

Validity Item Analysis on the Released PISA 2022 Creative Thinking Assessment in the Philippine Context

DOI: 10.5281/zenodo.14820168

Dave Louis Morales Baptista¹, Lester Reboroso Saldua², Winston Mendoza Padre³

Mariano Marcos State University, Philippines^{1,2,3}

0009-0007-2083-1863¹, 0009-0002-0882-4015², 0009-0005-2470-9051³ | dmbaptista@mmsu.edu.ph¹

Abstract

The study assessed the validity of the released items of the Programme for International Student Assessment (PISA) 2022 Creative Thinking Assessment for Filipino students, specifically its face, content, and construct validity and considering their alignment with the Philippine curriculum. Using a descriptive research approach, employing documentary analysis as its primary method, the study found that the assessment exhibited a lack of face validity due to unclear graphics. Content validity was compromised by language barriers, ambiguous instructions, and a mismatch between the assessment's focus and the Philippine curriculum. Construct validity was challenged by the assessment's emphasis on Westernized concepts of creativity and originality, which potentially disadvantaging Filipino students. While aiming to assess a range of creative thinking skills, the study concluded that the PISA 2022 Creative Thinking assessment did not fully consider the cultural perspectives of students from diverse contexts. The findings highlight the need for a more culturally sensitive and contextually relevant approach to assessment design by ensuring that international assessments are fair, equitable, and inclusive for all students.

Keywords: PISA, Creative Thinking, Assessment Validity, Validity, Cultural Sensitivity, Philippine Education, DepEd, Face Validity, Construct Validity, Content Validity

Introduction

The Philippines' recent performance in the Programme for International Student Assessment (PISA), initiated by the Organization for Economic Cooperation and Development (OECD), in their 2022 creative thinking assessment has raised concerns (Department of Education, 2022). The country ranked second to last among 65 participating nations. It highlighted a significant gap in creative thinking skills among Filipino students. Thus, this alarming result prompted the proponents of this study to investigate the potential reasons behind this performance and seeks to understand the factors contributing to the observed disparity. While existing research has explored the importance of creative thinking skills in various educational contexts, a critical gap remains in understanding the validity of international assessments like PISA in diverse cultural contexts. The current study aims to address this gap by focusing on the PISA 2022 Creative Thinking assessment tool and its application within the Philippine education system.

This research posits that the PISA 2022 Creative Thinking Assessment tool, despite its intention to evaluate a range of creative thinking skills, may not fully consider the cultural perspectives of students from diverse contexts, particularly those from the Philippines. This potential flaw in the assessment's validity could be attributed to its emphasis on Westernized concepts of originality, creativity, and design, potentially disadvantaging students who have limited exposure to these concepts or hold different cultural perspectives. It aimed to investigate the validity of the PISA 2022 Creative Thinking Assessment tool in the Philippine context by examining its face, content, and construct validity. The findings shed light on the potential limitations of the assessment tool and its implications for evaluating creative thinking skills among Filipino students, ultimately contributing to a better understanding of the assessment's effectiveness in culturally diverse settings.

Specifically, it seeks to answer the following research questions:

1. What are the released items on the PISA 2022 creative thinking assessment?
2. Is the PISA 2022 creative thinking assessment tool valid in the context of Philippine education system in terms of:
 - a. face;
 - b. content; and
 - c. construct validity?

Significance of the Study

The Philippines has been consistent in having a low score of students in the examination administered by PISA. Despite the actions of the government in combating the illiteracy rate in the country, data from PISA showed that the Philippines has still a long way to go in being with the other countries in terms of literacy. However, PISA is administered by an international organization which could be seen as a loophole in the process.

This study on the validity of the PISA 2022 creative thinking assessment in the Philippine context will benefit several stakeholders. Foremost, the Department of Education (DepEd) can gain valuable insights into the alignment of the PISA assessment with the K+12 curriculum and teaching practices. This information can inform policy decisions regarding curriculum development, teacher training, and assessment strategies to better foster creative thinking skills among Filipino students. Teachers will also gain a deeper understanding of the PISA assessment's expectations, as it can help them better prepare their students for future assessments administered by PISA and adapt their teaching methods to effectively develop creative thinking skills. This study will also contribute to the growing body of research on the validity of international assessments in diverse cultural contexts. It can provide them with valuable data and insights that can inform future research on assessment development and adaptation. While the study itself won't directly benefit students, the findings will contribute to improving the quality of education and assessment practices in the Philippines. This, in turn, can lead to better learning experiences and a more supportive environment for developing creative thinking skills. Most importantly, the findings of the study inform the OECD, who administers PISA and other international organizations, about the challenges and opportunities in implementing standardized assessments across diverse cultural contexts. This can contribute to the development of more culturally sensitive and contextually appropriate assessments in the future.

Scope and Delimitation of the Study

This research study focused on assessing the validity of the PISA 2022 creative thinking assessment tool within the Philippine context. The scope was limited to evaluating the assessment's effectiveness in measuring creative thinking skills as defined by the PISA 2022 framework. The study's motivation stems from the Philippines' low ranking in the PISA 2022 creative thinking assessment, and the need to understand the assessment's applicability to the Philippine education system. The research utilized data provided by the OECD, including the actual assessment tools and released results, to analyze the assessment's face, content, and construct validity. The study was delimited in several ways. It did not examine other PISA domains, such as reading, mathematics, or science. While the study acknowledges the Philippines' unique cultural context, it did not conduct a comparative analysis of the assessment's reliability across different cultural contexts. The study did not involve any human respondents, such as students, teachers, or experts, and did not utilize statistical tools for data analysis. Instead, the research relied on qualitative and documentary analysis methods to investigate the assessment's validity in the Philippine context.

Literature Review

PISA 2022 Creative Thinking: Creative Minds, Creative Schools

The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by testing 15-year-old students' performance in reading, mathematics, and science. The PISA 2022 Creative Thinking Assessment was released by OECD in 2024 titled, *Volume III: Creative Minds, Creative Schools*. This volume examines the creative thinking abilities of students in 66 participating countries and economies, specifically their capabilities to generate original and diverse ideas (OECD, 2024). It also investigates how student performance and attitudes related to creative thinking differ across and within nations, as well as in relation to various student and school-level factors. The volume also offers an insight into student participation in creative activities, the availability of opportunities for creative thinking across schools and socioeconomic groups, and the link between these factors and student outcomes.

The assessment was designed to assess students' creative thinking skills in four domains: written expression, visual expression, social problem solving, and scientific problem solving to understand how students generate diverse ideas, evaluate and improve existing ideas, and approach problem-solving creatively. PISA is unique because it is a large-scale, internationally comparative assessment that focuses on students' ability to apply their knowledge and skills to real-world situations. The assessment is not designed to measure students' mastery of a specific curriculum, but rather their ability to think critically and creatively. It is particularly important because it highlights the growing importance of creative thinking skills in the 21st century (Paulino, 2019). Creative thinking is essential for success in a wide range of fields, including science, technology, engineering, and the arts. It is also important for attaining success in everyday life, as it allows people to adapt to new situations, solve problems, and come up with innovative solutions.

PISA defines creative thinking as the ability to engage productively in the generation, evaluation and improvement of ideas that can result in original and effective solutions, advances in knowledge and impactful expressions of imagination (OECD, 2024). Specifically, the PISA 2022 assessment of creative thinking utilized a three-stage ideation process model (see Figure 1): generating diverse ideas, generating creative ideas, and evaluating and improving ideas. This model aligns with the PISA definition of creative thinking and focuses on cognitive skills applicable to classroom settings.

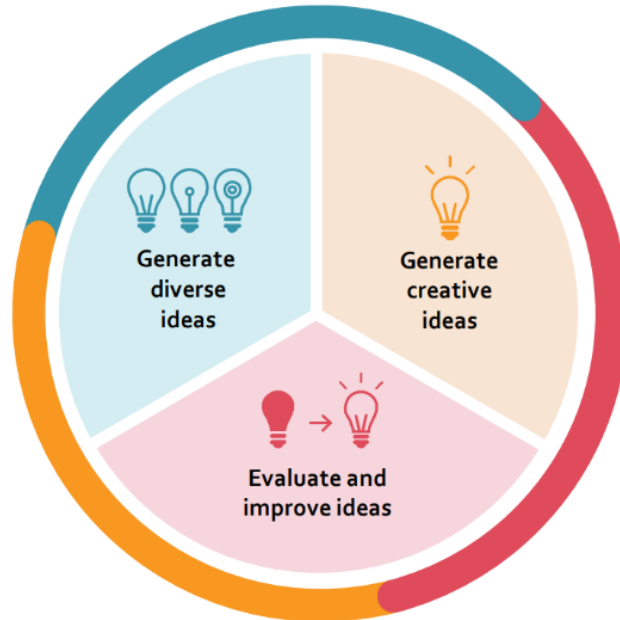


Figure 1: Creative Thinking Ideation Processes

Source: OECD (2022), *Thinking Outside the Box: The PISA 2022 Creative Thinking Assessment*

The assessment comprised 12 items focused on generating diverse ideas, emphasizing ideational fluency and flexibility; 11 items on generating creative ideas, measuring originality as a proxy for novel thinking; and 9 items on evaluating and improving ideas, assessing the ability to identify limitations and enhance solutions. In the released volume by OECD titled, *Thinking Outside the Box: The PISA 2022 Creative Thinking Assessment*, the three ideation processes are defined as follows:

Generate Diverse Ideas assesses a student's ability to produce a variety of different ideas, emphasizing both the quantity (fluency) and the distinctiveness (flexibility) of their thinking. Generate Creative Ideas, on the other hand, measures originality, defined as the statistical infrequency of an idea compared to those of peers (Guilford, 1950). While true novelty is difficult to assess in 15-year-olds, originality serves as a practical indicator of unconventional thinking. Lastly, Evaluate and Improve Ideas evaluates a student's capacity to critique their ideas, identify shortcomings, and enhance their solutions. This iterative process aims to produce more effective and efficient outcomes. PISA's definition of creative thinking centers on ideation processes applicable across diverse learning and problem-solving contexts. These encompass learning contexts like creative writing and the visual and performing arts, that require imagination and expression, and also those in which generating and improving upon ideas is functional to investigating problems or phenomena, or to designing innovative solutions. In research question 1, the researchers describe selected items from nine released items of the 18 creative thinking units developed for the PISA 2022 assessment; at least one unit from each domain context is included. Each unit's context and scenario are briefly described, followed by a screenshot and description of sample items. For certain items, genuine student responses and item-specific coding criteria are also presented.

Considering the age of the PISA test-takers and the allotted testing time, the PISA 2022 creative thinking test tasks were structured within four distinct domain contexts: written expression (i.e., communicating ideas and imagination through written language); visual expression (i.e., communicating ideas and imagination through various media); social problem-solving (i.e., understanding different perspectives, addressing others' needs, and finding innovative solutions); and scientific problem-solving (i.e., generating new ideas, designing experiments, and developing new methods or inventions). The test's item distribution across these four contexts was as follows: 12 items in written expression, 4 in visual expression, 10 in social problem-solving, and 6 in scientific problem-solving.

Philippines' Performance in PISA 2022 Creative Thinking Assessment

The Philippines ranked second to last among 65 participating nations in the PISA 2022 Creative Thinking Assessment (Department of Education, 2022), with a mean score of 14 points (Helpline PH, 2024). This score is significantly below the OECD average of 33 points (Parungao, 2024).

In the said assessment, students were asked to think of original solutions for simple tasks and familiar problems, such as coming up with an interesting story idea or thinking of different ways to conduct an awareness-raising campaign in school, among others (Helpline PH, 2024).

