

---

## Stealth Assessment Set in Enhancing Scientific Creativity and Critical Thinking Skills for Grade 10 Chemistry

Sheryl Lirio<sup>1\*</sup>, & Julie Fe Panoy<sup>2</sup>

<sup>1</sup>Bernardo Lirio Memorial National High School, **Philippines**

<sup>2</sup>College of Arts and Sciences, Laguna State Polytechnic University, **Philippines**

DOI - <http://doi.org/10.37502/IJSMR.2025.8703>

---

### Abstract

This study aimed to determine the effectiveness of a stealth assessment in enhancing students' creative and critical thinking skills through game-based learning. The study involved two groups of respondents: 40 Grade 10 students from Bernardo Lirio Memorial National High School and eight expert validators from different schools in Tanauan City. The study utilized a developmental research design. The instruments used included pre-test and post-test questionnaires to measure students' thinking skills, a stealth assessment set embedded in digital games, and an evaluation rubric for expert validation. Statistical treatments involved frequency and percentage to describe respondents' interest in digital games, weighted mean for analyzing game preferences and expert evaluations, and paired sample t-tests to determine significant differences in the students' performance before and after the intervention. According to the findings, pupils demonstrated a high interest in digital games, with a preference for genres such as action, adventure, and simulation. Points, rewards, levels, and difficulties were the aspects of the game that they liked best. The most popular game-based learning apps were Quizizz and Kahoot. The pupils' increased post-test scores demonstrated significant improvements in their critical and creative thinking abilities. Expert validators evaluated the stealth assessment tool as highly effective in achieving the desired learning objectives. According to the findings, pupils demonstrated a high interest in digital games, with a preference for genres such as action, adventure, and simulation. Points, rewards, levels, and difficulties were the aspects of the game that they liked best. The most popular game-based learning apps were Quizizz and Kahoot. The pupils' increased post-test scores demonstrated significant improvements in their critical and creative thinking abilities. Expert validators evaluated the stealth assessment tool as highly effective in achieving the desired learning objectives.

**Keywords:** Stealth Assessment, Inquiry-Based Learning, Creativity, Critical Thinking, Game-Based Learning.

---

### 1. Introduction

In an increasingly interconnected and rapidly evolving world, education systems worldwide face the critical challenge of preparing students with 21st-century skills essential for navigating complex, information-driven, and innovation-based economies. The global consensus emphasizes that excellent education must transcend traditional knowledge transmission to develop critical thinking, creativity, problem-solving abilities, and effective communication skills, enabling students to assess, create, and resolve real-world issues (Trilling & Fadel, 2009;

Saavedra & Opfer, 2012). This transformational educational approach equips learners with the adaptability necessary to thrive in a rapidly changing global landscape.

The imperative for developing these competencies is particularly pronounced in science education, where the primary objective extends beyond content mastery to fostering scientific literacy, critical reasoning, and creative problem-solving capabilities. According to the National Research Council (2012), science education must equip students with skills necessary for critical and creative thinking, including modelling, reasoning, and debate. Similarly, Bybee (2010) emphasizes that science education should promote inquiry-based learning where students pose questions, formulate hypotheses, conduct experiments, and draw data-supported conclusions—processes that inherently develop both critical thinking and creativity (Reyes & Alianzas, 2021).

The global commitment to quality education is formalized through international frameworks, particularly the United Nations Sustainable Development Goals (SDGs), which emphasize inclusive and equitable learning opportunities worldwide. These international standards acknowledge that 21st-century skills—encompassing critical thinking, creativity, communication, collaboration, and digital literacy—are essential components of high-quality education, enabling individuals to adapt, innovate, and address complex global challenges.

Within this global framework, the Philippines has committed to providing high-quality, future-ready education that aligns with international standards and the SDG objectives (Department of Education, 2020). The Department of Education recognizes that developing 21st-century skills, particularly creativity and critical thinking, is central to equipping Filipino students with the competencies needed for success in an increasingly complex global economy. However, despite these national commitments and the implementation of the K-12 curriculum reform designed to incorporate 21st-century learning outcomes, the Philippine educational system continues to face significant systemic challenges. These obstacles include inadequate teacher preparation, insufficient instructional materials, and a persistent reliance on traditional assessment methods that emphasize memorization over critical thinking, application, and analysis (Llego, 2019; Bernardo, 2017).

The severity of these challenges is starkly illustrated by the Philippines' performance in international assessments. In the 2018 Programme for International Student Assessment (PISA), Filipino students ranked among the lowest globally in reading, mathematics, and science (OECD, 2019). More concerning, the 2022 PISA results revealed that Filipino students scored exceptionally low in creative thinking, with an average score of 14 out of 60, placing them 63rd out of 64 participating countries, significantly below the OECD average of 33 (OECD, 2024). This persistent gap between the Philippines' aspirational educational goals and actual student performance reveals fundamental systemic issues, including the prevalence of rote learning methods, inadequate assessment strategies, and limited opportunities for students to engage in higher order thinking processes.

These national challenges manifest at the regional and local levels, where individual schools struggle to implement effective strategies for developing 21st-century competencies. At Bernardo Lirio National High School, for example, science performance has consistently fallen short of expectations over the past five years. The school's Science Mean Percentage Scores (MPS) have remained well below the Department of Education's 75% target, with scores of 52.98% in 2018-2019, 54.05% in 2019-2020, 56.14% in 2020-2021, 55.26% in 2021-2022,

and declining to 52.49% in 2022-2023. These figures demonstrate that students face difficulties not only in mastering scientific content but also in applying scientific knowledge critically and creatively to solve complex problems.

In response to these persistent challenges at both national and local levels, educational researchers and practitioners have begun exploring innovative assessment approaches that can simultaneously measure and develop 21st-century skills. Stealth assessment emerges as a particularly promising solution, offering a technology-enhanced approach that seamlessly integrates real-time skill evaluation into digital learning environments without disrupting the learning process (Shute, 2011). This cutting-edge assessment method enables continuous monitoring of students' creative outputs, reasoning techniques, and problem-solving processes while they engage in meaningful learning activities, providing both measurement and developmental opportunities for critical thinking and creativity.

