STEM Education
STEM Education, 5(2): 187–206
DOI: 10.3934/steme.2025010
Received: 03 October 2024
Revised: 15 January 2025
Accepted: 20 January 2025
Published: 19 February 2025

https://www.aimspress.com/journal/steme

Research article

Creating task-specific creativity assessment tools

Hong Liang Lee and Aik-Ling Tan*

Natural Sciences and Science Education, National Institute of Education, Nanyang Technological University, Singapore, Singapore; hongliang.1999.hl@gmail.com, aikling.tan@nie.edu.sg

* **Correspondence:** Email: aikling.tan@nie.edu.sg; Tel: +65 6790 3828.

Academic Editor: Feng-Kuang Chiang

Abstract: Development of creativity is one of the desired outcomes of integrated STEM learning and can serve as an indicator of high-quality education. Reliable evaluation of creativity can serve as an indicator of the success of integrated STEM learning. Current methods of creativity assessment such as the Torrance Test of Creativity (TTCT) have limitations as they are domain-generic and may not accurately reflect creativity valued by specific domains. Furthermore, creativity has been argued to be a dynamic phenomenon, which tests such as the TTCT fail to measure accurately. To consider disciplinary peculiarities, this study examined the value of considering disciplinary epistemic practices and task requirements in evaluating creativity. We adapted the Scientific Creativity Test (SCT) by incorporating elements that are more domain- and task-specific to vertical farming integrated STEM learning and trialed it with pre-service teachers. To measure changes in creativity over time, behavior of pre-service teachers when they are engaged with the integrated STEM activity was monitored through video recordings. Scoring of the adapted SCT revealed a decrease in creativity, but low sample size and "memory" effects meant that no meaningful conclusions were drawn. Behaviors during the integrated STEM learning associated with higher creativity were identified. Findings suggested that the two creativity assessment tools are not substitutable but are complementary—the SCT is suited as an easy-to-administer assessment for a larger number of students, while monitoring behavior of students during integrated STEM learning is suited for smaller groups and can shed light on more aspects of the student's creativity. Recommendations for future studies on the SCT and behavior-monitoring were made.

Keywords: STEM, education, creativity, assessment, task-specific

1. Introduction

STEM is the acronym referring to Science, Technology, Engineering, and Mathematics. Beyond being a term used to represent the four disciplines, STEM represents the interconnectedness between the four disciplines. Correspondingly, Bryan et al. [1] calls for STEM education to go beyond the teaching of these disciplines as discrete subjects and to anchor STEM education with the practices and knowledge distilled from the four disciplines. As such, the integration of the four disciplines gives rise to integrated STEM learning which has been pushed as an important pedagogical tool for STEM education [1,2].

There are many learning outcomes associated with integrated STEM learning [1,3,4], one of which is creativity. To ascertain the effectiveness of integrated STEM learning in achieving these learning outcomes and to provide feedback for improvement, accurate and reliable assessment tools would need to be developed. For creativity, many tests have been developed such as the Torrance Test of Creative Thinking (TTCT), Consensual Assessment Technique (CAT), self-/peer-/expert-assessments, and more [5,6], of which the most widely used and adapted is the TTCT which forms the basis for many Divergent Thinking (DT) tests.

There are, however, calls to reconsider the use of the TTCT and related DT tests which fail to consider the domain-specific nature of creativity [7]. Kaufman et al. [8] and Ugras [4] argued that there is still value in using DT tests with Kaufman et al. raising the example of Hu and Adey's [9] scientific structure creativity model (SSCM) as a domain-specific DT test. The Scientific Creativity Test (SCT) designed based on the SSCM by Hu and Adey assesses student creativity in the domain of science but can still be seen as generic within the wide domain of science itself. Furthermore, there is value in looking at designing DT tests that are task-specific on top of being domain-specific [10]. The dynamic nature of creativity necessitates the development of a dynamic micro-longitudinal measurement method [11]. This entailed monitoring of creative behaviors exhibited by students during the integrated STEM activity conducted. This study thus trialed a task-specific DT test designed based on the context of a vertical farming integrated STEM learning activity and monitored participants' behavior against a rubric to monitor their creative behavior throughout the activity.