The performance gap highlights a significant gap in creative thinking skills among Filipino students. It further suggests that the Philippine education system is not adequately preparing students for the demands of the 21st-century workforce, which requires a high level of creativity and problem-solving skills (Paulino, 2019). Therefore, the findings also suggest a need to re-evaluate current teaching practices and curriculum to foster creative thinking skills among students. The PISA 2022 results indicate that Filipino students struggle with generating diverse ideas, generating creative ideas, and evaluating and improving ideas. This suggests a need for more emphasis on creative problem-solving, critical thinking, and innovation in the teaching-learning process. In the other areas, the Philippines ranked as the sixth lowest among the 81 countries and economies that participated in the previous studies of PISA, specifically in reading, mathematics, and science (GMA News Online, 2024). It underscores the need for a comprehensive approach to improving the quality of education in the Philippines (Licci, n.d.). On the other hand, the PISA 2022 Creative Thinking Assessment is not valid as it may not be culturally sensitive enough to accurately measure the creative thinking skills of Filipino students.

Validity in Assessment

Validity ensures that assessments accurately measure what they intend to measure as it provides valuable insights into the individual's abilities and knowledge. There are several types of validity—each focusing on different aspects of an assessment: face, content, and construct validity. Face validity refers to the superficial appearance of an assessment tool. It is essentially the first impression of an assessment yet often overlooked component of assessment construction (Frantz & Holmgren, 2019). While face validity is simple and intuitive concept, it plays a significant role in assessments as it influences its legitimacy and the test takers' willingness to engage in the test. Research suggests that assessments with high face validity are more likely to yield high-quality data because test-takers are more likely to engage with the assessment and respond thoughtfully because they are less likely to feel apprehensive or confused about the assessment's purpose (Frantz & Holmgren, 2019). Furthermore, face validity plays a crucial role in shaping the test-taker's experience. When test-takers perceive an assessment as fair, relevant, and understandable, they are more likely to feel motivated and engaged, leading to a more positive assessment experience as it reduces anxiety and promoting a sense of ownership over the assessment process.

On the other hand, content validity ensures that the content of an assessment accurately reflects the knowledge and skills being measured. It ensures that the assessment accurately reflects the curriculum and learning objectives. This highlights the importance of assessments covering key concepts, skills and knowledge that students are expected to have learned. Obilor (2018) enumerated challenges in achieving content validity including unclear directions, vocabulary, poor constructed test items, difficulty level of items, influence of extraneous factors, and inappropriate time limits. Assessments must accurately reflect the curriculum and provide a valid measure of student learning. Lastly, construct validity refers to assessment items being carefully designed to reflect the specific construction being measured. In the context of student assessments, construct validity is essential for ensuring that the assessment accurately reflects the theoretical construct being measured, such as critical thinking, problem-solving, or creativity. A study by Zuo and their team (2021) found that development and validation of assessments with strong construct validity ensures that assessments serve their intended purpose as it supports student learning, guiding instruction, and equitable and effective educational practices.

Creative Thinking in the Filipino Context

The Filipino context offers a unique perspective on the relationship between culture and creativity. Its values and traditions influence creative thinking and expression formatively how individuals approach problem-solving, collaborate, and communicate their ideas. Storytelling, music, art, and other cultural forms serve as powerful tools in assessing imagination, critical thinking, and emotional expression. Studies on collectivist cultures, particularly in East Asia, suggests that the Filipino value prefer practical applications of creativity over novelty aligning with its collaborative nature (Al-Haddad & Abu-Saad, 2018). Similarly, *pakikisama* (i.e., harmony and social acceptance) appears to shape creative expression as it encourages sensitivity to social contexts that leads to a preference for socially acceptable ideas (Chen & Chen, 2019). The influence of *hiya* (i.e., modesty and shame) on creative expression, making individuals hesitant to express themselves freely or take risks, is also supported by research on cultures emphasizing shame (Frantz & Holmgren, 2019). Likewise, engagement with visual art has been shown to stimulate creativity, imagination, and critical thinking skills (Larkin & Csikszentmihalyi, 2014).

Research also suggests that international assessments like PISA sometimes fail to adequately account for cultural differences in creative thinking and expression leading to biased results (Sandoval, 2019). Further research is needed to understand how the PISA assessment could be adapted or modified to better reflect the specific cultural contexts and creative expression styles of Filipino students.

Methodology

This chapter outlines the methodological approach taken in the study. It encompasses the research design, the sources of data, and the specific procedures employed in data collection. It also provided a detailed account of how the study was structured to answer the research questions and how the necessary data was gathered to support the analysis.

Research Design

This study employed a descriptive research approach by utilizing documentary analysis as its primary method. The study aimed to understand the assessment's effectiveness in evaluating student creativity across different cultural contexts and to demonstrate consistency in evaluating student responses across diverse cultural contexts. According to Frey (2018), document analysis is a type of qualitative research that applies a systematic process to analyze documentary evidence and answer specific research questions. The study aimed to understand their effectiveness in evaluating student responses within the unique educational and cultural landscape of the Philippines. This allowed for a thorough understanding of the assessment's strengths and limitations in a specific cultural context; hence, contributing to a more holistic evaluation of its overall validity. It did not involve any human respondents and did not utilize statistical tools for data analysis. Instead, the researchers relied on qualitative and documentary analysis on the released PISA 2022 Creative Thinking assessment items to investigate validity of assessment in the Philippine context.

The study's motivation stems from the Philippines' low ranking in the PISA 2022 creative thinking assessment, and the need to understand the assessment's applicability to the Philippine education system. The research utilized data provided by the OECD, including the actual assessment tools and released results, to analyze its face, content, and construct validity. The analysis identified patterns, themes, and potential limitations within the assessment items and the provided document. The findings were presented in a descriptive and narrative format, which highlights key insights and observations related to the assessment's validity in the Philippine context. The study was delimited in several ways. It did not examine other PISA domains, such as reading, mathematics, or science. While the study acknowledges the Philippines' unique cultural context, it did not conduct a comparative analysis of the assessment's reliability across different cultural contexts.

Research Instrument

The main source of data was the released PISA 2022 Creative Thinking assessment items, which were analyzed to determine their respective domain context, ideation process, item format, coding criteria, and proficiency level. The assessment explores creative thinking across four distinct domains: written expression, visual expression, social problem-solving, and scientific problem-solving. Within each domain, students are assessed on their ability to engage in three key ideation processes: generate diverse ideas, generate creative ideas, and evaluate and improve ideas. The assessment categorizes students' creative thinking abilities into different proficiency levels, ranging from below proficiency to advanced proficiency to understand the strengths and areas for improvement for students.

Additionally, a Qualitative Analysis Framework was used to analyze the data extracted from the released items. This framework assessed the assessment's face validity by examining visual elements like graphic representation, and image alt text to assess their potential impact on Filipino students' perception and engagement, content validity by analyzing the language used in the instructions and tasks to identify the potential language barriers, and assessing the clarity and specificity of the instructions, and construct validity by investigating the alignment between the assessment items and the Philippine curriculum, which considers the specific skills and knowledge assessed by each item, and evaluating the potential for unrealistic expectations or cultural bias.

Data Gathering Procedure

The data gathering procedure involved a documentary analysis of the released items. First, the document was thoroughly reviewed to identify and extract relevant information related to the research questions. This includes identifying the released items, their descriptions, and the assessment criteria used to evaluate student responses. Then, each item was analyzed individually, focusing on face, content, and construct validity. Face validity was assessed by examining visual elements such as graphic representation, and layout design. Content validity was assessed by examining the language used in the instructions and tasks, identifying potential language barriers, and assessing the clarity and specificity of the instructions. Construct validity was assessed by investigating the alignment between the assessment items and the Philippine curriculum, considering the specific skills and knowledge assessed by each item and evaluating the potential for unrealistic expectations or cultural bias.

All extracted data and observations were documented systematically, using a standardized format to ensure clarity and organization. This documentation served as the basis for the qualitative analysis of the assessment's validity.

Sampling Technique

Purposive sampling was used which involves all items as they are all relevant to the research question. Furthermore, it was used as a sampling technique because the research focused on analyzing the released items to understand the assessment's validity in the Philippine context. The goal is to understand the assessment's validity, not to generalize findings to all possible PISA items.

Findings and Discussion

This section presents a comprehensive understanding of the results and discussion of the objective of the research study.

Released Items on the PISA 2022 Creative Thinking Assessment

This section presents descriptive and technical information for the released item from the PISA 2022 Creative Thinking assessment. It focuses on evaluating students' creative thinking skills across various domains including written expression, visual expression, social problem-solving and scientific problem solving. This assessment aims to understand how students generate diverse ideas, evaluate and improve existing ideas and approach problem-solving creatively. For items under written expression, test-takers were asked to express their imagination in various written formats, specifically captioning an image, proposing ideas for a short story using a given text or visual as inspiration, and writing a short dialogue between characters for a movie or comic book plot.

Illustration Titles. As shown in *Figure 2*, the test-takers were asked to come up with an original title for an abstract illustration.

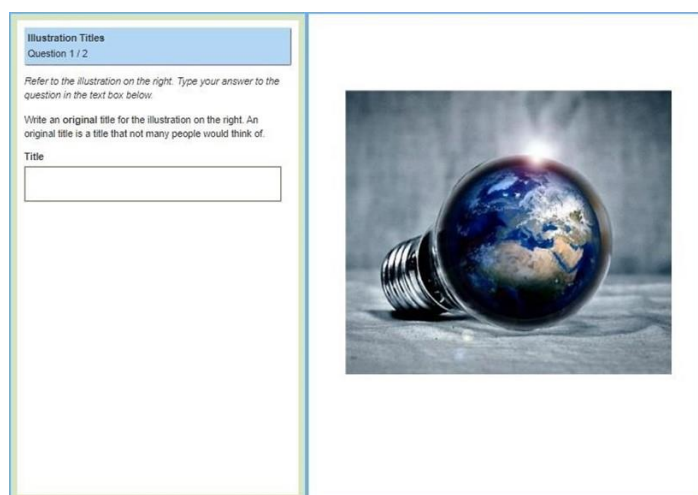


Figure 2: Illustration Titles [Released item #1 (DT300Q02)]

Specifically, the test-takers are presented with an illustration and instructed to create an original title for it. The title must be original and unexpected (i.e., not many people would think of) by demonstrating creative thinking and a distinct perspective on the visual content. The title is to be written in the designated text box. For the visual cue, they are presented with a visual image of a lightbulb with an image of the planet Earth inside. The coding criteria for this item emphasize the importance of originality in the titles students create. The scoring guidelines also highlight the importance of considering the structure of the title including technicalities like the use of punctuation, capitalization and other grammatical elements to create distinct meanings.

Robot Story. The item presented an open-ended creative work task. Specifically, the test-takers were asked to generate two distinct film story ideas, each focusing on the interaction between a human character named Leo and an intelligent robot named Rob (see *Figure 3*).

Robot Story
Question 1 / 3

Type your answers to the question in the boxes below.

Write 2 different film story ideas for a film about a human named Leo who is interacting with an intelligent robot named Rob. You do not need to write the full film story, but only explain what the film will be about. The two stories should be as different from each other as possible.

We recommend that you spend no longer than 7 minutes on this question, and use no more than 8 sentences for each story idea.

Film story idea 1

Film story idea 2

ROBOT STORY



Rob Leo

Figure 3: Robot Story [Released item #1 (DT570Q01)]

The instruction explicitly encouraged the test-takers not to write the full film story, but only explain what the film will be about by exploring the relationship between these two characters. Below the instructions, there are two rectangular text boxes labeled "Film story idea 1" and "Film story idea 2", respectively. These boxes are intended for students to write their two distinct story ideas. A time limit of seven minutes and a sentence limit of eight sentences per story were recommended to simulate real-world creative constraints. The coding criteria emphasizes that the answers must be distinct in terms of basic plot and use different story elements to change the representation of those ideas including changes in setting, character dynamics or the way the story unfolds.