Digital games, enhanced through stealth assessment capabilities, have demonstrated significant potential to transform educational experiences by creating engaging, performance-based assessment environments that mirror real-world problem-solving scenarios (Quian & Clark, 2016; Shute & Ventura, 2013). Research indicates that well-designed educational games not only capture student engagement but also provide ongoing measures of learning progress while fostering essential 21st-century competencies such as creativity, persistence, and collaboration (Udeozor, Abegao, & Glassey, 2023).

Given the urgent need to address declining 21st-century skills among Filipino students, particularly in science education, and recognizing the potential of innovative assessment approaches to bridge the gap between learning objectives and actual student performance, this study aims to develop a comprehensive stealth assessment framework specifically designed for Grade 10 Chemistry instruction. The research aims to investigate whether integrating digital game-based stealth assessment can significantly enhance scientific creativity and critical thinking skills among learners, ultimately contributing to the Philippines' goal of providing high-quality, future-ready education that prepares students for success in an increasingly complex and demanding global environment.

## **2. Methodology**

### **2.1 Research Design**

This research employed a developmental research design to create and evaluate a stealth assessment set designed to enhance students' creative and critical thinking abilities. Developmental research is suitable when the primary objective involves creating, developing, and validating educational resources (Richey & Klein, 2007). The design incorporated a try-out phase wherein pre-test and post-test results were gathered and compared following the utilization of stealth assessment among Grade 10 students.

### **2.2 Participants of the Study**

Participants were divided into two groups: eight expert validators and Grade 10 students. Expert validators were Master Teachers with Master of Arts in Education degrees specializing in Science Education from various schools in Tanauan City. They evaluated the developed stealth assessment set for accuracy, relevance, and curriculum alignment.

The second group consisted of 40 Grade 10 students from Bernardo Lirio Memorial National High School (academic year 2024-2025), purposively selected from five sections handled by the researcher. Purposive sampling ensured participants possessed appropriate cognitive abilities corresponding with the stealth assessment learning objectives and provided consistent implementation monitoring.

### 2.3 Research Instruments

Four instruments were utilized: (1) Learner's Interest in Digital Games and Game-Based Learning Assessment Survey - a researcher-developed tool determining student preferences for game-based learning, focusing on game genres, elements, and applications; (2) Teacher-made PISA-based Test - administered as pre-test and post-test to measure creative and critical thinking capabilities using science-based scenarios requiring higher-order thinking; (3) Learning Resources Management and Development System (LRMDS) Evaluation Tool - used by experts to assess the stealth assessment set's design, usability, instructional alignment, and content quality; and (4) Student Perception Survey on Stealth Assessment Effectiveness - gathered student feedback on engagement, motivation, and learning effectiveness.

### 2.4 Data Gathering Procedure

The study proceeded through structured phases: (1) administration of the interest survey to determine student game-based learning preferences; (2) development of the stealth assessment set based on survey results; (3) expert validation using the LRMDS tool with statistical analysis consultation; (4) pilot testing with pre-test administration, 10-day teaching intervention implementing the stealth assessment, and post-test administration; and (5) student feedback survey collection regarding their digital gaming assessment experiences.

### 2.5 Data Analysis

Frequency and percentage described pre-test and post-test performance levels. A paired t-test examined differences between respondents' pre-test and post-test scores to determine the stealth assessment's effect on developing scientific creativity and critical thinking skills.

## 3. Results and Discussion

**Table 1. Profile of the Respondents According to Interest in Digital Games**

Indicators	Mean	SD	Verbal Interpretation
1. I feel motivated to learn when using digital games.	3.35	.622	Highly Interested
2. I think playing games makes learning even more fun.	3.48	.554	Highly Interested
3. I like digital games because it helps me understand complex topics easier.	3.18	.501	Interested
4. I prefer using educational digital games over textbooks.	3.25	.630	Highly Interested
5. I am interested in digital games because it can improve creative and critical thinking skills.	3.60	.591	Highly Interested
6. I would like more learning activities that include digital games.	3.50	.506	Highly Interested

7. I prefer digital games because they help me retain information better than lectures.	3.00	.641	Interested
8. I believe applying digital games in a lesson improves students' engagement in school.	3.55	.552	Highly Interested
9. I would recommend digital game-based learning to others.	3.37	.490	Highly Interested
10. I believe my learning style is suited for digital game-based learning.	3.05	.597	Interested
<b>Total</b>	<b>3.33</b>	<b>.371</b>	<b>Highly Interested</b>

*Legend: 3.25-4 Highly Interested, 2.50-3.24-Interested, 1.75-2.49-Moderately Interested, 1.0-1.74-Not Interested*

Table 1 presents the learners' interest in digital games, which was evaluated using a 10-item survey. The average interest score was 3.33 (SD = 0.37), indicating a generally strong interest in digital games among the participants. The mean scores for individual items varied from 3.00 to 3.60, indicating that while most participants demonstrated a consistent interest in digital games, certain items elicited slightly stronger or weaker levels of agreement. Notably, the item with the highest means of 3.60 (item 5) indicates that students are interested in digital games because they believe they can improve critical and creative thinking skills. Upon closer examination, digital gaming enhances critical thinking skills among high school students by promoting problem-solving, strategic planning, and informed decision-making. Games that are typically presented digitally often feature complex scenarios that require players to analyze information, evaluate options, and anticipate potential consequences. This interactive environment encourages students to think critically as they navigate challenges and adapt their strategies.

