as a tool to improve learning rather than merely measure it, through timely, specific feedback and opportunities for revision. Together with DepEd Order No. 31, s. 2020 as legal basis, these perspectives support the integration of formative, summative, and authentic assessments in Grade 10 Biology.

Despite a growing body of literature on authentic and formative assessment including foreign studies on feedback loops (Lee & Salter, 2020), Philippine studies on STEAM authentic tasks (Garcia, 2025; Sarmiento et al., 2020), and Cebu-based evidence that science teachers value assessment for diagnosing learning (Tesorio & Canizares, 2018) there remains a clear gap. No study has directly compared the formative and summative performance of Grade 10 students on the three feedback mechanism MELCs at Cabancalan National High School, Mandaue City, while simultaneously triangulating these quantitative gains with teachers' lived experiences in designing and evaluating authentic tasks. Therefore, this study aimed to: (1) determine the distribution of students' scores in formative assessment, (2) determine the distribution in summative assessment, (3) test the significant difference between the two, (4) explore teachers' experiences in assessing Grade 10 Biology, and (5) propose a validated authentic task assessment package as output.

Literature Review:

Authentic assessment has emerged as a transformative approach in 21st-century biology education. Sambell, Brown, and McDowell (2023) defined authentic assessments as tasks that mirror real-world problems, requiring students to set their own targets, draft, revise, and manage resources, thereby fostering independence and responsibility. Brown and Sambell (2023) emphasized that such tasks develop time management and self-regulation, while McArthur (2021) directly linked authentic assessment to employability, arguing that scenario-based tasks allow students to practice communication, collaboration, and adaptability in contexts that simulate future careers. Villaroel, Bloxham, and Bruna (2023), Karunanayaka et al. (2020), and Garcia (2025) further contended that authentic assessment promotes equity and inclusion, as it enables students from diverse socioeconomic and cultural backgrounds to demonstrate competence through varied modalities — poems, comic strips, brochures, models, role-playing, case analyses, infographics, and mind maps — rather than through a single standardized test. In biology, this multiplicity is crucial because scientific literacy requires observation, questioning, evidence interpretation, and communication of conclusions.

In the Philippine K-12 context, however, the ideal of authentic assessment remains largely aspirational. Petalla (2021) observed that despite curriculum reforms emphasizing performance-based assessment, many biology tasks in the Philippines still emphasize memorization and remain disconnected from real-world practice. Sarmiento, Cruz, and Lim (2020), in an analysis of 106 STEAM teachers across 14 Philippine regions, found a hybrid pattern: teachers used both authentic and traditional multiple-choice assessments, driven by systemic constraints such as large enrolments, lack of instructional materials, and insufficient training in designing reliable rubrics. This persistence of conventional testing is reinforced by a culture of high-stakes assessment that, as the National Education Association (2023) noted, often fails to capture reasoning, problem-solving, and creative thinking. The result is a gap between policy intent and classroom reality.

The theoretical foundation for bridging this gap rests on three complementary perspectives. First, Dewey's Experiential Learning Theory, as reviewed by Pappas (2023) and the Gilliam Writers Group (2023), posits "learning by doing" — knowledge is constructed through active engagement, experimentation, and reflection on lived experience. In Grade 10 Biology, this translates to laboratory investigations, case studies, and health-related analyses where students observe phenomena, formulate questions, and relate findings to ecology, health, and environment. Second, Vygotsky's Social Constructivism emphasizes that learning is socially mediated within the Zone of Proximal Development (ZPD). As McLeod (2024) and Simply Psychology (2024) summarized, students accomplish more challenging tasks when scaffolded by teachers and peers through collaborative laboratory work, group analysis, and peer discussion. Third, Black and Wiliam's (1998) Assessment for Learning framework, further developed by Wiliam (2011), reframes assessment as a means to improve learning rather than merely measure achievement. Formative assessment, through questioning, observation, self- and peer-assessment, and timely, specific, and actionable feedback, provides evidence that teachers use to adjust instruction. This framework is legally supported by DepEd Order No. 31, s. 2020, which mandates varied, holistic, and formative-summative integrated assessment practices.

Empirical evidence strongly supports formative assessment in biology. Choudhary and Bano (2022), in a six-week quasi-experimental study with 46 secondary students, demonstrated that concept maps as a formative tool improved conceptual understanding, reduced dependence on rote learning, and promoted critical thinking and reasoning. Foster (2024), in a meta-analytical review, found that formative assessment had significant positive effects on student autonomy and performance in biology. Garcia (2025), studying 102 Grade 11 Filipino STEAM students, reported that varied authentic tasks enhanced scientific literacy, cognitive development, self-expression, and generally elicited favorable perceptions among learners. Andyhapsari and Djukri (2021) showed that blended learning strategies supplemented by Google Classroom improved cognitive, creative, and self-regulated learning in a digital environment. Mulyani and Doyan

(2023) found that contextual teaching, which mixes biology content with real-life situations, increased scientific literacy and the perceived relevance of biology to global issues. These findings collectively suggest that when formative assessment is intentional and contextualized, it moves students from partial to full mastery.

The specific content of this study — hormonal regulation, feedback mechanisms in the menstrual cycle, and nervous system coordination for homeostasis — presents unique cognitive and sociocultural barriers. Almarghani and Mijatović (2021) described hormonal physiology as inherently abstract because hormones are invisible chemical messengers with Latinate terminology and fluctuating levels. Students struggle to differentiate FSH, LH, estrogen, progesterone, and testosterone and to explain their age- and gender-specific roles. Lee and Salter (2020) identified feedback loops as a threshold concept: once understood, they transform understanding of physiology, but they are troublesome due to dynamic, non-linear causality and the need to distinguish positive from negative feedback. Villanueva and Chan (2023) advocated for visualizing feedback loops through histograms, flowcharts, and scaffolded systems mapping to make abstract interactions visible.