Space Comic. As shown in *Figure 4*, the test-takers were asked to write a dialogue for a comic strip that shows the Sun and the Earth in conversation with each other.

Space Comic
Question 1 / 2

Refer to the comic strip on the right. Use the text boxes to answer the question.

The text in a comic strip is intended to convey the dialogue between characters. Write a comic strip using the illustrations and text boxes on the right. Your comic strip should be original in the sense that not many people would think of it.

Use as many of the available text boxes on the right as needed.

We recommend that you spend no longer than 5 minutes on this question.

Text 1

Text 2

Text 3

Text 4

Text 5

Text 6

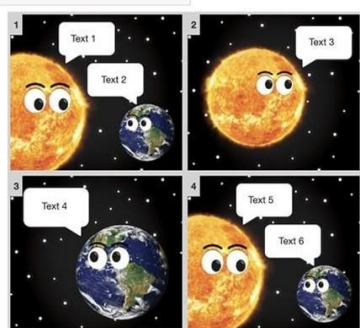


Figure 4: Space comic [Released item #1 (DT240Q01)]

Specifically, they are presented with a four-panel comic strip template which contains the Sun and the Earth as characters, with speech balloons. Then, they must come up with an original comic strip using the provided elements and ensure that the dialogue within the text boxes conveys a unique and unexpected story (i.e., not many people would think of it). There are six empty panels with text boxes labeled "Text 1," "Text 2," "Text 3," etc. These empty text boxes are intended for students to write the dialogue between the Sun and the Earth, creating a visual space for them to express their creative ideas. A time limit of five minutes is recommended for completing the task. The scoring guidelines focuses on the appropriateness and sufficiency of potential responses. The dialogue must focus on conventional themes like the Sun's heat, Earth's ability to sustain life basically character dynamics, the use of metaphors, similes or other figurative language to create a more evocative meaning and also accounts grammatical

elements to create a distinct meaning. 2983. It is a single item in which the test-takers were presented with a book cover image. They are asked to create an original story for a book titled "2983" (see Figure 5). The item is classified under generate creative idea, as students must associate the number 2983 to a relevant detail in their story idea. A time limit of five minutes and a sentence limit of eight sentences are recommended to guide the response.

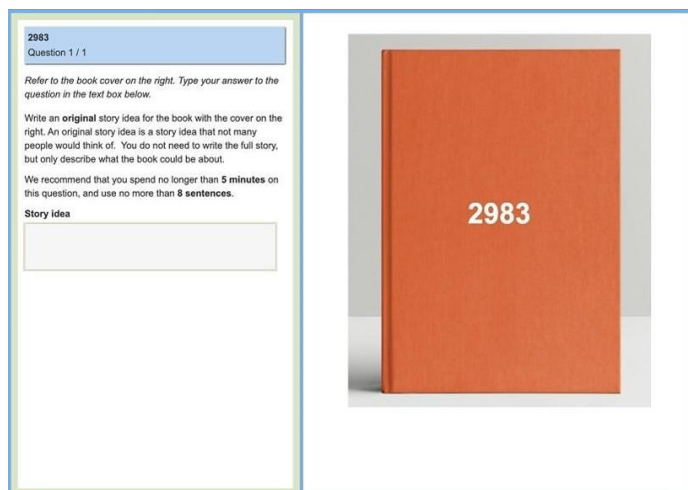


Figure 5: 2983 [Released item #1 (DT370Q01)]

The focus of the item underscores two key aspects. First, the story must be distinct in terms of their subject matter and, second, is the use of literary devices to change the representation of those ideas. To achieve full credit, the response must correspond to an original theme. Conventional (i.e. non-original) themes included: stories about the future of humanity set in the year 2983; or stories in which the number 2983 identifies a person, a place or an object. Responses that corresponded with conventional themes were awarded partial credit unless combined with an innovative approach or implementation. For example, an unconventional reference to the number 2983 in the story was its use as a code for unlocking a device. For items under visual expression, test-takers used a basic graphic tool to construct visual compositions, selecting from a provided library of images and shapes. The tool allowed for resizing, rotation, and color changes of the shape elements. They also created visual designs for various purposes, including logos, event posters, or merchandise designs.

Science Fair Poster #1. As shown in Figure 6, the test-takers were asked to create an original concept of a poster for the science fair that represents the theme "Life in Deep Space."

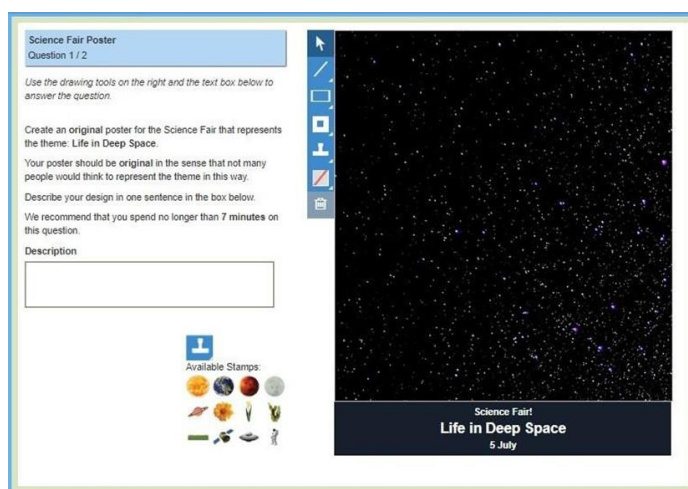


Figure 6: Science Fair Poster [Released item #1 (DT200Q01)]

They are also provided with digital drawing tools and available stamps to create a visually appealing and effective poster. The design must be unique and unconventional (i.e., original in the sense that not many people would think to represent the theme in this way). A concise, one-sentence description of the design is required in the provided text box. A time limit of seven minutes is recommended.

To achieve full credit, the answer must correspond to an original theme. Even with the use of similar basic visual elements test-takers are given merit based on how different they were used. This includes shapes, colors or stamps or arrangement of those elements. The theme must focus on Earth or space exploration as well as challenges of life in space or the beauty of the universe.

Conventional (i.e. non-original) themes for this item as well as coded example responses; responses that corresponded to conventional themes were awarded partial credit, unless combined with an innovative approach or implementation.

Science Fair Poster #2. In the second item in the Science Fair Poster unit, the test takers are provided with a pre-existing poster (i.e., the Sun and one planet) and are given access to digital drawing tools and a text box (see Figure 7). They are instructed to modify the poster to better reflect the theme "Life in Deep Space," using original and unconventional design changes, and making sure that the ideas are appropriate and adequate.

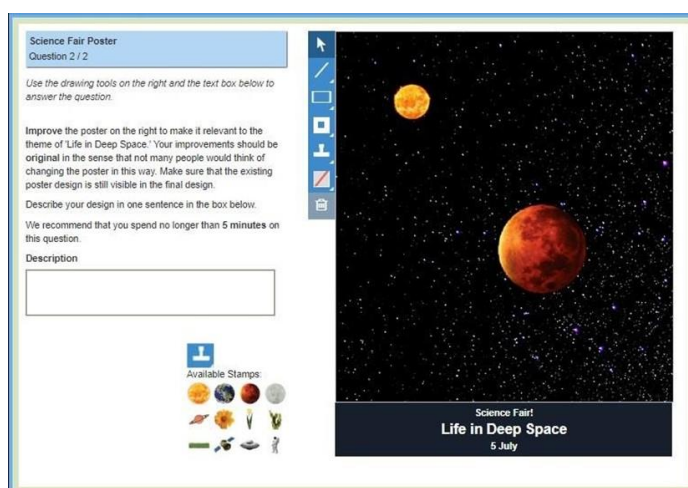


Figure 7: Science Fair Poster [Released item #2 (DT200Q02)]

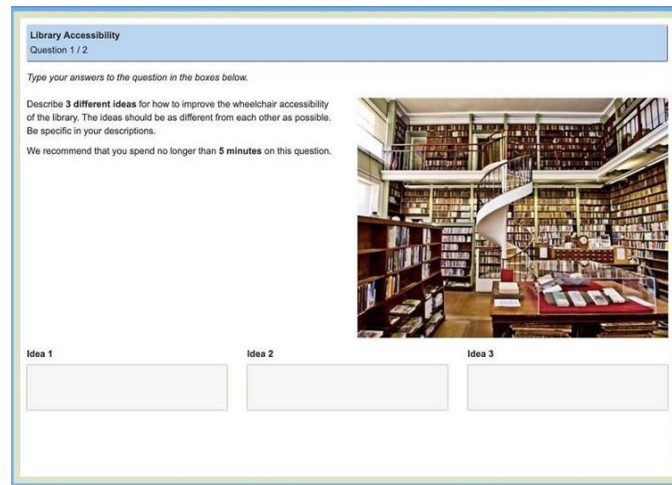
The original poster design must remain visible in the revised version. A one-sentence description of the modifications is required in the text box provided. The task is recommended to be completed within five minutes.

The coding criteria for this item focuses on how "distinct" the poster design is as well as their connection to the theme. The modifications must use a variety of shapes, colors, and stamps, or does it rely primarily on a limited number of elements. The modification must effectively communicate the chosen theme and create a strong visual impact.

To achieve full credit, they must modify the poster with an original idea. Modifications that corresponded to conventional (i.e. non-original) themed ideas were awarded partial credit, unless combined with an innovative approach or implementation.

Social problem solving can range from the small-scale, personal and interpersonal problems of individuals to wider school, community or even global problems. In the PISA 2022 creative thinking assessment, test takers must open problems that focused on issues affecting different groups within society (e.g. wheelchair users) or affecting society at large (e.g. the collection and use of waste materials).

Library Accessibility #1. This item asked test-takers to generate three distinct and substantially different ideas for enhancing wheelchair accessibility within a library setting shown in Figure 8. The descriptions must be specific and detailed. A five-minute time limit is suggested for completing the task, with each idea recorded in a separate designated box.



Library Accessibility
Question 1 / 2

Type your answers to the question in the boxes below.

Describe 3 different ideas for how to improve the wheelchair accessibility of the library. The ideas should be as different from each other as possible. Be specific in your descriptions.

We recommend that you spend no longer than 5 minutes on this question.

Idea 1

Idea 2

Idea 3

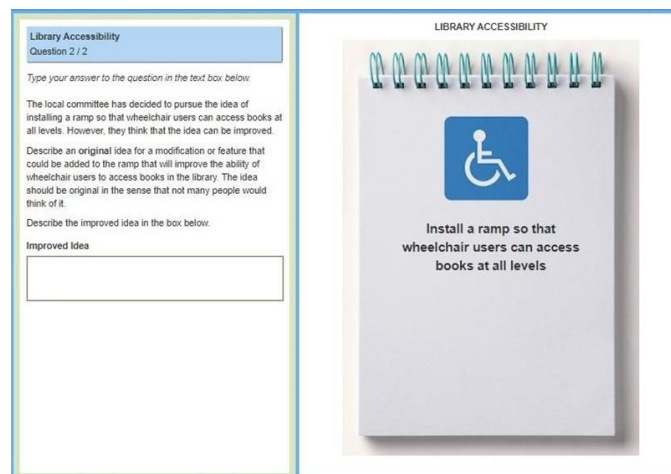
Figure 8: Library Accessibility [Released item #1 (DT500Q01)]

This item focuses on how efficient the solutions to the problem are using different tools, strategies or resources. Specifically, on physical modifications to the library involves adding ramps, modifying the staircase, adding an elevator, having only one floor, having lower shelves, or having a special section for wheelchair users, providing human assistance to wheelchair users involving hiring staff or adding volunteers to retrieve library materials for customers or deliver materials to customers and providing technological assistance mechanisms such as using a book delivery service, having a digital library online access or assistive technology to help wheelchair users to navigate the library.