Students today are often referred to as "digital natives," having grown up with computers, and as such, they are drawn to game-based environments, according to Prensky (2001). This statement was supported by the results of this table, which show that junior high school students have a great desire to incorporate digital games into their learning experience. Most respondents reported that playing digital games boosts motivation and engagement, and many believe that these resources make studying more enjoyable. Students specifically link playing digital games to improved knowledge of challenging subjects, enhanced memory retention, and the development of critical and creative thinking skills. Perhaps the reason for this interest is that, unlike more conventional teaching strategies such as textbook reading or passive lectures, digital games are immersive and interactive. Additionally, the replies indicate that students believe digital games align with their preferred learning styles, hence facilitating more individualized engagement and tailored instruction.

A study by Rideout and Robb (2019) analyzing data from Common Sense Media found that teens worldwide, aged 13 to 18, spend an average of 1 hour and 13 minutes per day playing video games. With a focus on junior high school students (ages 15–16), Anderson and Jiang (2018) found that 84% of teenagers routinely play video games, with boys playing more frequently and spending an average of more than two hours a day. Moreover, these results are echoed specifically in the Philippines. For example, Villaruel et al. (2022) found that high school students in the country prefer mobile platforms and spend two to three hours a day playing digital games. This high frequency implies that it is not merely a recreational activity

but also an important aspect of adolescents' everyday lives, which teachers can use to support meaningful learning.

Similarly, digital gaming can foster engagement among high school students (mean = 3.55, SD = .552) by providing interactive and immersive experiences that capture their interest. Aside from the dynamic nature of games that encourage active participation, gamification elements (such as rewards, levels, and achievements) also motivate students to invest their time and effort in their learning. Additionally, the social aspect of multiplayer games fosters collaboration and communication, enabling students to connect with their peers and share strategies. This combination of competition, instant feedback, and community creates a stimulating environment that enhances student engagement and encourages a positive attitude toward learning.

Digital gaming can also transform learning into an enjoyable and engaging experience (mean = 3.48, SD = 0.554) for high school students, most likely by incorporating elements of competition, storytelling, and interactivity. Games presented digitally also often feature immersive narratives and visually appealing graphics that capture students' attention, making educational content more relatable and exciting. Research likewise indicates that engaging with digital games enhances cognitive skills, motivation, and emotional development (Granic, Lobel, & Engels, 2014). Furthermore, gamified learning strategies have been shown to improve student engagement and motivation, making learning more interactive and enjoyable (Deterding et al., 2011). This intersection of gaming and education highlights the potential of digital games as practical tools for developing essential skills among high school students, making them an ideal consideration for developing lessons and learning materials that support student learning.

**Table 2. Student Preferences for Game-Based Assessment**

<b>Category</b>	<b>Component</b>	<b>Frequency (f)</b>	<b>Percentage (%)</b>
<b>Game Genre</b>	Adventure	29	72.5
	Action	28	70.0
	Simulation	25	62.5
	Puzzle	23	57.5
	Strategy	22	55.0
	Sports	20	50.0
	Role-Playing games	11	27.5
<b>Game Elements</b>	Points	30	75.0
	Reward	29	72.5
	Level	25	62.5
	Challenge	25	62.5
	Leaderboard	21	52.5
	Badges	20	50.0
	Social Interaction	15	37.5
<b>Learning Applications</b>	Kahoot	36	90.0
	Quizizz	24	60.0
	Helpful Games	14	35.0
	Legends of Learning	8	20.0
	Blooket	2	5.0
	Liveworksheet	1	2.5

The results reveal clear preferences among Grade 10 students for specific game-based assessment components that align with their learning motivations and technological experiences. In terms of game genres, students demonstrated strong preferences for adventure (72.5%) and action (70.0%) games, followed by simulation (62.5%) and puzzle (57.5%) games. These genres typically incorporate decision-making, challenges, exploration, and instant feedback mechanisms, making learning more dynamic and engaging. This approach reflects students' need for autonomy, stimulation, and interaction, which are often absent in traditional assessment methods (Hamari et al., 2021; Kapp, 2020).

Regarding game elements, students showed overwhelming preference for points (75.0%) and rewards (72.5%), indicating their attraction to components that provide immediate feedback and recognition of achievement. The strong preference for levels and challenges (both 62.5%) suggests that students value structured progression and task complexity as important motivators. These preferences align with extrinsic motivation factors that can enhance student engagement when implemented thoughtfully (Deci & Ryan, 2000). Notably, social interaction received the lowest preference (37.5%), possibly indicating limited exposure to collaborative learning environments or a preference for individual achievement over group-based activities.

Among learning applications, Kahoot emerged as the most preferred platform (90.0%), followed by Quizizz (60.0%). This strong preference for familiar, interactive, and competitive platforms demonstrates students' attraction to learning tools that resemble play, offer real-time feedback, and allow progress tracking. The significant gap between these popular platforms and less familiar applications, such as Blooket (5.0%) and Liveworksheet (2.5%), suggests that exposure and familiarity strongly influence student preferences.

These findings indicate that students respond positively to game-based assessments that provide control, visibility, and feedback—elements often lacking in traditional evaluation models. The preference patterns suggest that students are capable, motivated, and receptive to modern educational tools, although their choices are influenced by the constraints of their immediate educational environment. The researcher strategically integrated these preferred components into the stealth assessment design, utilizing puzzle-based games for diagnostic assessments, Quizizz for formative assessments, and Blooket for summative evaluations, ensuring alignment with student interests while maintaining educational effectiveness.