Beyond cognitive complexity, reproductive health topics face sociocultural barriers. Navarro and Torres (2022) documented that menstrual physiology is often underrepresented or euphemized in teaching materials due to sociocultural and religious sensitivities, leading to low self-efficacy, particularly among female students. Yilmaz and Kara (2021) found that cultural misconceptions about reproduction interfere with learning and require targeted interventions. Juma and Ellis (2021) argued that reproductive health education must be student-centered, building discussion and trust to directly address misconceptions. Batool et al. (2021) added that biology is regarded as more complex than other subjects due to dense terminology, and students with weak reading comprehension struggle even with basic concepts. These barriers were exacerbated by the COVID-19 pandemic, during which biology teaching became heavily reliant on technology and lost hands-on laboratory opportunities (Tijap & Pelgone, 2022; Khajeloo et al., 2021; Bird, 2021).

In contrast, nervous system coordination and homeostasis appear relatively more accessible. Singh and Kaur (2023) explained that systems-based teaching, where biological concepts feed into one another to show interdependence, improves outcomes because students activate prior knowledge from earlier grades. Fernandez and Liu (2024) noted that prior knowledge plus interlevel strengthening increases success in solving multi-level problems. Brown and Harris (2023) recommended contextualizing homeostasis with everyday examples such as thermoregulation and stress responses. However, Antipolo and Rogayan (2021) cautioned that while spiral progression in K-12 science reinforces learning, it can also create redundancy and confusion, leading to gaps in understanding if not carefully integrated.

Teachers' strategies and contextual constraints further shape assessment effectiveness. Tesorio and Canizares (2018), in a mixed-methods study of 125 science teachers from 30 private and public schools in western Cebu plus focus group discussion with eight teachers, found that teachers strongly viewed assessment as a means to improve learning and diagnose abilities, not just assign grades. Effective strategies identified in local and international literature include graphics, flowcharts, storytelling, comic strips, real-life examples, and collaborative prompts (Bhardwaj et al., 2025; Abulhul, 2021). Chen and Ramirez (2020) emphasized that assessment design must balance content and cognition, requiring higher-order thinking, problem-solving, and functional application to move students to full mastery.

Yet time constraints and student learning pace remain persistent challenges. UNESCO International Institute for Educational Planning (2023) asserted that adequate instructional time is vital for quality education; deficits in time produce deficits in learning. Petalla and Doromal (2021), interviewing six Filipino senior high school students with low performance-task scores, identified difficulties in task planning, execution, weak preparation, limited supervision, inadequate skills, and unreliable scoring, highlighting the

need for clear instructions, teacher guidance, fair rubrics, and meaningful feedback. Tijap and Pelgone (2022) documented that Philippine science teachers struggled to incorporate new teaching methods during remote learning and needed training in effective biology pedagogy.

Taken together, foreign studies, Philippine studies, Cebu-based evidence, theoretical perspectives, and legal basis indicate that authentic and formative assessments can promote conceptual understanding, scientific literacy, critical thinking, and active participation. However, their effectiveness depends on task design, teacher preparation, resources, technology access, supervision, and reliable scoring. The reviewed literature has not directly compared formative and summative performance of Grade 10 Biology students at Cabancalan National High School on the MELCs covering feedback mechanisms and homeostasis while simultaneously examining teachers' perceptions of authentic task design and evaluation. This gap provides the basis for the present study.

Methodology:

Research Design

A mixed-methods design was used to gain deeper insights into student performance on the Grade 10 biology assessment tasks. A combination of quantitative and qualitative approaches allowed the study to present a more comprehensive view of both student learning outcomes and teachers' assessment experiences, utilizing a mix of perspectives to judge the efficacy of different assessment approaches.

The study included a descriptive-comparative design (for the quantitative component), statistical direction: the mean, the standard deviation, frequency, and percentage of formative and summative assessment scores. Using a paired-sample t-test (or Wilcoxon signed-rank test for non-normally distributed data), findings indicated whether a significant difference exists between students' formative and summative scores, representing learning gains and instructional effectiveness. Teachers' reflections on their assessment experiences of Grade 10 biology students are analyzed thematically to gain insights into student learning, instructional challenges and effective practices.

Flow of the Study

The study is conducted systematically following the input, process, and output model as the focus of the study. Data are being collected on students' scores in formative assessments of the third-quarter competencies in biology. These include explaining the functions of endocrine glands in males and females, describing feedback mechanisms in the menstrual cycle, and explaining how the nervous system coordinates these actions for homeostasis. Also, students' scores in summative assessments within the same scope of competencies, as well as the obstacles and difficulties in evaluating their performance, are considered.

The procedure has several main steps, such as acquiring authorization from the school principal through a transmittal letter. Following authorization, data gathering is done, and then analysis and interpretation of the information using relevant statistical tools are conducted.

In any case, the study works toward an output by creating an authentic task assessment, which is int
to meet the competencies to which students can meaningfully relate in biology.