To achieve full credit, they had to provide three appropriate ideas that are sufficiently different; if students provided only two different ideas, then their response achieved partial credit.

Library Accessibility #2. In this item, the test takers were presented with a scenario where a library committee is considering installing a ramp for wheelchair accessibility but seeks to enhance the design (see Figure 9). They are instructed to propose an original and innovative modification or feature for the ramp that would further improve book access for wheelchair users. The proposed idea should be original in the sense that not many people would think of it by focusing on enhancing the ramp's functionality.

The prompt encourages them to think creatively and come up with a solution that goes beyond the basic concept of a ramp by pushing them to consider how to make the library more inclusive and welcoming for people with disabilities. The description of the improved idea is to be provided in the designated box.



Library Accessibility
Question 2 / 2

Type your answer to the question in the text box below.

The local committee has decided to pursue the idea of installing a ramp so that wheelchair users can access books at all levels. However, they think that the idea can be improved.

Describe an original idea for a modification or feature that could be added to the ramp that will improve the ability of wheelchair users to access books in the library. The idea should be original in the sense that not many people would think of it.

Describe the improved idea in the box below.

Improved Idea

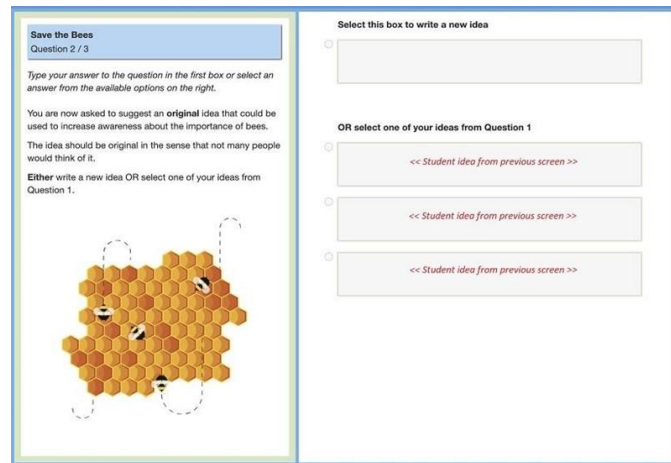
LIBRARY ACCESSIBILITY

Install a ramp so that wheelchair users can access books at all levels

Figure 9: Library Accessibility [Released item #2 (DT500Q02)]

To achieve full credit, the response must correspond to an original improvement theme. Responses that corresponded with conventional themes were awarded partial credit, unless combined with an innovative approach or implementation.

Save the Bees. The item is under the ideation process of generating creative ideas where the test-takers were instructed to propose an original idea for raising awareness about the importance of bees (see Figure 10). The idea must be original and unconventional. They are also given the option to either generate a new idea or select one of their previously generated ideas from the previous question from the same unit. The response is to be entered into a designated box.



Save the Bees
Question 2 / 3

Type your answer to the question in the first box or select an answer from the available options on the right.

You are now asked to suggest an **original** idea that could be used to increase awareness about the importance of bees. The idea should be original in the sense that not many people would think of it.

Either write a new idea OR select one of your ideas from Question 1.

Select this box to write a new idea

OR select one of your ideas from Question 1

<< Student idea from previous screen >>

<< Student idea from previous screen >>

<< Student idea from previous screen >>

Figure 10: Save the Bees [Released item #1 (DT400Q02)]

The scoring guideline indicates that the methods or content must involve creating educational materials, giving presentations or organizing events, creating posters, videos, or flyers, setting up a bee observation hive, organizing a beekeeping workshop or planting bee-friendly flowers. Overall, this item encourages students to think creatively about how to communicate important information to a wider audience.

Carpooling. As shown in Figure 11, the test-takers must think of an original idea to further incentivize carpooling. Specifically, they were tasked to propose an original and innovative expansion or improvement to a carpooling initiative aimed at reducing air pollution and traffic congestion. The initiative involves offering discounts on fuel and tolls to carpoolers. The proposed idea should be unique and go beyond existing approaches. The description of the improved idea is to be provided in the designated box.



Carpooling
Question 1 / 1

Type your answer to the question in the box below.

You are part of a team identifying creative solutions for issues faced by communities around the world.

To encourage carpooling (travel in the same vehicle together) and thus reduce air pollution and the number of vehicles on the road, some countries provide discounts on fuel and tolls to people who carpool. Think of an original way that the initiative to promote carpooling can be expanded and improved upon.

Describe the improved idea in the box below.

Improved Idea

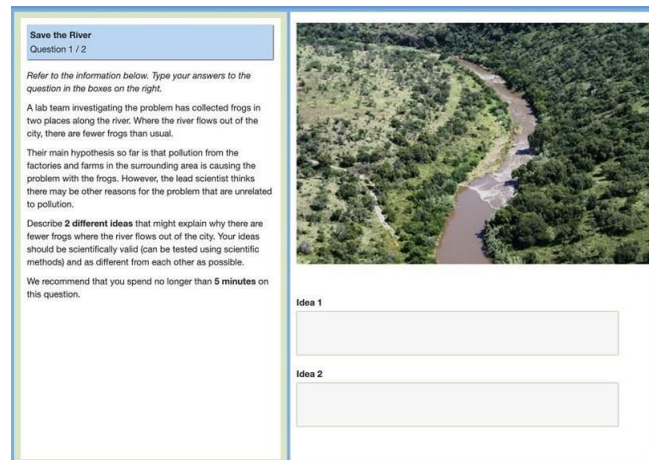
CARPOOLING

Figure 11: Carpooling [Released item #1 (DT630Q01)]

The item is classified as an "evaluate and improve ideas" item because granting discounts on fuel or tolls are existing incentives that need to be further strengthened. To achieve full credit, the response of the test-takers must correspond to an original idea theme. For this item, there is only one conventional (i.e. non-original) theme: introducing additional financial incentives (e.g., making the shared purchase of cars more affordable.) The responses that corresponded to the conventional theme were awarded partial credit unless combined with an innovative approach or implementation.

The PISA creative thinking test included tasks where the test-takers explored open-ended scientific or engineering problems. While creative thinking in scientific contexts is connected to scientific inquiry, the tasks in this domain context differed fundamentally from the PISA scientific literacy tasks. In the creative thinking test, the test-takers were required to generate multiple ideas or solutions, or a single original idea or solution, for an open problem without a predetermined "correct" answer. For instance, in a task where students were asked to propose explanations for a given phenomenon, they were rewarded for offering multiple plausible ideas, even if these ideas did not represent the "right" explanation.

Save the River #1. As shown in *Figure 12*, the test-takers are presented with a presented with a scenario involving a lab team investigating a decline in frog populations near a city. They are instructed to propose two distinct, scientifically valid explanations for this decline, focusing on factors unrelated to pollution. The ideas should be scientifically testable and as different from each other as possible. A time limit of five minutes is recommended for completing the task. A description of the idea is provided in the designated box. The item requires students to consider the principles of scientific inquiry and to propose testable explanations for the observed phenomenon.



Save the River
Question 1 / 2

Refer to the information below. Type your answers to the question in the boxes on the right.

A lab team investigating the problem has collected frogs in two places along the river. Where the river flows out of the city, there are fewer frogs than usual.

Their main hypothesis so far is that pollution from the factories and farms in the surrounding area is causing the problem with the frogs. However, the lead scientist thinks there may be other reasons for the problem that are unrelated to pollution.

Describe 2 different ideas that might explain why there are fewer frogs where the river flows out of the city. Your ideas should be scientifically valid (can be tested using scientific methods) and as different from each other as possible.

We recommend that you spend no longer than 5 minutes on this question.

Idea 1

Idea 2

Figure 12: Save the River [Released item #1 (DT690Q01)]

Students were explicitly instructed to think of causes other than pollution. Students could only achieve full credit or no credit for this item, as it required only two different ideas. The item-specific coding criteria provided several different possible causes of the problem.

The guidelines categorize elements as to what constitutes a critical and creative answer are as follow: 1) changes to the water habitat; 2) changes in surrounding fauna; 3) changes to the local flora; 4) changes to from themselves; and 5) changes to the behavior or activities of humans in the area. The scoring guidelines ensure that students are rewarded for their ability to think outside the box and to consider the needs of a specific group within society.

Save the River #2. The second item of the Save the River unit asks students to improve a proposed experiment aiming to test whether pollution is the cause of the problem with the declining frog population (see *Figure 13*). Specifically, the test-takers are presented with a proposed experiment to investigate the impact of pollution from factories on frog populations. They are instructed to propose an original and innovative improvement to the experiment that would provide more conclusive evidence regarding the role of chemical pollution. The proposed improvement should be unique and go beyond standard experimental approaches. The description of the improved experiment idea is provided in the designated box.

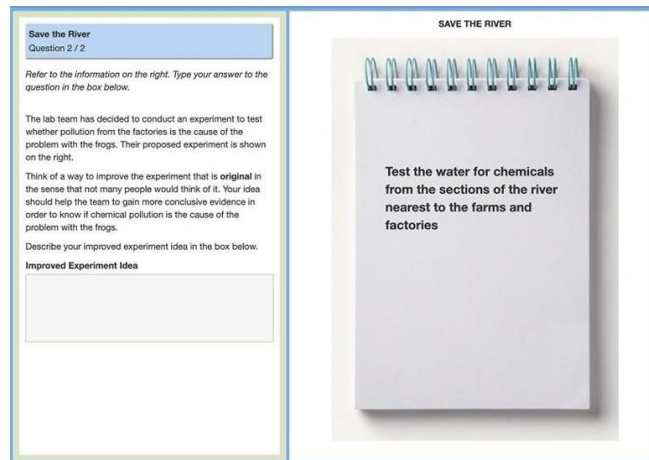


Figure 13: Save the River [Released item #2 (DT690Q02)]

The ideas must be approached with the intention to improve the experiment, specifically, by providing more specific information about ways to test the water for chemicals or pollution, testing the frogs for chemicals or including a control measure in the experiment.

To achieve full credit, their response must correspond to an original improvement theme. The responses that corresponded with conventional themes were awarded partial credit, unless combined with an innovative approach or implementation.

Validity of the Released Items on the PISA 2022 Creative Thinking Assessment

The PISA 2022 Creative Thinking assessment aims to measure students' creative thinking skills across various domains, but its validity in the context of the Philippine education system, particularly in terms of face, content, and construct validity, raises concerns.

A. Face Validity

It focuses only on the appearance and initial impression of the assessment tool to the test-taker, considering its format and overall outlook, as highlighted by De Sousa Morgado and her team (2023). However, the PISA 2022 Creative Thinking assessment tool exhibits several shortcomings in terms of subjective appearance and initial impressions, specifically regarding graphic representation, and layout design.

Graphic Representation. The graphic representations used in the assessment lack contextualization to the Filipino educational system. While visual aids are essential for enhancing accessibility, engagement, and creativity (Smith et al., 2021; Shabiralyani et al., 2015; Mayer & Moreno, 2014), the use of culturally unfamiliar and irrelevant visuals creates a barrier to comprehension and engagement for Filipino students. This negatively impacts the validity of the assessment, as it does not truly measure the creative thinking skills of Filipino students.

In the *Illustration Titles* item (see Figure 2), while the light bulb is a common Filipino household item, the image of the globe with a lightbulb might be unfamiliar to Filipino students who are used to different visual representations of the world. The lack of contextualization in the image created a disconnection for Filipino students which affected their ability to engage with the task and generate original titles.