**Table 3. Expert Evaluation of Stealth Assessment Set Effectiveness**

<b>Evaluation Criteria</b>	<b>Overall Mean</b>	<b>Standard Deviation</b>	<b>Verbal Interpretation</b>
Content	3.88	0.258	Very Satisfactory
Prints	3.94	0.116	Very Satisfactory
Illustrations	3.83	0.236	Very Satisfactory
Design and Layout	3.84	0.229	Very Satisfactory
Paper and Binding	3.81	0.259	Very Satisfactory
Size and Weight	3.94	0.177	Very Satisfactory
Presentation and Organization	4.00	0.000	Very Satisfactory
Accuracy and Up-to-datedness	3.92	0.178	Very Satisfactory

The expert evaluation results consistently demonstrate high ratings across all assessment criteria, with overall means ranging from 3.81 to 4.00, all of which fall within the "Very

Satisfactory" range. Eight expert validators, comprising Master Teachers with Master of Arts in Education degrees specializing in Science Education, evaluated the developed stealth assessment set using the Department of Education's Learning Resources Management and Development System (LRMDS) Evaluation Tool.

The highest-rated criterion was Presentation and Organization ( $M = 4.00$ ,  $SD = 0.00$ ), indicating unanimous expert agreement that the material demonstrates an engaging presentation, a logical flow of ideas, an appropriate vocabulary level, suitable sentence length, and varied structure. This perfect rating reflects the assessment set's clear and structured presentation of information, which builds systematically on prior knowledge, aligning with effective instructional design principles (Merrill, 2002; Gagné et al., 2005).

Size and Weight ( $M = 3.94$ ,  $SD = 0.177$ ) and Prints ( $M = 3.94$ ,  $SD = 0.116$ ) received the second-highest ratings, indicating excellent physical characteristics that enhance usability. The lightweight, portable design, along with an appropriate font size, spacing, and print quality, facilitates classroom implementation and reduces cognitive load during instruction (Mayer, 2009).

Accuracy and Up-to-datedness ( $M = 3.92$ ,  $SD = 0.178$ ) received strong ratings, with experts particularly noting the absence of factual and computational errors. This high accuracy level is crucial for chemistry instruction, where scientific principles must align with current IUPAC guidelines and contemporary applications.

The remaining criteria—Content ( $M = 3.88$ ), Design and Layout ( $M = 3.84$ ), Illustrations ( $M = 3.83$ ), and Paper and Binding ( $M = 3.81$ )—all received "Very Satisfactory" ratings. Content evaluation highlighted the material's developmental appropriateness, objective alignment, and potential for developing higher-order thinking skills, including creativity and critical thinking. The design elements effectively support learning without causing cognitive overload, while illustrations successfully clarify and supplement textual content.

The consistently low standard deviations (ranging from 0.000 to 0.259) indicate strong inter-rater agreement among experts, suggesting reliable evaluation outcomes. The overall "Very Satisfactory" ratings across all criteria demonstrate that the developed stealth assessment set meets professional standards for educational materials and is ready for classroom implementation. These results validate the tool's potential effectiveness in enhancing science instruction while seamlessly integrating assessment into game-based learning activities.

**Table 4. Student Performance in Critical and Creative Thinking Skills**

Thinking Skills Component	Performance Level	Pre-test		Post-test	
		f	%	f	%
<b>Analysis</b>	Advanced (13-15)	0	0	21	52.5
	Proficient (10-12)	0	0	14	35.0
	Approaching Proficiency (7-9)	6	15	3	7.5
	Developing (4-6)	8	20	2	5.0
	Beginning (0-3)	26	65	0	0
<b>Logical Reasoning</b>	Advanced (13-15)	0	0	21	52.5
	Proficient (10-12)	0	0	14	35.0

	Approaching Proficiency (7-9)	6	15	3	7.5
	Developing (4-6)	8	20	2	5.0
	Beginning (0-3)	26	65	0	0
<b>Logical Justification</b>	Advanced (13-15)	0	0	25	62.5
	Proficient (10-12)	1	2.5	13	32.5
	Approaching Proficiency (7-9)	11	27.5	2	5.0
	Developing (4-6)	22	43.0	0	0
	Beginning (0-3)	6	15.0	0	0
<b>Application of Scientific Concepts</b>	Advanced (13-15)	0	0	24	60.0
	Proficient (10-12)	2	5.0	14	35.0
	Approaching Proficiency (7-9)	8	20.0	2	5.0
	Developing (4-6)	21	52.5	0	0
	Beginning (0-3)	9	22.5	0	0

The pre-test and post-test results demonstrate remarkable improvements in students' critical and creative thinking skills following the implementation of the stealth assessment intervention. The data reveal a dramatic transformation in student performance across all four measured components: analysis, logical reasoning, logical justification, and application of scientific concepts.

Initial assessment results showed concerning performance levels, with the majority of students (65%) performing at the "Beginning" level in both analysis and logical reasoning skills. Similarly, substantial portions of students demonstrated limited abilities in logical justification (15% at the Beginning level, 43% at the Developing level) and the application of scientific concepts (22.5% at the Beginning level, 52.5% at the Developing level). These baseline results indicated that students initially struggled with interpreting scientific data, identifying relationships between variables, constructing logical arguments, and applying scientific principles to real-world scenarios.

Following the stealth assessment intervention, student performance showed remarkable improvement across all components. The most striking transformation occurred in the elimination of students performing at the "Beginning" level, as no students remained in this category across any of the four components. Furthermore, substantial percentages of students achieved "Advanced" performance levels: 52.5% in analysis, 52.5% in logical reasoning, 62.5% in logical justification, and 60.0% in application of scientific concepts.

Logical justification showed the highest proportion of advanced performers (62.5%), followed closely by application of scientific concepts (60.0%). This suggests that the stealth assessment approach was particularly effective in developing students' abilities to construct well-reasoned arguments and apply scientific principles creatively to novel situations. The consistent improvement pattern across analysis and logical reasoning (both achieving 52.5% advanced performance) indicates balanced development of fundamental critical thinking components.

The intervention not only improved quantitative performance but also enhanced the quality and sophistication of student responses. Post-test answers demonstrated a deeper conceptual understanding, the use of scientific terminology, the integration of chemical equations, and the

ability to connect theoretical knowledge with practical applications. Students progressed from basic, memorization-based responses to sophisticated explanations that demonstrated systematic reasoning and a scientific rationale.