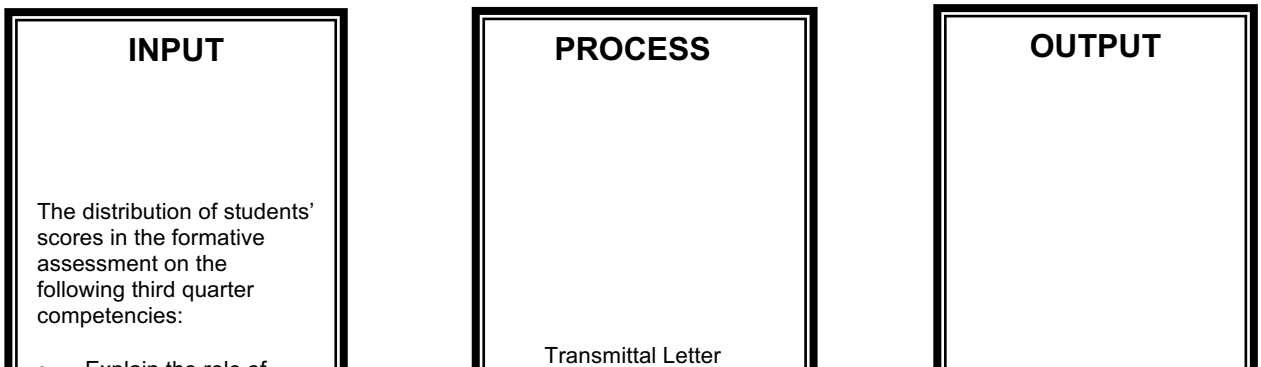


Figure 2. Flow of the Study

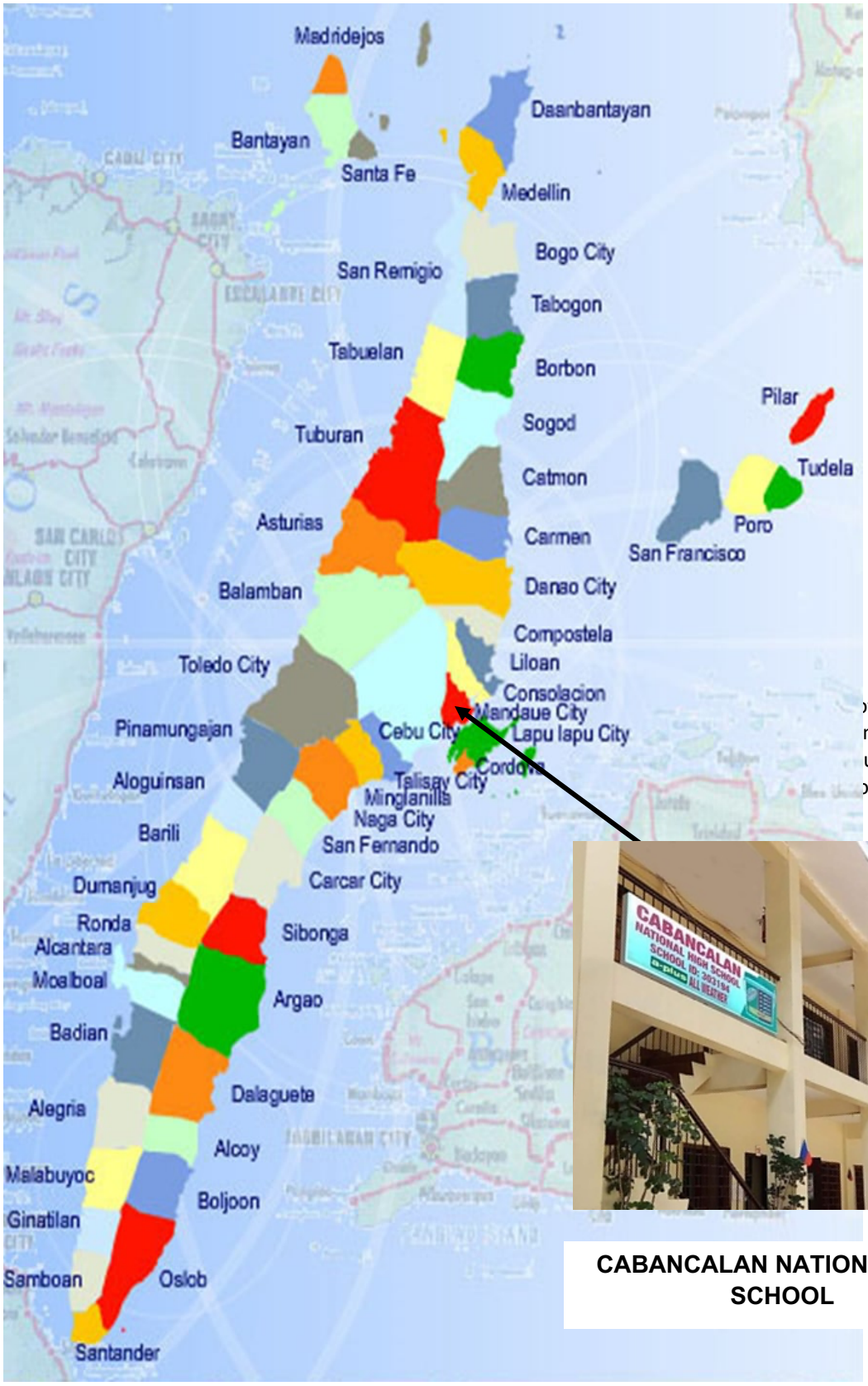
Research Environment

The study was carried out at Cabancalan National High School. It is situated along Manuel Quezon Street, in Cabancalan, Mandaue City, Cebu, under the Mandaue City Division. The school opened on June 1, 1985, and is currently considered one of the best-performing schools in the City of Mandaue. It has a Junior High and Senior High School curriculum. The school comprises one school head, one assistant school head, 107 teaching staff, and six non-teaching staff.