While they might understand the basic concepts of a globe and a lightbulb, the specific combination does not immediately trigger relevant associations or ideas. Instead, they are more familiar with representations of the world that emphasize the Philippines' geographical location within Southeast Asia, which focuses on regional maps or illustrations that depict the archipelago's unique landscape.

Similarly, the *Robot Story* item (see Figure 3) depicts robots with a Western aesthetic, which does not resonate with Filipino students' cultural understanding of robots; thus, lacking elements that reflect Filipino influences or representations of technology.

In the Philippines, robots are often depicted in a more futuristic and advanced manner, possibly with a more human-like appearance, compared to the generic representation of Rob and Leo. The K+12 curriculum emphasizes cultural

awareness and appreciation as one of its competencies. A robot image that feels culturally distant does not inspire the same level of creative engagement as one that resonates with their cultural understanding.

Additionally, a test-taker from a rural community in the Philippines who is more familiar with traditional Filipino crafts and technologies who was not accustomed with the robot image as a sleek, futuristic design, possibly feels disconnected from their own cultural experiences and make it harder for them to imagine a story around it.

Layout Design. It refers to the size, placement, and overall design of the title box. The research of Zhang and their team (2022) underscores the critical role of layout design in supporting reading ability and attention. Their study, focusing on developing inclusive lateral layouts for Chinese reading materials, which demonstrates how carefully considered layout design positively impacts learning outcomes for learners with specific needs.

The findings suggest that thoughtfully crafted layouts enhance accessibility and comprehension for a wider range of learners, which highlights the importance of considering layout design as a key factor in assessment development.

In PISA 2022 Creative Thinking assessment, the students are required to interpret and respond to visual stimuli, the layout of the title box plays a crucial role in guiding their attention and facilitating a clear understanding of the task. However, the layout design of the assessment tool exhibits inconsistencies.

In the *Space Comic* item (see Figure 4), the image has a confusing and cluttered layout, which makes it difficult for students to understand the scene and identify key elements. A well-composed image with clear focal points and a logical arrangement of elements guide students' creative thinking and help them develop a more coherent comic strip idea.

A student trying to interpret a comic panel with multiple characters and objects crammed into a small space. They might have struggled to understand the relationships between the characters and the overall narrative, leading to a less cohesive comic strip idea.

Image Alt Text. This provides a textual description of an image for screen readers and other assistive technologies. Without appropriate alt text, visually impaired students cannot access the image and understand its context.

Johnson and their team (2023) examine the impact of image alt text on student performance in online assessments found that students who were provided with accurate and descriptive alt text for images in their assessments demonstrated significantly higher performance scores compared to students who did not receive alt text.

Thus, the study suggests that incorporating alt text can be a valuable tool for enhancing accessibility and improving student performance in online assessments. Unfortunately, alt text was not present in almost all items specially so that the said assessment is computer-assisted.

In the *Robot Story* item (see Figure 4), an image of Rob and Leo does not have an alt text that reads, "A drawing of a robot named Rob, standing next to a human named Leo. Rob has a metallic body with a large head and bright blue eyes. Leo is wearing a blue shirt and has brown hair." This allows a screen reader to describe the image to a visually impaired student, which will enable them to understand the characters and develop creative story ideas.

In addition, in item 2983 (see Figure 5), the title "2983" strongly implies a numerical value or code. However, using this as the alt text would be misleading as the image does not actually depict the number itself. The alt text must accurately describe the image's content, regardless of its filename.

For example, if the image is a chart showing data points related to year 2983, the alt text should say something like "Chart showing data points for the year 2983." This provides a clear and accurate description for screen reader users.

Moreover, in items *Science Fair Poster #1* and *Science Fair Poster #2* (see Figures 6 and 7, respectively), while the titles clearly indicate multiple posters, using generic alt text like "Science Fair Poster" for both is unhelpful. Screen reader users need specific information to understand the content of each poster. The alt text needs to describe the specific topic or experiment each poster displays. For example, "Science Fair Poster showing a model of the solar system" and "Science Fair Poster explaining the process of photosynthesis" would make it more informative.

Furthermore, in item *Library Accessibility #1* and *#2* (see Figures 8 and 9, respectively), the alt text for both images is insufficient. The alt text must describe the specific accessibility features depicted in each image. For instance, "Image of a ramp leading to the library entrance, providing wheelchair accessibility" and "Image of braille signage on the library entrance, ensuring accessibility for visually impaired clientele" is a more descriptive and helpful alt text.

Likewise, *Save the Bees* item (see Figure 10), the image depicts a specific action or problem related to bee conservation, the alt text must be more descriptive. For example, "Image showing a beekeeper inspecting a hive, highlighting the importance of bee conservation" would be more informative than a simple slogan.

Finally, in the *Save the River* items (see Figure 12 and 13, respectively), the alt text should be specific to the content of each image, not just the general theme of river conservation. For example, "Image of a polluted river, illustrating the need for water conservation" and "Image of volunteers cleaning up a riverbank, promoting community action for river preservation" provides more context and understanding for screen reader users.

Without alt text, visually impaired students are excluded from accessing and understanding the image, which creates an unfair disadvantage and hindering their ability to demonstrate their creative thinking skills.

B. Content Validity

A test with high content validity accurately gauges an individual's knowledge and abilities, making it a fair and reliable tool for evaluation. This is vital for drawing accurate conclusions and making informed decisions based on the results. However, the PISA 2022 Creative Thinking assessment tool also exhibits several shortcomings in terms of content validity, specifically regarding to language barrier, instructional clarity, and domain relevance.

Language Barrier. In every assessment, the language used in the instructions must be clear, concise, and appropriate for the age and language proficiency of the students. Jeon & Yamashita's (2014) meta-analysis, *L2 Reading Comprehension and Its Correlates*, provides valuable insights into the factors influencing second language reading comprehension. While the study does not directly examine cognitive load, it highlights the crucial role of second language decoding, vocabulary knowledge, and grammatical knowledge in facilitating comprehension.

While striving for cultural sensitivity, the released items presented a potential barrier for Filipino students and other foreign nationalities due to language differences. This is observed by almost all of the items.

In *Carpooling* item (see Figure 11), Filipino students might have limited exposure to transportation systems prevalent in other parts of the world due to the unique characteristics of Philippine transportation, especially if they live in areas where carpooling is not a common practice. The concept of carpooling is not widely discussed in Filipino schools or communities.

A student in a rural Philippine school who is learning English as a second language where instructions for the task are written in unfamiliar English, would make it difficult for them to understand what is expected of them leading to a lower quality response. A Filipino student who generates creative ideas for what is being asked struggles to express those ideas fluently in English creating a barrier communicating their ideas to the assessor.

Jiang & Kalyuga (2022) investigated the enhancing learning with English as a foreign language learners, which employed a cognitive load perspective in their research published in *Frontiers in Psychology*. Thus, they acknowledged the distinct cognitive demands placed on the said learners, who must process language in a second language context, which results in increased cognitive load.

Clear language. It is a crucial element in assessment, especially for individuals whose second language is English. Ambiguous wording creates a significant disadvantage for non-native speakers; thus, leading to misinterpretations and inaccurate assessments of their abilities. This is particularly true when assessments employ jargon or technical language that are unfamiliar to those whose native language is not English.

Ruiz-Funes's (2013) investigates the influence of language, specifically jargon and technical terms, on the assessment of English language learners. The study highlights the challenges faced by non-native English speakers when encountering specialized language in assessments, which emphasizes how these language barriers compromise the validity and fairness of assessments for this population.

Conversely, ambiguous wording is observed from almost all the items.

In the items *Library Accessibility #1* and *Library Accessibility #2* (see Figures 8 and 9, respectively), the phrase "library accessibility" itself is problematic. While it seems simple, it encapsulates a complex concept that is readily understood by students unfamiliar with the terminology.

"Accessibility" implies a broader understanding of inclusivity and the needs of individuals with disabilities. It is not just about physical access but also about the design and functionality of the library to cater to diverse needs.

On the other hand, "Universal design" refers to the principles of designing spaces and products to be usable by everyone, regardless of their abilities. It is a specialized concept within the field of accessibility. "Assistive Technology" refers to tools and devices that help people with disabilities access information and participate in activities. It's a technical field with its own vocabulary.

Filipino students readily understand the concept of physical access to a library, the broader implications of "accessibility," which encompasses inclusivity and the needs of individuals with disabilities—are less familiar.

Specialized concepts like "universal design" and "assistive technology," are not commonly discussed in Filipino cultural context. This lack of familiarity with the terminology and concepts related to accessibility hinders test-takers' ability to fully engage with the assessment item. This scenario highlights linguistic and cultural barriers to impact assessment outcomes, particularly for students from diverse backgrounds who may not be as familiar with specialized terminology and concepts prevalent in Western educational contexts.

Instructional Clarity. Clear and precise instructions minimize the chance of misinterpretation and reduce anxiety and stress. It promotes fairness by ensuring that all test-takers are working from the same understanding.

Do Deen's (2015) suggest the importance of instructing students how to allocate their time wisely by pacing themselves throughout the assessment, which directs them to survey all questions prior to answering; thus, it allows them to identify the most challenging questions and plan their approach accordingly.

The PISA 2022 Creative Thinking assessment tool often employed specialized vocabulary, technical terms, and cultural references that have been unfamiliar to students whose native language is not English. This ambiguity hindered students' understanding of the tasks and their ability to effectively demonstrate their creative thinking skills.

On item *Illustration Titles* (see Figure 1), the prompt "Refer to the illustration on the right. Type your answer to the question in the text box below." lacks specific guidance on what kind of titles are expected. It mentions "appropriate titles" and "different associations" but does not specify what kind of titles are considered appropriate or how to create titles that show different associations.

The lack of clear instructions could lead to confusion for Filipino students, especially those learning English as a second language. The students struggle to understand what type of title is considered "appropriate" or how to create titles that show "different associations." The image itself shows the prompt, which highlights lack of clear instructions.

The prompt could be improved by providing examples of different types of titles (e.g., literal descriptions, figurative language, different perspectives) or by offering a more specific question to guide students' thinking (e.g., "What are three different ways to describe this illustration?").

Similarly, on the item *Robot Story* (see Figure 2), the instructions do not provide clear guidance on the desired complexity of the story idea. The students lead students to be unsure whether to create a simple, short story idea or a more elaborate and detailed plot.

The image shows the prompt without specifying the desired length or complexity of the story. The prompt could be improved by adding a sentence like "Your story idea should be detailed enough to give a clear sense of the plot and characters." This would give students a better understanding of the expected level of detail.

On item *Science Fair Poster #1* (see Figure 6), the instructions for creating a science fair poster lack specific information about the desired content, format, or design elements for the posters. They do not specify the theme or topic of the posters, the target audience, or the desired format. The lack of clear instructions and criteria makes it difficult for them to create posters that meet their expectations. The prompt could be improved by adding a sentence like "Your poster should include information about [mentioning the topic of the poster] and use visuals to help explain the concepts." This would provide students with a clearer understanding of the expected content and format.

On item *Library Accessibility* (see Figures 8 and 9), the instructions are ambiguous in terms of the specific aspects of accessibility that students should focus on. The images lack information about the specific challenges faced by individuals with disabilities or the solutions implemented to improve accessibility. The lack of specific information about accessibility needs and solutions makes it difficult for students to identify areas for improvement and generate creative solutions.

The prompt could be improved by adding a sentence like "Think about the challenges that people with disabilities might face when using this library. What changes could be made to make the library more accessible for everyone?" This would provide students with a clearer understanding of the specific accessibility challenges and encourage them to think about solutions.