These results provide strong evidence that the stealth assessment approach effectively transforms students from passive information consumers into active critical thinkers capable of meeting international scientific literacy standards. The dramatic reduction in lower-performing categories and substantial increases in advanced performance levels suggest that game-based stealth assessment creates an engaging, non-threatening environment that facilitates deeper learning and skill development in science education.

**Table 5. Test of the Difference between the Pre-test and Post-test Scores of the Student Respondents in Critical Thinking Skills**

Critical Thinking	Pre-test		Post test		Paired Differences					t	Df	Sig. (2-tailed)
	Mean	SD	Mean	SD	Mean	SD	SEM	95% CID Lower	95% CID Upper			
Analysis	3.13	2.45	12.00	2.43	8.87	2.99	0.474	7.916	9.834	18.727	39	0.000
Logical Reasoning	3.08	2.45	11.98	2.44	8.9	2.97	0.471	7.948	9.852	18.909	39	0.000

*SD- Standard Deviation SEM - Standard Error of the Mean CID - Confidence Interval of the Difference*

Table 5 shows the test difference between the pre-test and post-test scores of the student respondents in critical thinking skills. The findings of the paired samples t-test indicated a statistically significant difference between the student respondents' pre- and post-test ratings regarding their capacity for creative thought. In particular, the mean score for analysis skills increased from 3.13 to 12.00, resulting in a mean difference of 8.88, which was statistically significant ( $t(39) = 18.727, p < 0.001$ ). The significant increase in analysis in the pre-test and post-test results indicates that students' critical thinking abilities—especially their capacity to break down information, identify connections, and critically evaluate data—have improved substantially. The significant increase in students' post-test scores in the analysis can be directly attributed to the employment of stealth assessment. Stealth assessment observes critical thinking during problem-solving exercises rather than evaluating it on a single occasion. When teaching chemical reactions, for instance, students may be required to use a novel reaction to build an inventive fuel, which would require them to generate, evaluate, and refine ideas. Additionally, there is no rigid "right answer" in the process. Students are instead encouraged to provide multiple interpretations of occurrences, such as the reasons behind a balloon's size changes in response to variations in pressure and temperature. According to Torrance (2000), this fosters the growth of fluency and flexibility, two essential elements of creativity. Moreover, students' replies are tracked and analyzed by the system or teacher as they finish tasks. Using creativity markers such as "original idea generation" or "elaboration of explanation," the stealth evaluation technique enables real-time feedback.

The Stealth assessment describes observable creative behaviors at every level, from beginner to expert. Rubrics were developed based on originality, elaboration, and flexibility. It even

challenges learners to think more deeply (for example, "Can you find a more efficient or eco-friendly method?"). By providing instructors with organized support, students might continuously foster creativity instead of seeing it as a distinct skill. Additionally, these assessments have seen a significant increase in use across different platforms in the diagnostic, formative, and summative parts of the developed stealth assessment set, as well as the inclusion of adventure and action game genres in the learning process. This improvement implies that an interactive and stimulating learning environment, where students could hone their analytical abilities without the stress of formal assessment, was beneficial to them. By its very nature, stealth assessment permits ongoing, discreet evaluation, allowing students to consider their ideas while they engage and play. While the game-based framework encouraged trial-and-error learning, pattern recognition, and strategic thinking, tools like Kahoot and Quizizz offered instant feedback and repeated exposure to analytical tasks.

The mean score for Logical Reasoning increased similarly, rising from 3.08 to 11.98, with a mean difference of 8.90, which is also extremely significant. Additionally, the mean score for creative thinking improved significantly from 6.20 on the pre-test to 23.98 on the post-test, resulting in a mean difference of 17.78, which was statistically significant. These results demonstrate that the intervention significantly enhanced students' abilities in both logical reasoning and analytical creativity. Students' increased capacity for logical reasoning can be attributed to several important components of the lesson plan and instructional design found in the stealth assessment set. Further intensifying this, Shute and Ventura (2013) emphasized that stealth assessments, when integrated into video games, promote deep learning by requiring players to practice systems thinking, problem-solving, and adaptability skills closely tied to logical and creative reasoning. This was also confirmed by Kuo and Chuang (2016), who discovered that by making learning interesting, iterative, and reflective, gamified platforms greatly enhanced learners' higher-order thinking.

Table 6 presents the test difference between the pre-test and post-test scores of the student respondents in terms of creative thinking skills. The results from the paired samples t-test indicate a statistically significant improvement in students' creative thinking skills from pre-test to post-test, specifically in the areas of Logical Justification and Application of scientific concepts. The mean score for Logical Justification was 5.43 (SD = 2.171) before the test, and it rose to 12.90 (SD = 1.972) after the test. Significant learning gains are indicated by the mean difference of 7.475. The result is highly significant, indicating that the increase was not the result of chance, with a t-value of 17.161, degrees of freedom (df) = 39, and a p-value of .000. The reliability of the result is further supported by the confidence interval (CI) for the mean difference, which spans from 6.594 to 8.356. The improvement in this area can be attributed to lesson segments that focused on argument construction, evidence evaluation, and explanation building. For instance, during inquiry-based discussions or problem-solving exercises, students were asked to use logical reasoning to support their decisions or predicted outcomes. The inclusion of scenario-based evaluations in the lesson guide, which required students to evaluate scientific scenarios and support their responses with knowledge of facts or principles, was one element that might have had a role in this. This was further supported by formative tests integrated into assessment platforms such as Quizizz and Kahoot, where students were prompted to improve their thinking in real-time by receiving instant feedback.