Most of the students for each section are at multiple grade levels, with 12 sections per grade. Twice-per-day classes are divided into two periods, and one of the twelve sections is a special science class. The Junior High School is composed of 1,795 students and has a total enrollment of. In addition, there are 326 students in Senior High School, Grade 11 with five sections and Grade 12 with three sections, as of school year 2024-2025.

In addition, the General Academic Strand, Humanities and Social Sciences Strand, and the Technical Vocational are in the school's Senior High Curriculum. The school also has a working library and three computer laboratories.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 9



CABANCALAN NATIONAL HIGH SCHOOL

otal of 175
nts served
urposively
ogy classes

The sample consisted of 175 students in four sections of Grade 10 and four Grade 10 teachers from Cabancalan National High School. Participation in this study was voluntary; however, all participants were made aware of the purpose of the study prior to providing consent. Student and guardian consent was obtained, as well as teacher consent. Throughout the study, subjects' confidentiality and anonymity were maintained, and the data collected will only be used for academic and research purposes. This research conformed to ethical guidelines and ensured that all participants' rights and well-being were respected.

Research Instrument

The research used formative and summative tests as tools. The tests were based on the validated Self-Learning Modules of the Department of Education (DepEd), Division of Mandaue City. The formative tests were centered on aiding student understanding and progress throughout the instructional period, while summative assessments measured achievement after instruction. The validated modules provided within DepEd's scope ensured the assessment tools were appropriate, authoritative, and coherent with educational standards set by DepEd. In addition, teachers were asked to reflect on their live instructional experiences through guided questions.

Data Gathering Procedure

An official letter requesting permission to undertake the study was sent to the principal of Cabancalan National High School. After securing consent, the formative assessment was conducted over three weeks, with each week targeting particular competencies. The assessment materials were handed out to the students every week.

After the instruction was completed, students took a summative assessment aimed at testing the extent of knowledge gained, after which all data were collected and handed over to the statistician for examination. Teachers were asked to share their challenges, strategies, and insights while assessing students' understanding of complex biological concepts. These experiences provided a qualitative perspective that complemented the quantitative results from student assessments.

Thereafter, the interpretable results were made, reviewed, and presented with the research adviser guiding the process, which was subsequently scrutinized alongside the draft of the study and was polished, put in order, and checked for errors and requested edits.

Statistical Treatment of Data

A Simple Percentage. This approach was quite helpful in interpreting the distribution of students' scores in both formative and summative assessments. This statistical tool provided insightful information regarding the proportion of students attaining given score ranges in the assessments.

Mean and Standard Deviation. On the other hand, the mean was used to compute the average score of the students in both formative and summative assessments, while the standard deviation indicated the measure of dispersion of the scores, which also indicated how consistently the students performed.

Paired Sample T-Test/Wilcoxon Signed-Rank Test. The paired-sample t-test was used to decide whether there was a significant difference between students' formative and summative assessment scores using paired-sample t-tests on the assumption of normal data distribution, using a p-value of 0.05.

Thematic Analysis. With such a wide scope of the teachers' experiences when assessing students' performance in Grade 10 biology, thematic analysis was found to be effective in drawing out qualitative responses, enabling the identification of the most common themes and patterns based on collected data.

After collecting the scores of the students, the researcher intended to compute their performance by applying descriptive aids such as mean, standard deviation, frequency distribution, and expectation, which was to assess the difference in the candidates' performances in formative and summative assessments.

Scoring Procedure

The scoring procedure is effective in evaluating student performance because it assists with their scoring progression. It places students into three groups based on their scores: Mastery Achieved (5-6), Developing Mastery (3-4), and Limited or No Mastery (0-2). With this model, educators monitor performance gaps and successes and stretch support while encouraging student growth.

Table 2
Scoring Guide for Student Assessments

Score Range	Category
5-6	Mastery Achieved
3-4	Developing Mastery
0-2	Limited or No Mastery

Results and Discussion:

Distribution of Students Scores in Formative Assessment

Formative assessment (n=175, max 6 per competency) clustered at Developing Mastery. Competency 1 (Hormones role): Mastery 47 (26.86%), Developing 79 (45.14%), Limited 49 (28.00%), M=3.38 SD=1.52. Competency 2 (Feedback mechanisms menstrual cycle): hardest, Mastery 20 (11.43%), Developing 87 (49.71%), Limited 68 (38.86%), M=2.92 SD=1.47. Competency 3 (Nervous system & homeostasis): Mastery 55 (31.43%), Developing 93 (53.14%), Limited 27 (15.43%), M=3.78 SD=1.16. Data show gap between menstrual feedback and homeostasis (20pp difference in mastery). This aligns with Almarghani & Mijatovic (2021) abstract nature, Lee & Salter (2020) threshold concepts, Navarro & Torres (2022) sociocultural underrepresentation.

Table 1. Distribution of Respondents

Respondent Groups	Frequency	Percentage
Group 1	45	25.71%
Group 2	45	25.71%
Group 3	42	24.0%
Group 4	43	24.57%
Teachers	4	2.23%

Table 2. Scoring Guide for Student Assessments

Score Range	Category
5-6	Mastery Achieved
3-4	Developing Mastery
0-2	Limited or No Mastery

Table 3. Distribution of Students' Scores in the Formative Assessment (n=175)

Competency	Mastery (5-6) f (%)	Developing (3-4) f (%)	Limited (0-2) f (%)	Mean	SD
Hormones role, age, gender	47 (26.86%)	79 (45.14%)	49 (28.00%)	3.38	1.52

Feedback mechanisms menstrual cycle	20 (11.43%)	87 (49.71%)	68 (38.86%)	2.92	1.47
Nervous system & homeostasis	55 (31.43%)	93 (53.14%)	27 (15.43%)	3.78	1.16

Figure 1 below shows formative distribution visual. The menstrual feedback competency has the smallest blue segment (Mastery).