In the sections that follow, we will first present the literature regarding STEM education, creativity, and creativity assessments that framed this study. We will then present the Methods section where we explain the design of the adapted DT test and the categories of creative behaviors monitored during integrated STEM learning. Finally, we present the results of the study and compare the use of the two assessment tools in creativity assessment. We conclude our discussion by proposing future areas of research in scientific creativity assessments.

2. Literature review

This section covers the existing literature in the three key areas: (1) STEM education and its link to creativity, (2) the definition of creativity used in this study, and (3) the assessment of creativity.

2.1. STEM education and creativity

STEM education seeks to establish STEM literacy, which can be divided into four aspects: (1) the understanding and application of STEM knowledge, (2) the recognition of research and problem-solving as integral to STEM, (3) the acknowledgement that humankind has been influenced by STEM, and (4) the capacity of students to be active contributors to the community in the field of STEM [3]. These goals of STEM education are encompassed within the 21st Century Competencies (21CC) for a globalized world that the Singapore Ministry of Education (MOE) seeks to develop in Singapore students [12], as well as goals stated in the World Economic Forum Education 4.0 Framework [13]. Such an overlap of goals underscores the importance of STEM education in Singapore and the need for a more comprehensive understanding of how best to deliver STEM education.

Bryan et al. [1] and English [2] advocated for the integration of the four STEM domains to design integrated STEM activities to enable learners to develop a more holistic understanding of the complexities and interrelatedness of problems in the 21st century. Engaging with integrated STEM learning provides opportunities for students to develop their creativity which is one of the learning goals of the 21CC [1,4]. One of the challenges in delivering STEM education using integrated STEM learning is the assessment of creativity. An accurate and reliable assessment tool to assess creativity is needed for determining the development of students' creativity which is often used as a proxy measure of the effectiveness of STEM programs.

2.2. Defining creativity

Before presenting the tools currently available for assessing creativity, it is necessary to define the construct of creativity used in this study. There are many definitions in literature for creativity and, consequently, many assessment tools were developed by researchers. Kanlı [14] raised the need for creativity to be defined before we can develop an accurate and reliable creativity assessment tool. To comprehensively define creativity, literature on creativity research was reviewed, and the factors that influence creativity and the characteristics of creativity were elucidated (Figure 1).

Many factors can influence one's creativity, which Barbot et al. [1] categorized into Education, Sociocultural, and Personal factors. Education factors include normative practices in schools and high-stakes summative assessments that can impede the nurturing of creativity as well as the use of pedagogies that help nurture creativity in students. Sociocultural factors refer to familial influence and prevailing stereotypes in the society the individual lives in, both of which influence creativity. Personal factors include one's motivation when engaging in the creative task, confidence, and domain-specific knowledge, to name a few.

Research in creativity describes creativity as multifaceted. Sternberg [15] presented creativity as the integrated product of an individual's various cognitive abilities, and this idea was supported by Barbot et al. [10], Kaufman and Beghetto [7], Corazza [16], and Plucker et al. [17] who all, in one way or another, argued that many skill sets come together to give rise to creativity.

Second, creativity is domain-specific and task-specific [1,18]. Barbot et al. [10] argued that an individual might have a higher potential for creativity in certain fields than in others, suggesting that creativity is not a general ability that can be applied across multiple disciplines. This was further supported by the poor correlation of creativity scores across various disciplines [19]. Barbot et al. also presented the case for creative expression being dependent on the context of the situation or task that the creativity is being expressed in. Amabile and Pillermer [18], on the task-specific nature of creativity, discussed how intrinsic task-specific motivation could improve levels of creativity.

Third, creativity is dynamic. Corazza [16] proposed the inclusion of "potential" in creativity definitions to give credit to expressions of creativity that are not captured in static "snapshot" measurements of creativity. Beghetto and Karwowski [11], Sternberg [15], and Mullen [20] were also of the view that creativity fluctuates with time and context. Having established the multifaceted nature of creativity, the dynamic nature of creativity can be attributed to the continual interplay between individual and environmental factors [10], which causes creativity to be in constant flux.