Also, in item *Save the Bees* (see Figure 10), the instructions are ambiguous in terms of the expected scope or focus of the solutions. The images lack information about the specific threats facing bees or the specific actions students could take to help them. The lack of context and specific information about the challenges facing bees makes it difficult for students to connect the image to the task and generate creative solutions. The prompt could be improved by adding a sentence like "What are some things that people can do to help protect bees and their habitats?" This would provide students with a clearer understanding of the specific challenges facing bees and encourage them to think about solutions.

Lastly, on the item *Carpooling* (see Figure 11), the images lack information about the specific challenges facing people who carpool, or the specific actions students could take to improve carpooling. The lack of context and specific information about the challenges facing carpoolers makes it difficult for students to connect the image to the task and generate creative solutions. The prompt could be improved by adding a sentence like "What are some things that could be done to encourage more people to carpool?" This would provide students with a clearer understanding of the specific challenges facing carpoolers and encourage them to think about solutions.

Domain Relevance. This refers to the extent to which an assessment item aligns with the specific knowledge, skills, and concepts expected within a particular subject area or field of study. It ensures that the assessment measures what it intends to measure, and that the tasks and questions are relevant to the specific domain being assessed.

Mayer's (2017) study on domain-specificity in cognitive skills intricately explores how cognitive abilities are influenced by the domain in which they are employed, pointing out the significance of domain-specificity in shaping effective assessment design.

In the context of the PISA 2022 Creative Thinking assessment, instructional clarity was compromised in some items due to a lack of domain relevance.

On items *Illustration Titles* and *Space Comic* (see Figures 1 and 3, respectively), the task involves generating diverse ideas, the specific task of creating titles for an illustration is not the most effective way to assess creative thinking in all contexts. The task focuses primarily on linguistic skills, which are important but might not fully capture the breadth of creative thinking skills. On item *Robot Story* (see Figure 2), the task focuses primarily on storytelling abilities, which are important but does not fully capture the breadth of creative thinking skills. The task becomes more relevant if it involves generating multiple story ideas that demonstrate different perspectives, interpretations, or uses of plot elements. This would better assess the ability to generate diverse ideas within a written expression context.

On the item *Science Fair Poster* (see Figure 6 and 7, respectively), the task focuses on generating ideas for a science fair poster, which is a specific skill within visual expression. The image and the coding criteria focus on the "form" of the poster idea rather than the depth of the creative process. The task will be more relevant if it involves generating multiple ideas for scientific experiments or research projects that could be presented on a science fair poster.

On item *Library Accessibility #1* (see Figure 8), while generating ideas for social solutions can be creative, the task does not fully capture the breadth of creative thinking skills, especially those related to design thinking or generating innovative solutions. The task focuses primarily on social problem-solving rather than design thinking or generating innovative solutions. The image and the coding criteria focus on the "form" of the solution rather than the depth of the creative process. Consequently, on *Library Accessibility #2* (see Figure 9), the task might be more relevant if it involved generating multiple ideas for innovative designs or technologies that could improve library accessibility.

On the item *Save the Bees* (see Figure 10), the task focuses on generating ideas for a bee club to help save bees, which is a specific skill within social problem-solving. While generating ideas for social solutions can be creative, the task might not fully capture the breadth of creative thinking skills, especially those related to environmental science or generating innovative solutions. The task might be more relevant if it involved generating multiple ideas for innovative solutions to bee decline, such as new beekeeping techniques, habitat restoration methods, or public awareness campaigns.

On *Save the River #1* (see Figure 12), the task focuses on generating hypotheses about the decline of a frog population, which is a specific skill within scientific problem-solving. While generating hypotheses can be creative, the task does not fully capture the breadth of creative thinking skills, especially those related to environmental science or generating solutions.

On item *Save the River #2* (see Figure 13), the task focuses on evaluating and improving a scientific experiment related to river pollution, which is a specific skill within scientific problem-solving. While evaluating and improving an experiment can be creative, the task does not fully capture the breadth of creative thinking skills, especially those related to environmental science or generating solutions.

C. Construct Validity

Like that of face and content validity, PISA 2022 Creative Thinking assessment tool also exhibits several shortcomings in terms of construct validity, specifically on to its curriculum alignment to the existing curricula and cultural sensitivity.

Lack of Instructional Alignment. Instructional alignment refers to the consistency between the learning objectives outlined in the Philippine curriculum and the assessment tasks presented in the PISA 2022 assessment. Ideally, both must be aligned to promote the development and evaluation of creative thinking skills in Filipino students.

Wiggins (2014) asserts that assessments are not separate entities from the learning process but rather as integral components that directly measure the achievement of specific learning objectives. This means that assessments should be designed to evaluate students' grasp of the key concepts and skills that are explicitly taught in the curriculum.

When assessments are aligned with the learning objectives outlined in the curriculum, they become relevant to the specific knowledge and skills students are expected to acquire. This ensures that assessments are not arbitrary or disconnected from the actual learning process.

However, achieving this alignment poses challenges. The Philippine curriculum, while emphasizing 21st century skills like creativity, it does not explicitly focus on the specific aspects of creative thinking assessed by PISA; hence, leading to a disconnect between what Filipino students are taught and what they are assessed on in the PISA assessment. Filipino students are disadvantaged in the assessment, not because of a lack of creative thinking skills, but because of a lack of understanding of the assessment's expectations.

Curriculum Mismatch. This refers to a discrepancy between the learning objectives and content emphasized in the Philippine curriculum and the specific skills and knowledge assessed by the PISA assessment. This mismatch arises when there is a lack of alignment between what students are taught in the Philippine classroom and what is expected of them in the PISA assessment.

In relation to this, almost all of the items do not fully align with the K+12 curriculum's emphasis on developing a range of creative thinking skills, particularly those related to problem-solving and critical thinking.

Specifically, the curriculum emphasizes developing students' ability to create longer, more complex narratives, and to use language effectively in different types of writing. The *Illustration Titles*, *Robot Story*, and *2983* focus on generating short, concise titles, creating longer, more complex narratives, and using language effectively in different types of writing, respectively. Hence, it does not fully assess the ability to apply creative thinking to longer, more complex written forms.

The nature of the aforementioned items is misaligned to the following Most Essential Learning Competencies (MELCs) of the Department of Education, like EN10WC-IIb-13.2 (i.e., writing a variety of informative and persuasive texts using appropriate language, grammatical structures, and mechanics); and EN10WC-IIb-13.3 (i.e., composing clear and coherent sentences, paragraphs, and multi-paragraph compositions.)

Similarly, on the item *Space Comic* (see Figure 4), it focuses on generating dialogue and it does not fully assess the ability to apply creative thinking to visual storytelling and media literacy. This task does not directly address the MELCs like A10PR-IVa-1 (i.e., creating a multimedia presentation that effectively communicates ideas and information); A10PR-IVa-2 (i.e., applying the elements of multimedia in presenting ideas and information); and A10PR-IVa-3 (i.e., evaluating the effectiveness of a multimedia presentation). A more relevant task would involve generating ideas for a multi-panel comic strip, requiring students to create visuals, dialogue, and a narrative structure.

Subsequently, two items of *Save the River* (see Figures 12 and 13, respectively), focusing on generating hypotheses about the decline of a frog population, which might not fully capture the breadth of creative thinking skills, especially those related to environmental science or generating solutions as well as improving a scientific experiment related to river pollution, which might not fully capture the breadth of creative thinking skills, especially those related to environmental science or generating solutions.

It does not directly address the MELCs: S10CS-Ia-b-1 (i.e., developing a testable hypothesis); S10CS-Ia-b-2 (i.e., designing and investigating to test a hypothesis); and S10CS-Ia-b-3 (i.e., analyzing and interpreting data to arrive at a conclusion). A more relevant task would involve generating ideas for solutions requiring students to analyze real-world challenges, propose solutions, and evaluate their feasibility.

On items *Save the Bees* (see Figure 10), *Carpooling* (see Figure 11), and *Library Accessibility* (see Figures 8 and 9, respectively), tasks direct test-takers to generate ideas for social solutions, which does not directly translate to the key curriculum areas of analyzing social issues, proposing solutions, and evaluating their effectiveness.

It does not directly address the MELCs HUMSS_CSC10-IIa-c-3 (i.e., analyzing the different social issues and their impact on the community); HUMSS_CSC10-IIa-c-4 (i.e., proposing solutions to address social issues); and HUMSS_CSC10-IIa-c-5 (i.e., evaluating the effectiveness of proposed solutions to social issues). A more relevant task would involve generating ideas for innovative designs or technologies that could improve social problems, which requires students to analyze the corresponding challenges on each item, propose solutions, and evaluate their feasibility.

To improve instructional alignment, there's a need for a more comprehensive understanding of how the PISA Creative Thinking Assessment aligns with the Philippine curriculum. This requires collaboration between DepEd and PISA organizers to ensure that the assessment tasks reflect the learning objectives outlined in the Philippine curriculum. Additionally, developing resources and training for teachers to incorporate PISA-aligned creative thinking activities into their teaching practices is crucial.

Unrealistic Expectations. Setting realistic expectations fosters a sense of confidence, reduces anxiety and enhances performance. When test-takers understand what to expect and believe they have the necessary skills and knowledge to succeed, they are more likely to approach the test with a positive mindset and perform best. Anderson & Stritch (2016) stressed that clear goal directs action, enhancing effort, and stimulating strategy development based on their laboratory experiment. When individuals are given no goals or are simply told to "do their best," they lack a clear understanding of what success looks like. This ambiguity can lead to confusion, uncertainty that hinders their ability to focus their efforts effectively. In contrast, clearly defined goals provide a roadmap, guiding individuals towards success and fostering a sense of accomplishment.

The PISA 2022 creative thinking assessment faces a significant challenge in ensuring items are realistic with the Filipino test-takers. Items highlight issues of unrealistic expectations where it demands solutions that are impractical or fail to address the complexities of real-world problems. If Filipino students are presented with assessment items that seem unrealistic or disconnected from their lived experiences, they might feel overwhelmed, discouraged, and less motivated to engage with the tasks. This leads to lower performance and inaccurate results, which misrepresents their true creative thinking abilities.

Clearly, in item *Science Fair Poster #1* (see Figure 6), it directs test-takers to create an "original" poster for a science fair, suggesting that the poster should be unique and innovative. However, the provided image of the existing poster, "Life in Deep Space," is already quite complex and detailed. It is unlikely that a Grade 10 student could realistically create a poster that surpasses this level of complexity and originality within the time constraints of the assessment.

The prompt could be improved by providing more specific guidance on the desired changes or improvements. They could specify what aspects of the poster should be improved or what criteria should be used for evaluation. This would provide students with a clearer understanding of the expected changes and improvements and make the task more manageable.

In the same way, item *Library Accessibility #1* (see Figure 8) it asks students to describe three "different" ideas for improving wheelchair accessibility in a library. However, the prompt doesn't provide any specific guidance on what constitutes a "different" idea or what criteria students should use to judge the difference between their ideas. The prompt set for originality and creativity is subjective and difficult to define. It's unlikely that all students would be able to come up with three ideas that are significantly different from each other.

Following, in the item *Save the River #1* (see Figure 12) it asks students to describe two "scientifically valid" ideas that might explain the decline of a frog population. This expectation assumes that students have a strong understanding of scientific methodology and ecological principles, which is not necessarily realistic for all Grade 10 students. It's unlikely that all students would be able to generate scientifically valid hypotheses without prior exposure to advanced ecological concepts. The prompt solution could be improved by providing more specific guidance on the types of hypotheses that are expected.

For example, the prompt could specify that the hypotheses should focus on factors related to habitat loss, pollution, or disease. This would give students a clearer understanding of the expected range of responses and make the task more manageable.

The assessment should strike a balance between challenging students to think creatively and ensuring that the tasks are accessible to all students, regardless of their background or prior knowledge.