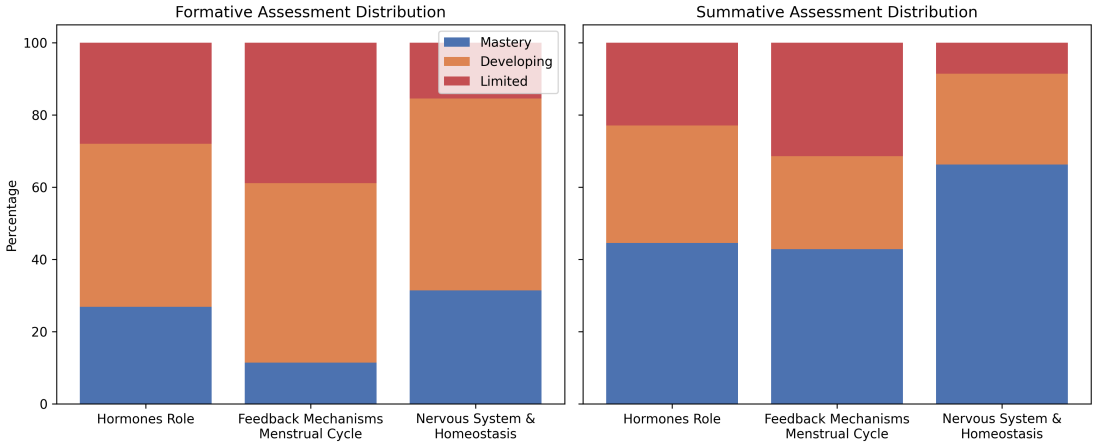


Figure 1. Stacked percentage distribution of mastery categories in formative (left) vs summative (right) assessments.

Distribution of Students Scores in Summative Assessment

After three weeks of instruction with formative feedback, summative mastery improved: Competency 1: Mastery 78 (44.57%, +17.71), Developing 57 (32.57%), Limited 40 (22.86%), M=3.99 SD=1.62. Competency 2: Mastery 75 (42.86%, +31.43), Developing 45 (25.71%), Limited 55 (31.43%), M=3.72 SD=2.01, still weakest with highest limited and increased SD (2.01) indicating widening dispersion. Competency 3: Mastery 116 (66.29%, +34.86), Developing 44 (25.14%), Limited 15 (8.57%), M=4.79 SD=1.33, strongest. Improvement aligns with Foster (2024) on formative assessment impact and Brown & Harris (2023) on real-world contextualization.

Table 4. Distribution of Students' Scores in the Summative Assessment (n=175)

Competency	Mastery (5-6) f (%)	Developing (3-4) f (%)	Limited (0-2) f (%)	Mean	SD
Hormones role, age, gender	78 (44.57%)	57 (32.57%)	40 (22.86%)	3.99	1.62
Feedback mechanisms menstrual cycle	75 (42.86%)	45 (25.71%)	55 (31.43%)	3.72	2.01
Nervous system & homeostasis	116 (66.29%)	44 (25.14%)	15 (8.57%)	4.79	1.33

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 9

Figure 2 highlights mastery jump between assessments.

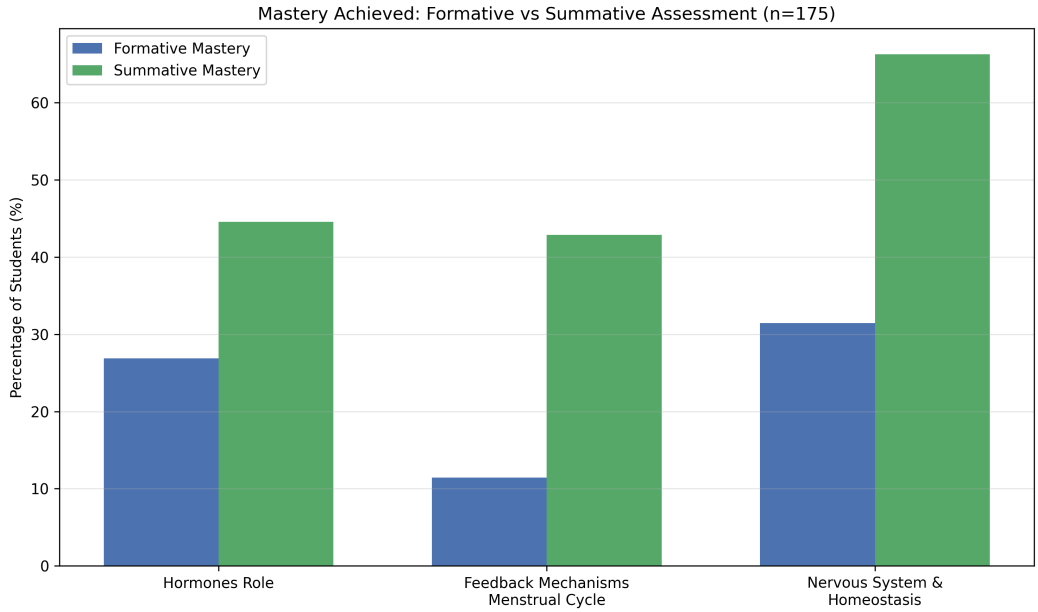


Figure 2. Mastery Achieved: Formative vs Summative Assessment (n=175). Green bar (summative) consistently higher than blue (formative), largest gap for nervous system & homeostasis.