**Table 6. Test of the Difference between the Pre-test and Post-test Scores of the Student-Respondents in Creative Thinking Skills**

Creative Thinking	Pre-test		Post test		Paired Differences					t	df	Sig. (2-tailed)
	Mean	SD	Mean	SD	Mean	SD	SEM	95% CID				
								Lower	Upper			
Logical Justification	5.43	2.17	12.90	1.97	7.47	2.75	.43	6.59	8.35	17.1	3	.000
Application of Scientific Concepts	5.25	2.36	12.85	1.96	7.60	2.88	.45	6.67	8.52	16.6	3	.000

*SD- Standard Deviation SEM - Standard Error of the Mean CID - Confidence Interval of the Difference*

Likewise, the mean score for Application of Scientific Concepts increased from 5.25 (SD = 2.362) in the pre-test to 12.85 (SD = 1.968) in the post-test, resulting in a mean difference of 7.600. This result also showed high statistical significance ( $t = 16.687$ ,  $p = .000$ ), with a 95% CI ranging from 6.679 to 8.521. These results show that students' ability to apply scientific concepts in a creative thinking environment improved consistently and significantly because of the intervention. The increase in scores for Application of Scientific Concepts, which involves applying scientific understanding to issues or explaining real-world events, can be linked to lectures that required students to apply their academic knowledge to practical situations. The lesson guide likely included theoretically rich subjects that are ideally suited for experiential or game-based learning. For example, students were encouraged to apply the knowledge they had gained in a meaningful context by using adventure or simulation game genres, where they utilized scientific ideas to navigate real-life challenges. This genre frequently features components that resemble real-life decision-making and necessitate the application of scientific principles to advance in the game, such as virtual labs, experimentation assignments, and problem-solving quests. In line with this, Shute and Ventura (2013) further elaborated on this by demonstrating how stealth assessments embedded in games enhance the application of scientific concepts, enabling students to use what they have learned in dynamic, real-world scenarios. This game-based approach helps students bridge theory with practice, encouraging creative thinking through problem-solving and decision-making.

Table 7 shows the students' perceived effectiveness of the stealth assessment-integrated lesson. Students believe that the science lesson combined with the stealth assessment was very effective, as evidenced by the overall mean score of 4.43 (SD = 0.61). The result showed that students thought that the incorporation of stealth assessment added excitement and vitality to the classroom. When students characterize a lecture as "engaging," it implies that the clear distinction between learning and evaluation was blurred, allowing them to feel more like active participants than passive recipients. This aligns with Shute and Ventura's (2013) findings, who highlighted how stealth assessment enhances student engagement by integrating evaluation into real-world, interactive activities such as games or simulations, thereby lowering stress levels and promoting organic learning.

**Table 7. Students' Perceived Effectiveness of the Stealth Assessment-Integrated Lesson**

<b>Indicators</b>	<b>Mean</b>	<b>SD</b>	<b>Verbal Interpretation</b>
1. The science lesson that included a stealth assessment is engaging.	4.73	0.55	Very Effective
2. I feel that stealth assessment (hidden assessment through activities, games, or simulations) helped me learn better.	4.63	0.54	Very Effective
3. I think the stealth assessment helped in improving my understanding of science concepts.	4.65	0.48	Very Effective
4. I feel more motivated to participate in the lesson because of the use of stealth assessment	4.50	0.64	Very Effective
5. The activities in the stealth assessment method are clear and easy to understand.	4.53	0.55	Very Effective
6. I think stealth assessment is effective and should be used more often in science lessons	4.48	0.64	Very Effective
<b>Overall</b>	<b>4.59</b>	<b>0.61</b>	<b>Very Effective</b>

*Legend: For Indicators 2-8 and 12; 4.20- 5.0-Strongly Agree/ Very Effective; 3.40-4.19-Agree/Effective; 2.60-3.39- Neutral/Moderately effective; 1.80-2.59- Disagree/slightly effective; 1.0-1.79 Strongly Disagree/Not at all effective*

According to many students, stealth assessment improved their learning and deepened their comprehension of scientific ideas. This illustrates how students perceive meaningful learning when they believe they genuinely understand fundamental concepts through investigation and engagement, rather than merely memorizing material for an exam. Because stealth assessment fosters critical thinking, teamwork, and real-time problem-solving—all of which are ingrained in the learning activity itself—it enables this kind of deep learning, claim Bai et al. (2020). Additionally, students were more inclined to engage in classes that included stealth assessment. This incentive results from the assessment's implicit format, which eliminates the stress of being judged in a traditional, high-stakes setting. Rather, students view the evaluation as a component of their educational process. According to Tiefenthaler and Yau (2020), students are more likely to participate when evaluation appears as an organic part of the activity rather than as a stressful, stand-alone event.

Another key feature in students' perceptions is the ease of understanding and clarity. Students' general comfort and desire to participate were probably influenced by their perception that the stealth assessment tasks were understandable and accessible. Students are more likely to concentrate on learning objectives rather than worrying about their evaluation when learning assignments are clear and easy to understand (Shute & Rahimi, 2021). Lastly, a substantial percentage of students stated that stealth assessment should be utilized more frequently in

science courses, indicating a favorable shift in perceptions of alternative forms of evaluation. This view suggests that students are receptive to new methods of assessment that better fit the learning objectives of the twenty-first century, which include critical thinking, creativity, and teamwork (Baker & Siemens, 2014).

Table 8 presents the students' preferences for the Use of Stealth-Assessment in Learning Selected Topics in Grade 10 Chemistry. According to a survey conducted with 40 respondents, students strongly favour the use of stealth assessment in science classes due to its positive effects on cognition and emotional responses. Most students (57.5%) said that stealth assessment "makes learning more fun." This indicates that students appreciate the game-like and captivating nature of stealth assessment activities, which often incorporate interactive media, simulations, or problem-based exercises. This choice reflects a shift in learners' expectations toward more engaging, interactive, and experiential learning methods, where evaluation is an integral component of the teaching process rather than a stand-alone activity (Shute & Ventura, 2013). According to Vogt et al. (2020), engaging in enjoyable learning activities not only boosts motivation but also promotes improved information retention and a lasting interest in the subject.