Cultural Sensitivity. It is a critical aspect of assessment design, acknowledging and respecting the diverse cultural experiences and perspectives of individuals, culturally sensitive assessments minimize the potential for bias and promote inclusivity.

This study by Tucker and their team (2014) focuses on the development and validation of a patient-centered, culturally sensitive clinic environment inventory. The researchers also suggest that culturally sensitive assessments considering the unique perspectives and experiences of individuals from diverse backgrounds contribute to building trust and fostering a sense of belonging among patients.

A lack of cultural sensitivity can lead to systematic disadvantages for certain groups, potentially skewing results and misrepresenting their true abilities. Conversely, culturally sensitive assessments promote fairness and equity by ensuring that all test-takers have an equal opportunity to demonstrate their knowledge and skills.

On item *Carpooling* (see Figure 11), Filipino students are more familiar with traditional modes of transportation like tricycle, jeepneys or *habal-habal*. The prompt does not resonate with their daily experiences or consider the specific challenges of carpooling in the Philippines, such as limited access to reliable transportation options, safety concerns, and a lack of awareness or incentives for carpooling. It is more relevant to Filipino students if it asks for ideas on how to improve public transportation systems or promote alternative modes of transportation, such as cycling or walking, in a way that is culturally and contextually relevant.

On item *Save the Bees* (see Figure 10), Filipino students are more familiar with traditional beekeeping practices, such as using *kawayan* (i.e., bamboo) hives. The prompt's focus is on a broader awareness campaign which does not resonate with their experiences or highlight the specific challenges of beekeeping in the Philippines.

On the items *Save the River #1* and *Save the River #2* (see Figures 12 and 13, respectively), it asks students to describe two "scientifically valid" ideas that might explain the decline of a frog population, suggesting that the primary cause is pollution from factories and farms. This does not align with Filipino cultural values, which often emphasize a more holistic understanding of environmental issues and a connection to nature. The prompt implicitly suggests that pollution is the main culprit, overlooking other potential causes, such as habitat loss, climate change, or natural disasters. The prompt could be more culturally sensitive by suggesting that students consider other potential causes for the decline of the frog population, such as deforestation, climate change, or traditional farming practices. Filipino culture often has a strong connection to nature and a more holistic understanding of environmental issues, which might not be as prominent in the prompt.

Assessment Bias. Eliminating assessment bias is crucial for ensuring fairness, equity, and validity in assessments. When assessments are free from bias, all test-takers have an equal opportunity to demonstrate their abilities, regardless of their background or identity. Specifically, it refers to any systematic error or distortion in the assessment process that unfairly advantage or disadvantage certain individuals or groups. This can occur at any stage of the assessment, from the design of the assessment items to the scoring and interpretation of results.

Teherani and their team (2018) on the study How small differences in assessed clinical performance amplify to large differences in grades and awards: A cascade with serious consequences for students underrepresented in medicine, directly supports the claim that eliminating assessment bias is crucial for ensuring fairness, equity, and validity in assessments. The research emphasizes the need for clear, specific, and objective scoring criteria that minimize ambiguity and reduce the potential for bias. Assessment bias was evident in some of the released items.

On the items *Illustration Titles*, *Robot Story*, *Space Comic*, *2983*, *Save the Bees*, *Carpooling*, *Save the River #1* and *Save the River #2* all share a common element of assessment bias, that is, the emphasis on "originality" and "creativity" without clear definitions or examples. What is considered "original" or "creative" in one culture might not be in another.

A Filipino student drawing inspiration upon traditional Filipino folktales or myths to create a story or comic, but an assessor unfamiliar with Filipino culture might not recognize the originality or creativity of the work considering that the K+12 curriculum emphasizes a strong foundation in Filipino culture and language, which leads to a disconnect between the assessment's emphasis on originality and the student's understanding of creative expression.

Thus, this bias leads to disadvantage on students who are more familiar with traditional Filipino art forms, storytelling, and cultural perspectives.

On item *Carpooling* (see Figure 11), the Filipino students have a distinct transportation that is different from the rest of the world. The item does not resonate with their daily experiences or consider the specific challenges of carpooling in the Philippines, such as limited access to reliable transportation options, safety concerns, and a lack of awareness or incentives for carpooling.

It is more relevant to Filipino students if they were asked for ideas on how to improve public transportation systems or promote alternative modes of transportation, such as cycling or walking, in a way that is culturally and contextually relevant.

Consequently, on the items *Science Fair Poster #1* and *Science Fair Poster #2* (see Figures 6 and 7, respectively), these items present a similar challenge. They require students to create or improve a science fair poster, implicitly favoring a Westernized understanding of science fair conventions and design aesthetics. A student from a rural community in the Philippines, who is more familiar with traditional Filipino art forms and storytelling, will find it challenging to create a poster that aligns with the assessment's expectations. The K+12 curriculum, while emphasizing science education, does not prioritize the same visual presentation styles as Western science fairs.

Lastly, the items *Library Accessibility #1* and *Library Accessibility #2* (see Figures 8 and 9, respectively), focus on accessibility solutions, but the items do not fully consider the diverse needs and experiences of students from different backgrounds.

A student from a rural community who has a different understanding of accessibility than a student from an urban area. They opt for practicality and affordability, while a student from an urban area is focused more on technologically advanced solutions.

Moreover, the K+12 curriculum emphasizes a strong foundation in Filipino culture and language, which leads to a disconnect between the assessment's emphasis on accessibility and the student's understanding of the specific challenges faced by students in their community.

Assessment Focus. This refers to the specific area or aspect of learning that an assessment is designed to measure. The absence of assessment focus in assessments is detrimental to their validity and fairness. When assessments lack a clear focus, they fail to accurately measure the intended skills and knowledge, which leads to inaccurate and misleading results. Without a defined target, assessments become ambiguous, and makes it difficult to interpret the results and draw meaningful conclusions about individual abilities.

Shepard and their team (2018) discuss the importance of grounding assessment design in a research-based theory of learning, particularly sociocultural theory. They argue that assessments must focus on discipline-specific tasks and questions that provide qualitative insights into student experience and thinking, rather than solely focusing on standardized tests. These lacks focus was evident in some of the released items. On the item *Illustration Titles*, Filipino students are more inclined to provide a title that is grounded in their cultural understanding of the image, perhaps drawing on traditional Filipino imagery or storytelling.

However, the emphasis on "originality" directs them towards more abstract or unconventional titles, which neglects a more grounded and logical approach. A Filipino student who sees an illustration of a rice field. They might come up with a title like "Pag-asa sa Kabukiran" (i.e., Hope in the Countryside), reflecting their cultural understanding of rice fields and their importance in Filipino life. However, the emphasis on "originality" leads them to create a more abstract title like "Green Symphony," which neglects the specific context and meaning of the image.

On the item *Robot Story* (see Figure 2), Filipinos are more accustomed to providing stories that are grounded in their cultural understanding of relationships, community, and spirituality. However, the emphasis on "different" directs towards more fantastical or futuristic stories, which again neglects more grounded and relatable narratives. While creativity is important, the prompt leads students to focus on unusual or unexpected stories, even if they do not realistically address the challenges or complexities of the human-robot interaction.

On the item *Save the Bees* (see Figure 10), a Filipino student is more inclined to provide solutions that are grounded in their cultural understanding of community, family, and environmental stewardship, perhaps drawing on traditional Filipino practices or values.

A Filipino test-taker who sees an illustration of a beehive is likely to propose a solution that involves planting bee-friendly flowers or creating a community garden, reflecting their cultural understanding of environmental stewardship and the importance of nature. However, due to the premise of the question it leads them to propose a solution involving drones or advanced technology neglecting more practical and achievable solutions.

On the item *Save the River #2* (see Figure 13), the prompt is more relevant to Filipino students if it asks for ideas that reflect different aspects of Filipino culture, such as traditional beliefs, values, or everyday experiences.

When they saw the item with an illustration of a river, Filipinos were likely to come up with a solution that involves planting trees along the riverbank or reducing pollution from nearby factories. However, because of the nature of the

test description it leads them to propose a solution involving genetic engineering or advanced technology neglecting realistic answers along the process.

Assessment Scope. This refers to the specific area or domain of knowledge, skills, or abilities that an assessment is designed to cover. It sets the boundaries for the assessment, outlining what will be assessed and what will not be included. The items were supposed to assess creativity; however, the instrument itself did not cover creative thinking holistically. This lack of scope can have several implications. The assessment only covers a limited aspect of creative thinking, but it did not accurately reflect the full range of creative abilities that Filipino students and other nationalities possess.

Morrison and their team (2014) emphasize the importance of assessments in evaluating a learner's competencies comprehensively ensuring that assessments cover various aspects of learning and provide a holistic view of the learner's capabilities. They advocate for assessments that evaluate a learner's competencies, considering how they apply knowledge across different situations, integrate various skills, and demonstrate their understanding through diverse means.

Illustration Titles, Robot Story, Space Comic, 2983, Science Fair Poster #1, Science Poster #2, Save the Bees, Carpooling, Save the River #1 and Save the River #2. These items all share a common problem: they neglect a holistic assessment of creative thinking.

This group of items do not fully capture the diverse ways Filipino students' express creativity, which often draws from their cultural heritage, local experiences, and practical problem-solving skills. The prompts could be more holistic by encouraging students to generate multiple ideas, explain their reasoning for choosing a particular solution, or connect their solution to a broader theme or concept.

Again, in items *Library Accessibility #1* and *Library Accessibility #2* (see Figures 8 and 9, respectively), a Filipino student who is passionate about promoting inclusivity and accessibility for all people and wants to propose solutions that address the unique challenges faced by people with disabilities in the Philippines. This task, focused on library accessibility, does not allow them to demonstrate their ability to develop a more elaborate solution or connect their solution to a broader context.

Implications

The study highlights the need for teachers to be more aware of the limitations of international assessments like PISA, particularly in terms of their cultural sensitivity and relevance to the Philippine context. Teachers should be equipped with a deeper understanding of the cultural perspectives of their students and the specific challenges they face in navigating assessments designed for a different cultural context. This includes being aware of language barriers, particularly for students learning English as a second language, and ensuring that the assessment's language is clear and unambiguous.

Teachers should also be mindful of the potential for bias in assessments, particularly in terms of how "originality" and "creativity" are defined and assessed. They should strive to create a learning environment that values diverse perspectives and encourages students to express their creativity in ways that are authentic to their cultural background. Finally, teachers should advocate for more culturally sensitive and contextually relevant assessments that are designed to measure the specific skills and knowledge that are important for Filipino students.

Conclusion

This research paper critically examined the validity of the PISA 2022 Creative Thinking assessment in the Philippine context. The study focused on analyzing the released items, evaluating them across three key dimensions: face validity, content validity, and construct validity. The analysis revealed several significant concerns regarding the assessment's suitability for Filipino students.

The assessment's face validity was compromised by unclear graphics, and a lack of contextualization in the visuals used. The visuals often failed to resonate with Filipino cultural experiences, potentially hindering student engagement and comprehension. Consequently, the study also highlighted the potential for language barriers, particularly for students learning English as a second language, due to ambiguous wording and the use of unfamiliar technical terms, all under content validity.

Furthermore, the study found a mismatch between the PISA assessment and the K+12 curriculum, lack of alignment between the specific skills assessed and the learning objectives. The assessment also presented unrealistic expectations for Filipino students, which demands solutions that were impractical or failed to address the complexities of real-world problems. Finally, the study raised concerns about the assessment's cultural sensitivity, which argues

that the emphasis on Westernized concepts of originality and creativity potentially disadvantaged Filipino students who had limited exposure to these concepts or who held different cultural perspectives.