Significant Difference Between Formative and Summative Scores

Paired analysis showed significant gains in all competencies (p<.001). MD hormones=0.611, SE=0.122, d=0.379 small-to-medium; MD menstrual feedback=0.800, SE=0.145, d=0.418 small-to-medium; MD nervous/homeostasis=1.011, SE=0.117, d=0.654 medium-to-large. All null hypotheses rejected at .05. Largest effect for homeostasis indicates educationally substantial improvement. Consistent with Singh & Kaur (2023) systems-based prior knowledge and Chen & Ramirez (2020) formative assessment catching misconceptions early. Menstrual feedback remains abstract requiring creative implementation (Villanueva & Chan, 2023).

Table 5. Paired Differences Between Formative and Summative Scores

Competency	Mean Difference	SE	p-value	Cohen's d	Remarks
Hormones role	0.611	0.122	<.001*	0.379	Significant, small-medium
Feedback mechanisms menstrual cycle	0.800	0.145	<.001*	0.418	Significant, small-medium
Nervous system & homeostasis	1.011	0.117	<.001*	0.654	Significant, medium-large

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 9

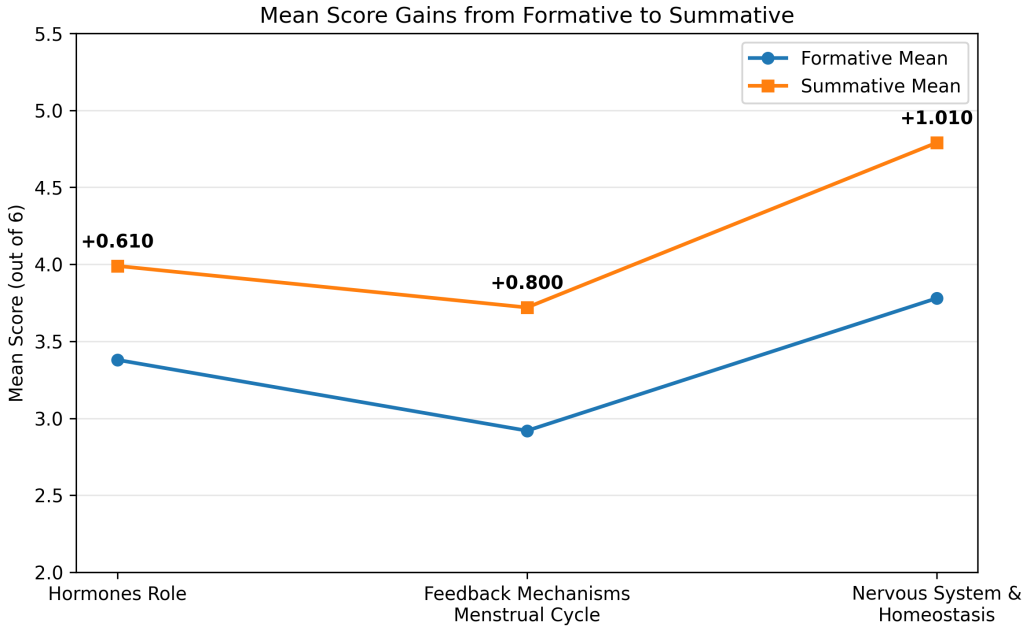


Figure 3. Mean Score Gains from Formative to Summative (max score 6). Annotations show +0.611, +0.800, +1.011.

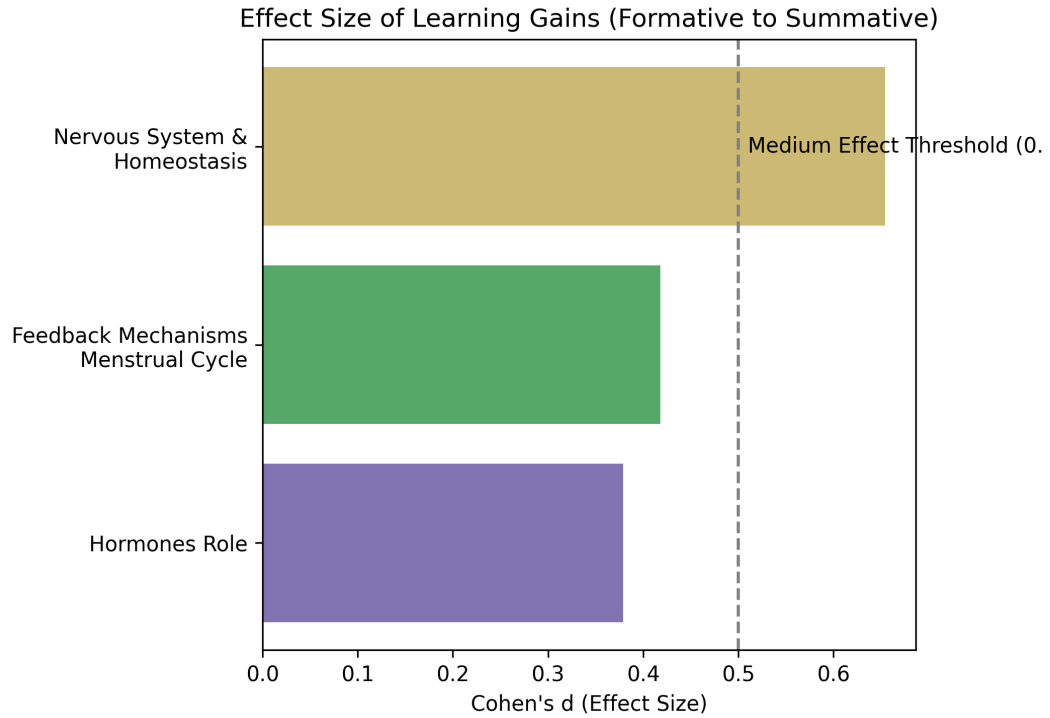


Figure 4. Effect Size (Cohen's d) of Learning Gains. Vertical dashed line indicates medium effect threshold (0.5). Only nervous system & homeostasis exceeds medium.