**Table 8. Students' Preference on the Use of Stealth-Assessment in Learning Selected Topics in Grade 10 Chemistry**

Indicators	Frequency	Percentage
Advantages of a stealth assessment		
Makes learning more fun	23	57.5
Reduces test anxiety	4	10
Helps in better understanding concepts	8	20
Provides instant feedback	4	10
Others	1	2.5
<b>Total</b>	<b>40</b>	<b>100</b>

One instance that exemplifies the above concept is when a teacher combines the virtual lab simulations PhET Simulation and Legends of Learning in a lesson on Chemical Reactions. Students mix various chemicals to observe different types of reactions, including synthesis, decomposition, and single replacement. As a mission-based game, the task requires students to identify unknown chemical reactions and select the appropriate neutralizing agents to "save" a virtual city's water supply. Technology continuously monitors their decisions, theories, and reaction equations, acting as a stealth assessment of their comprehension. In response, the students said, "I thought we were just experimenting, but it turns out we were being assessed the whole time." "Among them, one said, "I still had to consider the kind of reaction that was taking place, even though it didn't feel like a test". One of them commented, "Finding the solution was enjoyable. We could keep trying until we got it right, which I appreciated. I gained a better understanding of equation balancing from that experience. These responses demonstrate how stealth assessment promotes an adventurous, stress-free environment where students learn by doing rather than memorizing facts. Students' comprehension of chemical reactions expanded because of their increased engagement and concentration on the task at hand.

Additionally, 20% of the students said that stealth assessment "helps in better understanding concepts." This underscores the cognitive advantage of stealth assessment, as students believe that being evaluated in an integrated, non-traditional way promotes deeper conceptual learning rather than merely memorizing at a surface level. Stealth assessments incorporate learning objectives into relevant situations, encouraging critical thinking and the application of knowledge in real-time, as contrasted to traditional assessments that frequently test rote knowledge (Bai et al., 2020). In science education, where understanding systems and processes is crucial, this type of learning is extremely beneficial (Postrado & Alianzas, 2024).

Stealth assessment was praised by 10% of students because it "reduces test anxiety." This is noteworthy since it implies a favorable emotional reaction to a low-stakes learning environment. Typical tests frequently cause stress, which can affect confidence and performance (Ifenthaler & Yau, 2020). Conversely, stealth exams conceal the evaluation element, allowing students to demonstrate their learning without the stress of a formal exam. Moreover, an additional 10% of respondents agreed that stealth assessment "provides instant feedback." With immediate feedback during learning activities, students can correct and reflect on their understanding as they interact with the material. According to Shute and Rahimi (2021), this approach promotes metacognitive development and aligns with the principles of formative assessment. It demonstrates that students appreciate systems that educate and direct their development in addition to seeking out interesting learning opportunities.

#### **4. Conclusion**

Based on the findings of this study, the implementation of stealth assessment in Grade 10 Chemistry instruction demonstrated significant effectiveness in enhancing students' critical and creative thinking skills. The research revealed that students initially performing at "Beginning" and "Developing" levels across multiple cognitive dimensions showed remarkable improvement, with substantial percentages achieving "Advanced" performance levels in analysis, logical reasoning, logical justification, and the application of scientific concepts following the intervention. The stealth assessment approach, which integrated students' preferences for adventure and action game genres, points-based reward systems, and familiar platforms like Kahoot and Quizizz, successfully created an engaging, non-threatening learning environment that facilitated deeper conceptual understanding and skill development.

The study's findings provide compelling evidence that game-based stealth assessment represents a viable pedagogical solution to the persistent challenges in Philippine science education, particularly in developing 21st-century skills. Expert validation confirmed the tool's high quality across all evaluation criteria, while student feedback indicated strong positive perceptions of the assessment's effectiveness. The dramatic reduction in students performing at lower proficiency levels and the substantial increase in advanced performance demonstrate that stealth assessment not only measures learning effectively but also actively promotes cognitive growth. These results suggest that when educational interventions align with student interests and employ innovative assessment strategies, they can successfully transform passive learners into active, critical thinkers capable of meeting international scientific literacy standards, thereby contributing to the Philippines' broader educational goals of providing high-quality, future-ready instruction.

**References**

- 1) Anderson, M., & Jiang, J. (2018). Teens, social media & technology 2018. Pew Research Center. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>
- 2) Bai, H., Pan, W., Hirumi, A., & Kebritchi, M. (2020). Assessing critical thinking and creative thinking in the context of game-based learning: A case for stealth assessment. *Educational Technology Research and Development*, 68(2), 877-900. <https://doi.org/10.1007/s11423-020-09757-4>
- 3) Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. In *Learning Analytics* (pp. 253-274). Springer. [https://doi.org/10.1007/978-1-4614-3305-7\\_10](https://doi.org/10.1007/978-1-4614-3305-7_10)
- 4) Baker, R. S., Clarke-Midura, J., & Ocumpaugh, J. (2020). Towards reliable stealth assessment of creativity in digital learning environments. *Computers in Human Behavior*, 106, 105849. <https://doi.org/10.1016/j.chb.2019.105849>
- 5) Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, 25(9), 6-8. <https://doi.org/10.3102/0013189X025009006>
- 6) Bernardo, A. B. I. (2017). Critical thinking in Philippine education: Rethinking the role of critical thinking in 21st-century learning. *Asia Pacific Journal of Multidisciplinary Research*, 5(3), 1-6.
- 7) Bybee, R. (2010). Advancing STEM education: A 2020 vision. *Technology and Engineering Teacher*, 70(1), 30-35.
- 8) Child Hope Philippines. (2023). Educational development and 21st-century skills in the Philippines. Child Hope Philippines Foundation.
- 9) Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268. [https://doi.org/10.1207/S15327965PLI1104\\_01](https://doi.org/10.1207/S15327965PLI1104_01)
- 10) Department of Education. (2016). DO No. 39, s. 2016 – Policy guidelines on the implementation of the senior high school career guidance program. Department of Education, Philippines.
- 11) Department of Education. (2020). The Philippines in PISA 2018: A second look. Department of Education, Republic of the Philippines. <https://www.deped.gov.ph/2020/12/03/the-philippines-in-pisa-2018-a-second-look/>
- 12) Department of Education. (2023). DO No. 24, s. 2023 – Guidelines on the provision of supplementary learning resources for public school libraries and library hubs. <https://www.deped.gov.ph>
- 13) Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9-15). ACM. <https://doi.org/10.1145/2181037.2181040>
- 14) Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2020). Gamification: Using game-design elements in non-gaming contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (pp. 2425-2428). ACM. <https://doi.org/10.1145/1979742.1979575>
- 15) Gagné, R. M., Wager, W. W., Golas, K. C., & Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Wadsworth/Thomson Learning.