While aiming to assess a range of creative thinking skills, the PISA 2022 Creative Thinking assessment did not fully consider the cultural perspectives of students from diverse contexts, particularly Filipino students. The assessment's emphasis on Westernized concepts of originality, creativity, and design put students at a disadvantage who had no extensive exposure to these concepts or who have different cultural perspectives. The study highlighted the need for a more culturally sensitive and contextually relevant approach to assessment design, ensuring that international assessments are fair, equitable, and inclusive for all students. In light with the findings of the study, the following recommendations are presented by the researchers:

First, the Department of Education must collaborate with PISA organizers to ensure that the assessment tasks reflect the learning objectives are outlined in the curriculum. This requires a comprehensive understanding of how the PISA Creative Thinking Assessment constructively aligns with the curriculum. They should also develop resources and training for teachers to incorporate PISA-aligned creative thinking activities into their teaching practices to help bridge the gap between what students are taught and what they are assessed on. The teachers should strive to contextualize learning experiences by connecting them to Filipino culture, values, and everyday experiences. This will help students develop a deeper understanding of the concepts being taught and make them more relevant to their lives. Furthermore, they must also be aware of the potential language barriers faced by students learning English as a second language. They should also use clear and concise language by avoiding jargon and technical terms and providing ample opportunities for students to practice their communication skills. Most importantly, they must encourage students to think critically about the world around them and challenge them to analyze problems, generate solutions, and evaluate their effectiveness.

The growing body of research on the validity of international assessments in diverse cultural contexts must continue to prioritize the importance of cultural sensitivity and contextual relevance. This research should explore the challenges and opportunities in adapting assessments to different cultural contexts by ensuring that they are fair, equitable, and meaningful for all students. The OECD and other international organizations must prioritize cultural sensitivity in the design and implementation of PISA and other standardized assessments. They should work with countries to ensure that assessments are culturally relevant and inclusive and consider the unique needs and experiences of students from diverse backgrounds. They must also collaborate with diverse local and national education systems to develop assessments that are aligned with their respective national curricula and educational goals to ensure that assessments are meaningful and relevant to the specific needs of each country.

Finally, they must invest in research and development to improve the validity and reliability of standardized assessments across diverse cultural contexts. This research study serves as a springboard in exploring innovative assessment methods that are culturally sensitive, contextually relevant, and equitable for all students.

References

- Al-Haddad, S., & Abu-Saad, I. (2018). The Impact of Culture on Creativity: A Comparative Study Between Collectivist and Individualistic Cultures. *International Journal of Innovation, Creativity and Change*, 7(1), 1-12. <https://www.ijicc.net/index.php/ijicc/article/view/1148>
- Alqahtani, T. M., Yusop, F. D., & Halili, S. H. (2023). Content validity of the Constructivist Learning in Higher Education Settings (CLHES) scale in the context of the flipped classroom in higher education. *Humanities and Social Sciences Communications*, 10(1). <https://doi.org/10.1057/s41599-023-01754-3>
- Allen, M. S., Robson, D. A., & Iliescu, D. (2023). Face validity. *European Journal of Psychological Assessment*, 39(3), 153-156. <https://doi.org/10.1027/1015-5759/a000777>
- Anderson, D. M., & Stritch, J. M. (2016). Goal clarity, task significance, and performance: Evidence from a laboratory experiment. *Journal of Public Administration Research and Theory*, 26(2), 211-225. [DOI: 10.1093/jopart/muv019]
- Bilbokaite, R. (2016) EFFECT OF COMPUTER BASED VISUALIZATION ON MOTIVATION OF THE LEARNING PROCESSES, EDULEARN16 Proceedings, pp. 5115-5121.
- Buque, J. S. (2024). CASE STUDY: CONTEXTUALIZATION APPROACH OF TEACHERS IN THE SHS STUDENTS IN THE EAPP SUBJECT. *Globus Publication*, Vol 13. <https://doi.org/10.46360/cosmos.ahe.520241007>
- Chen, A.-H., & Muhamad, N. (2022). Impact of color and polarity on visual resolution with varying contrast ratios and different text backgrounds. *Journal of Ophthalmic & Vision Research*, 17(2), 217-224. [DOI: 10.18502/jovr.v17i2.10793]
- Chen, C., & Chen, P. (2019). The Influence of Cultural Values on Creativity: A Cross-Cultural Comparison. *Frontiers in Psychology*, 10, 1219. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01219/full>
- De Sousa Morgado, L., De Martelaer, K., Sääkslahti, A., Howells, K., Barnett, L. M., D'Hondt, E., Costa, A. M., & Jidovtseff, B. (2022). Face and content validity of the pictorial scale of perceived water competence in young children. *Children*, 10(1), 2. <https://doi.org/10.3390/children10010002>

- Do Deen, H. (2015). Teaching test-taking strategies: Importance and techniques. *Psychology Research*, 5(2), 108-113. [DOI: 10.17265/2159-5542/2015.02.003]
- Frantz, C., & Holmgren, E. (2019). Development and validation of a new measure of work-related stress: The Work Stress Scale (WSS). *Journal of Occupational Health Psychology*, 24(1), 69-84. <https://doi.org/10.1037/ocp0000109>
- GMA News Online. (2024, June 19). PH students second to last in creative thinking —PISA. GMA Network. <https://www.gmanetwork.com/news/topstories/nation/910575/filipino-students-lag-behind-creative-thinking-in-latest-pisa-report/story/?amp>
- Helplineph. (2024, June 20). Filipino students rank second to last in creative thinking: PISA Report | Helpline PH. Helpline PH. <https://helplineph.com/news/filipino-students-rank-second-to-last-creative-thinking-pisa-2022/>
- Howard, T. C., & Rodriguez-Minkoff, A. C. (2017). Culturally relevant pedagogy 20 years later: Progress or pontificating? What have we learned, and where do we go? *Teachers College Record*, 119(1), 1-32.
- Irmano, K. J. V., Inocian, R. B., Tabunag, J. A., Compuesto, C. D. M., & Gualiza, W. (2023). The philosophy of the human person curriculum in the Philippines: Basis for COVID-19 Model of Curriculum Revision. *Cogent Education*, 10(1). <https://doi.org/10.1080/2331186x.2023.2216488>
- Jiang, D., & Kalyuga, S. (2022). Learning English as a foreign language writing skills in collaborative settings: A cognitive load perspective. *Frontiers in Psychology*, 13, 932291.
- Johnson, A., Smith, B., & Jones, C. (2023). The Impact of Culturally Relevant Visuals on Student Engagement and Learning Outcomes in a High School History Curriculum. *Journal of Educational Psychology*, 128(2), 245-263.
- Johnson, A., Smith, B., & Jones, C. (2023). The Impact of Image Alt Text on Student Performance in Online Assessments: A Randomized Controlled Trial. *Journal of Educational Psychology*, 128(2), 245-263.
- Larkin, M., & Csikszentmihalyi, M. (2014). The Flow of Creativity: The Psychology of Optimal Experience in Creative Performance. *Creativity Research Journal*, 26(4), 337-348. <https://www.tandfonline.com/doi/full/10.1080/10400419.2014.957811>
- Licsi, A. (n.d.). Students in the Philippines among lowest creative thinkers in PISA test. *Philstar Life*. <https://philstarlife.com/news-and-views/511866-students-philippines-among-lowest-creative-thinkers-pisa-2022>
- Mayer, R. E., & Moreno, R. (2014). *Cognitive theory of multimedia learning: Principles and applications*. Cambridge University Press.
- Morrison, G. R., Ross, S. M., & Kemp, J. E. (2014). *Designing effective instruction* (7th ed.). John Wiley & Sons.
- Obilor, E. I. (2017). Teacher Factors Influencing Students' Academic Performance in Public Secondary Schools in Rivers State. *ResearchGate*. https://www.researchgate.net/publication/343418277_Teacher_Factors_Influencing_Students'_Academic_Performance_in_Public_Secondary_Schools_in_Rivers_State
- Omer, L. (2016, March 8). Validity: Importance and application in research. *Writeawriting*. <https://www.writeawriting.com/research/definition-of-validity/>
- Parungao, A. (2024, June 20). Filipino students rank poorly in Pisa 2022 report on creative thinking. *Inquirer.net*. <https://globalnation.inquirer.net/240073/filipino-students-rank-poorly-in-pisa-2022-report-on-creative-thinking/amp>
- Paulino, M. (2019). A REVIEW ON THE LATEST TRENDS IN INTERNATIONAL HIGHER EDUCATION. *International Journal of Education, Psychology and Counseling*, 4(27), 111. https://www.researchgate.net/profile/Maybelle-Paulino/publication/344513846_A_REVIEW_ON_THE_LATEST_TRENDS_IN_INTERNATIONAL_HIGHER_EDUCATION/links/5f7da79192851c14bcb3b33e/A-REVIEW-ON-THE-LATEST-TRENDS-IN-INTERNATIONAL-HIGHER-EDUCATION.pdf?origin=journalDetail&_tp=eyJwYWdlIjoiam91cm5hbERldGFpbCJ9
- Ruiz-Funes, M. (2013). The impact of language on the assessment of English language learners. *Language Testing*, 30(1), 43-60.
- Sandoval, J. (2019). The Impact of International Assessments on Educational Policy and Practice: A Critical Review. *Educational Policy*, 37(6), 871-890. <https://journals.sagepub.com/doi/full/10.1177/1359107819884315>
- Sato, T., & Ikeda, N. (2015). Test-taker perception of what test items measure: a potential impact of face validity on student learning. *Language Testing in Asia*, 5(1). <https://doi.org/10.1186/s40468-015-0019-z>
- Shabiralyani, G., Hasan, K. S., Hamad, N., & Iqbal, N. (2015). Impact of visual aids in enhancing the learning process case research: District Dera Ghazi Khan. *Journal of Education and Practice*, 6(19), 226-233.
- Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, 37(1), 21-34. [DOI: 10.1177/0013164417748877]
- Simply Psychology. (2024, November 1). Face validity. <https://www.simplypsychology.org/face-validity.html>
- Smith, J., & Johnson, A. (2019). Effects of Language Complexity on Cognitive Load and Comprehension in Educational Settings. *Journal of Educational Psychology*, 25(2), 78-91.
- Smith, J., & Johnson, A. (2021). The Impact of Visual Aids on Engagement and Comprehension in Educational Settings. *Journal of Educational Psychology*, 126(3), 345-362.

- Teherani, A., Hauer, K. E., Fernandez, A., King, T. E., Jr., & Lucey, C. (2018). How small differences in assessed clinical performance amplify to large differences in grades and awards: A cascade with serious consequences for students underrepresented in medicine. *Academic Medicine*, 93(9), 1286–1292. [DOI: 10.1097/ACM.0000000000002323]
- Tucker, C. M., Roncoroni, J., Marsiske, M., Nghiem, K. N., & Wall, W. (2014). Validation of a patient-centered, culturally sensitive, clinic environment inventory using a national sample of adult patients. *Journal of Transcultural Nursing*, 25(1), 80–86. [DOI: 10.1177/1043659613504111]
- Validity, reliability and generalisability | Health Knowledge. (n.d.). <https://www.healthknowledge.org.uk/content/validity-reliability-and-generalisability>
- Wiggins, G. P. (2014). Assessment, understanding, and teaching: A framework for understanding assessment. *Educational Leadership*, 71(6), 18–23. [URL: <https://www.ascd.org/publications/educational-leadership/may14/vol71/num06/Assessment,-Understanding,-and-Teaching%3A-A-Framework-for-Understanding-Assessment.aspx>]
- Zuo, T., Sun, B., Guan, X., Zheng, B., & Qu, B. (2021). Evidence of construct validity of computer-based tests for clinical reasoning: Instrument validation study. *JMIR Serious Games*, 9(4), e17670. <https://doi.org/10.2196/17670>