Teachers' Experiences in Assessing Grade 10 Biology

Thematic analysis of four teachers yielded three themes. Theme 1: Barriers to Science Learning – T1: 'Challenging to assess because you need to determine their level of comprehension, especially

terminologies and application/impact on life.' T1: 'Not all students can comprehend basic concepts, even reading simple thoughts.' T1: 'Students find it hard to grasp feedback mechanisms, cause-effect in hormone production.' T4: 'Students struggle with specific hormones (estrogen, progesterone, testosterone, LH, FSH) and how they affect reproductive systems at different life stages.' Subthemes: linguistic complexity (Batool et al., 2021), prior knowledge gaps, hands-on deficits post-pandemic (Tijap & Pelgone, 2022), spiral progression confusion (Antipolo & Rogayan, 2021), inclusivity gaps (Villaroel et al., 2023). Theme 2: Effective Strategies – T2: 'Using diagrams and flowcharts was most effective to simplify FSH, LH, estrogen, progesterone rise/fall.' T2: 'I encouraged storytelling, comic strips narrating body response step by step.' T3: 'Adapt assessment to balance scientific accuracy with student comfort. Visual aids, real-life scenarios, interactive discussions help grasp hormonal feedback while ensuring respectful environment.' Strategies: visual scaffolds, storytelling, real-life contextualization (Mulyani & Doyan, 2023), collaborative discourse (Bhardwaj et al., 2025), blended tools Google Classroom (Andyhapsari & Djukri, 2021), gender-responsive respectful discussion (Navarro & Torres, 2022). Elkhidir (2020) agency and reflection. Theme 3: Time Constraints and Learning Pace – UNESCO IIEP (2023) adequate time essential. Teachers noted limited time restricts deep formative assessment and differentiation in large classes (42-45) and double shifts. Sarmiento et al. (2020) large numbers drive multiple-choice persistence. Petalla & Doromal (2021) need clear instructions and fair rubrics. Integration: quantitative gains align with qualitative strategies—where visual/real-life scaffolds used, mastery rose, especially homeostasis. Where barriers highest and time shortest, menstrual feedback weakest.

Proposed Authentic Task Assessment

Based on findings, three authentic tasks designed: Task 1 Health Workers of the Day: Know It, Share It! (Hormones) – role-play barangay health workers creating brochures on hormonal changes across age/gender, rubric: Understanding, Accuracy/Believability, Cooperation, Presentation. Task 2 Hormonal Feedback and Mark the Calendar (Menstrual feedback) – comic strip narrative + 28-day cycle calendar with hormone graph, reflection on health relevance, rubric: Completeness, Understanding, Reflection, Organization. Task 3 Homeostasis Health Analysis – case studies (fever, dehydration, stress) showing nervous-endocrine coordination, produce infographic and short video. Each includes objectives, scheme (2-3 sessions), materials, 4-point rubric (Exceptional to Attempted). Embodies Dewey, Vygotsky, Assessment for Learning.

Conclusion:

Formative assessment revealed Developing Mastery dominance and menstrual feedback as most difficult (11.43% mastery, M=2.92). Summative showed significant improvement (44.57%, 42.86%, 66.29%; $p<.001$; $d=0.379,0.418,0.654$) with largest gain for homeostasis. Teacher experiences explained patterns: barriers (abstract terminology, sociocultural sensitivity, prior gaps, time), effective strategies (visual scaffolds, comic strips, storytelling, real-life contextualization, collaborative discourse, gender-responsive discussion), and time constraints. Integration validates authentic tasks making invisible hormonal feedback visible and personally relevant. Proposed output offers practical, equitable assessment package. Limitations: single-site convenience, 3-week duration, SLM-based tests. Future: quasi-experimental across schools, include student voice, teacher training on rubrics and culturally responsive reproductive health education, DepEd time allocation review. Assessment significance lies in informing instruction, not just measuring.

References

Abulhul, Z. (2021). The use of PowerPoint and interactive methods in science teaching. *Journal of Science Education, 12*(2), 45–58.

Almarghani, E. M., & Mijatović, I. (2021). Teaching hormonal physiology: Challenges and real-life connections. *Journal of Biological Education, 55*(3), 234–247.

Andyhapsari, S., & Djukri, D. (2021). Blended learning strategies and self-regulated learning in biology. *International Journal of Instruction, 14*(3), 567–584.

Antipolo, A. M., & Rogayan, D. V. (2021). Spiral progression approach in K-12 science: Reinforcement or redundancy? *Journal of Education and Practice, 12*(15), 88–98.

Batool, S., et al. (2021). Language complexity and science learning difficulties. *Journal of Education and Learning, 10*(4), 112–125.

Bhardwaj, S., et al. (2025). Collaborative projects and student engagement in biology. *Science Education International, 36*(1), 23–38.

Bird, K. (2021). Impact of COVID-19 on science laboratory learning. *Journal of Science Teaching, 9*(2), 101–115.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Brown, S., & Harris, L. (2023). Real-world examples and understanding of homeostasis. *Biology Teaching Journal, 65*(2), 145–158.