- 16) Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice* (2nd ed.). Teachers College Press.
- 17) Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Palgrave Macmillan.
- 18) Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist*, 69(1), 66-78. <https://doi.org/10.1037/a0034857>
- 19) Halverson, R., & Owen, V. E. (2014). Game-based assessment: An integrated model for capturing evidence of learning in play. *International Journal of Learning and Media*, 4(2), 1-30. [https://doi.org/10.1162/ijlm\\_a\\_00068](https://doi.org/10.1162/ijlm_a_00068)
- 20) Hamari, J., Koivisto, J., & Sarsa, H. (2021). Does gamification work? A literature review of empirical studies on gamification. In 2014 47th Hawaii International Conference on System Sciences (pp. 3025-3034). IEEE. <https://doi.org/10.1109/HICSS.2014.377>
- 21) Harlen, W. (2010). *Principles and big ideas of science education*. Association for Science Education.
- 22) Kapp, K. M. (2020). *The gamification of learning and instruction: Game-based methods and strategies for training and education* (2nd ed.). Wiley.
- 23) Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. Springer.
- 24) Lee, J., & Kim, S. (2023). Digital game-based learning in science education: A systematic review of effectiveness and engagement. *Computers & Education*, 195, 104721. <https://doi.org/10.1016/j.compedu.2022.104721>
- 25) Llego, M. A. (2019). Challenges in implementing the K-12 curriculum in the Philippines. TeacherPH. <https://www.teacherph.com/challenges-implementing-k-12-curriculum/>
- 26) Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
- 27) Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, 50(3), 43-59. <https://doi.org/10.1007/BF02505024>
- 28) National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. <https://doi.org/10.17226/13165>
- 29) OECD. (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD Publishing. <https://doi.org/10.1787/5f07c754-en>
- 30) OECD. (2024). *PISA 2022 results (Volume III): Creative minds, creative schools – Philippines*. OECD Publishing. [https://www.oecd.org/en/publications/pisa-results-2022-volume-iii-factsheets\\_041a90f1-en/philippines\\_68d369a7-en.html](https://www.oecd.org/en/publications/pisa-results-2022-volume-iii-factsheets_041a90f1-en/philippines_68d369a7-en.html)
- 31) Postrado, F. & Aliazas, J.V. (2024). Designing Enrichment Program in Science Utilizing Blended Learning Models. *International Journal of Innovation Scientific Research and Review* 6 (6), 6568-6577.
- 32) Prensky, M. (2001). *Digital game-based learning*. McGraw-Hill.
- 33) Quian, M., & Clark, K. R. (2016). Game-based learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, 63, 50-58. <https://doi.org/10.1016/j.chb.2016.05.023>
- 34) Reyes, A., & Aliazas, J. V. (2021). Online integrative teaching strategies: Thematic and focus inquiry for improved science process skills. *International Journal of Science, Technology, Engineering and Mathematics*, 1(2), 19-38.

- 35) Richey, R. C., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Lawrence Erlbaum Associates.
- 36) Rideout, V., & Robb, M. B. (2019). *The Common Sense census: Media use by tweens and teens, 2019*. Common Sense Media. <https://www.commonsensemedia.org/>
- 37) Saavedra, A. R., & Opfer, V. D. (2012). Learning 21st-century skills requires teaching 21st-century skills. RAND Corporation. [https://www.rand.org/pubs/occasional\\_papers/OP315.html](https://www.rand.org/pubs/occasional_papers/OP315.html)
- 38) Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503-524). Information Age Publishing.
- 39) Shute, V. J., & Ventura, M. (2013). *Stealth assessment: Measuring and supporting learning in video games*. MIT Press.
- 40) Shute, V. J., Ventura, M., & Kim, Y. J. (2016). Assessment and learning of 21st century skills. In A. A. Rupp & J. P. Leighton (Eds.), *The handbook of cognition and assessment: Frameworks, methodologies, and applications* (pp. 407-436). Wiley.
- 41) Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4(4), 295-312. [https://doi.org/10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5)
- 42) Tiefenthaler, D., & Yau, J. Y. K. (2020). Utilizing learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961-1990. <https://doi.org/10.1007/s11423-020-09725-y>
- 43) Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.
- 44) Torrance, E. P. (2000). The millennium: A time for looking forward and looking backward. *Creative Child and Adult Quarterly*, 25(4), 1-8.
- 45) Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Jossey-Bass.
- 46) Udeozor, C., Abegao, F. R., & Glassey, J. (2023). Digital game-based learning for chemical engineering education: A systematic literature review. *Computer Applications in Engineering Education*, 31(2), 492-516. <https://doi.org/10.1002/cae.22598>
- 47) Villaruel, J. A., Ramos, M. C., & Dela Cruz, A. D. (2022). Digital game preference and usage among Filipino junior high school students: Implications for instructional design. *Philippine Journal of Educational Technology*, 11(1), 35-49.
- 48) Vivekanandan, K. (2019). 21st century skills and tools in education. *International Journal of Research in Engineering, Science and Management*, 2(9), 456-459.
- 49) Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). ASCD.