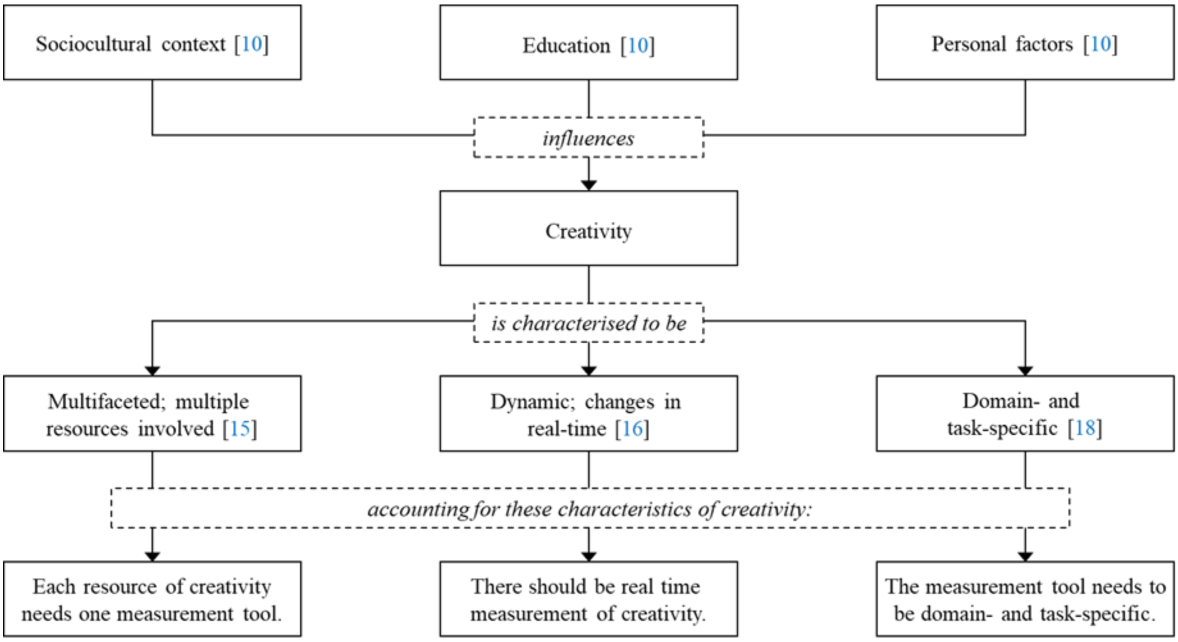

Figure 1. Flow diagram showing the factors that influence creativity, the characteristics of creativity, and the corresponding criteria for the effective assessment of creativity.

These factors that influence creativity and characteristics were aptly captured in the Componential Theory of Creativity [18], which identified the construct of creativity to be made of three intra-individual components: (1) domain-specific skills, (2) creativity-related skills, and (3) task-specific intrinsic motivation, and one external component: the social environment. The social environment influences and mediates the expression of one's creativity, which is based on the interaction of domain-specific skills, creativity-related skills, and task motivation. The Componential Theory of Creativity thus encapsulated the factors influencing creativity and the characteristics of creativity.

Based on these identified characteristics of creativity, we can add to the definition proposed by Plucker et al. [17] and define creativity to be the generation of a product that has potential novelty and usefulness in a specific domain/task resulting from the interplay between personal and environmental factors.

2.3. Assessment of creativity

The abovementioned characteristics of creativity can be observed in the 4C Model of Creativity [6,20] and Csikszentmihalyi's Systems Model of Creativity [20]. In the STEM learning context, we are concerned with the little-c level of creativity in the 4C Model, dealing with the expression of creativity in everyday activities at the interpersonal level [6]. The Systems Model of Creativity comprises three system levels of Person, Field, and Domain in which creative ideas from the Person need to be validated (gatekeepers; Field) before being accepted into the Domain. Mullen associated little-c creativity with the Person level of creativity, requiring less gatekeeping, and proposed the use of the TTCT or other associated DT tests to assess creativity.