Brown, S., & Sambell, K. (2023). Authentic assessment and student independence. *Assessment & Evaluation in Higher Education, 48*(4), 512–526.

Bulut, O., et al. (2024). Technological developments in assessment implementation. *Educational Assessment, 29*(1), 34–52.

Chen, M., & Ramirez, P. (2020). Designing biology assessments for conceptual depth: Balancing content and cognition. *Journal of Science Assessment, 28*(4), 201–218.

Choudhary, F., & Bano, R. (2022). Concept maps as an effective formative assessment tool in biology at the secondary level. *Journal of Education and Educational Development, 9*(1), 157–175. https://doi.org/10.22555/joeed.v9i1.454

Department of Education. (2020). *DepEd Order No. 31, s. 2020: Interim guidelines for assessment and grading in light of the Basic Education Learning Continuity Plan*. https://www.deped.gov.ph/

Elkhidir, M. (2020). Student-centered approaches in biology education. *Journal of Biological Education, 54*(2), 101–115.

Fernandez, L., & Liu, H. (2024). Prior knowledge and interlevel strengthening in multi-level problem solving. *Cognitive Science Education, 18*(1), 67–82.

Foster, H. (2024). The impact of formative assessment on student learning outcomes: A meta-analytical review. *Academy of Educational Leadership Journal, 28*(S1), 1–3.

Foster, L. (2024). Formative assessment strategies in biology education. *Philippine Journal of Science Teaching, 15*(1), 23–38.

Garcia, M. (2025). Authentic assessment and scientific literacy in secondary biology. *Journal of Education Research, 20*(4), 55–70.

Gilliam Writers Group. (2023). *John Dewey's experiential learning: Transforming education through hands-on experience*. https://www.gilliamwritersgroup.com/blog/john-deweys-experiential-learning-transforming-education-through-hands-on-experience

Juma, F., & Ellis, R. (2021). Addressing student misconceptions in reproductive health education. *Health and Science Education, 15*(3), 145–160.

Karunanayaka, S., et al. (2020). Employability and skills development through authentic assessment. *International Journal of Education, 18*(2), 77–89.

Khajeloo, P., et al. (2021). Digital divide and student performance during online learning. *Educational Studies, 33*(1), 88–103.

Lee, J., & Salter, D. (2020). Scaffolding complex concepts: Teaching physiological feedback loops in secondary science. *International Journal of Science Education, 42*(11), 1830–1846.

McArthur, J. (2021). Authentic assessment and employability: Bridging the gap between education and the workplace. *Higher Education Studies, 8*(3), 67–79.

McLeod, S. (2024). *Vygotsky's social development theory*. Simply Psychology. https://www.simplypsychology.org/vygotsky.html

Mulyani, S., & Doyan, A. (2023). Contextual teaching methods and their role in enhancing scientific literacy in biology education. *Journal of Biology Education Research, 18*(3), 87–98.

National Education Association. (2023, March 30). Standardized testing is still failing students. *NEA Today*. https://www.nea.org/nea-today/all-news-articles/standardized-testing-still-failing-students

Navarro, C., & Torres, A. (2022). Overcoming menstrual education barriers: Enhancing curriculum inclusivity in biology classrooms. *Educational Research and Reviews, 17*(6), 122–132.

Pappas, C. (2023). *Using John Dewey's theory to reform education*. eLearning Industry. https://elearningindustry.com/john-deweys-theory-reforming-education-through-experiential-learning

Petalla, J. (2021). Performance-based assessments in the Philippine K to 12 biology. *Philippine Journal of Science Education, 50*(2), 88–102.

Petalla, J., & Doromal, A. (2021). Challenges in low performance tasks in senior high school. *Philippine Journal of Education, 52*(1), 45–60.

Sambell, K., Brown, S., & McDowell, L. (2023). Designing authentic assessments for self-directed learning. *Journal of Educational Assessment, 15*(2), 123–135.

Sarmiento, A., Cruz, R., & Lim, E. (2020). STEAM education and authentic assessment in Philippine schools. *Asian Journal of Education, 21*(3), 205–220.

Simply Psychology. (2024). *Vygotsky's social development theory*. https://www.simplypsychology.org/vygotsky.html

Singh, V., & Kaur, T. (2023). Systems-based teaching: Improving student outcomes in human physiology. *Biology Teaching International, 64*(3), 189–202.

Tesorio, J., & Canizares, M. (2018). Assessment perspectives of science teachers in Cebu. *Cebu Journal of Science Education, 5*(1), 12–28.

Tijap, M. S., & Pelgone, R. P. (2022). Online teaching in the Philippines: Struggles of science teachers during the pandemic. *Education in Science, 45*(4), 291–306.

UNESCO International Institute for Educational Planning. (2023). *Adequate instructional time for quality education*. UNESCO Learning Portal. https://learningportal.iiep.unesco.org/en/issue-briefs/improve-learning/instructional-time-and-classroom-management

Vasilev, K., & Peters, S. (2021). Gender-responsive pedagogy and reproductive health outcomes. *Gender and Education, 33*(5), 567–583.

Villanueva, C. J., & Chan, M. D. (2023). Visualizing feedback loops: A scaffolded approach to hormonal systems. *Biology Teacher Journal, 85*(4), 312–318.

Villaroel, V., Bloxham, S., & Bruna, D. (2023). Inclusive practices in authentic assessment: Supporting diverse learners. *Educational Research Review, 20*(4), 89–102.

Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Yilmaz, R., & Kara, Y. (2021). Overcoming challenges in teaching reproductive biology: A mixed-method study. *Science Education International, 32*(2), 112–121.