However, Baer [7] criticized the present-day TTCT and DT tests as being domain-general. Such criticism is valid, and there is thus the impetus to re-evaluate current assessment tools which are domain-general and even redesign TTCT-derived DT tests to be more domain-specific [7]. Kaufman et al. [8] raised Hu and Adey's [9] Scientific Structure Creativity Model (SSCM) as a domain-specific assessment tool. In the SSCM, domain-specific knowledge is represented by the Product dimension—Technical Product, Science Knowledge, Science Phenomena, and Science Problem—providing four ways to represent scientific knowledge. In the SSCM, creativity is represented by the Trait dimension—Fluency, Flexibility, and Originality (derived from the TTCT as the three scoring criteria for creativity). Creativity is also encapsulated in the Process dimension—Imagination and Thinking—representing two methods students can use to generate creative ideas.

Based on the SSCM, Hu and Adey [9] designed seven questions for their Scientific Creativity Test (SCT). They have ascertained that the SCT is a valid measure of the construct of scientific creativity through factor analysis with principal components. While the SCT is specific to science, it is still considered generic since there are disciplinary differences between biology, chemistry, and physics. As recommended by Barbot et al. [10], there is value in adapting the domain-specific SCT items to be task-specific bearing in mind the requirements of the task students are expected to complete.

Lastly, Beghetto and Karwowski [11] reported on the dynamic nature of creativity and urged researchers to adopt a dynamic approach to assessing creativity. They call for measurements of creativity to be conducted at much shorter intervals to capture expressions of creativity that are transient, not acted upon or not acknowledged during the activity. To do so, close monitoring of participants' behavior during an integrated STEM activity would need to be conducted. Crismond and Adams [21] designed a matrix that captured the various phases of integrated STEM learning and identified how creativity is expressed in each of these phases. This Informed Design Teaching and Learning Matrix formed the basis for the rubrics used in this study's "micro-longitudinal measurements" of creative behavior [11, p. 8].

Thus, the research questions guiding this study are: (1) Is the modified SCT useful in measuring changes in scientific creativity?, and (2) How can the SCT and behavior-monitoring be used as creativity assessment tools for students?

3. Methods

Approval for the study was granted (IRB No. IRB-2019-09-004-05). The study involved 17 pre-service teachers, 11 females and six males ranging from ages 20 to 23, who are undergraduates in a Bachelor of Science (Education) course from the National Institute of Education at the Nanyang Technological University. Before and after the activity session, participants individually completed Hu and Adey's SCT [9] that was modified to be more domain- and task-specific to the context of the vertical farming integrated STEM learning activity.

This vertical farming integrated STEM learning activity was first developed by Tan et al. [22] for use in schools to bring together the scientific concepts of photosynthesis, engineering, and technological expertise, and the real-world problem of food production in a land-scarce country like Singapore. For the purposes of this study, it was assumed that participants already had a baseline understanding of photosynthesis and food production in Singapore, and the series of lessons on photosynthesis and agriculture was omitted. Participants were grouped into groups of four or five. On separate sessions, each group was provided with the same materials (e.g., cardboard, plastic bottles, ice cream sticks, aluminium foil, straws, string, metal wires) and tasked to craft a prototype of a vertical farming system to address a problem of their choice related to vertical farming in Singapore. The effectiveness of their vertical farm prototype was tested using chlorophyll extract trapped in calcium alginate gels and acidified potassium permanganate, and measured by the time taken for the purple acidified potassium permanganate to decolorize. Groups were then tasked to discuss improvements and modifications they would make to their current prototype before identifying and carrying out modifications. Their revised prototype was then subjected to the same test as before, and a third round of feedbacking and prototyping occurred. The activity ended either when groups completed three rounds of prototype testing or the three-hour limit lapsed. The activity session was recorded and reviewed to monitor participants' creative behavior as detailed in section 3.2.

The following section details the modifications that were done to Hu and Adey's SCT [9] and how the items were scored.

3.1. Domain- and task-specific SCT

3.1.1. Adapting the SCT

There were seven items in the SCT by Hu and Adey [9], of which six items were adapted to be more domain- and task-specific within the context of vertical farming. The seven items in the original SCT and the adapted SCT can be found in Table 1.

For item one, students were to list possible uses of aluminium foil to assess their creativity in applying science knowledge. The item in the original SCT was "glass" and the change was made to include "aluminium foil", which was a material provided during the task, thus making the question task-specific.

For item two, students had to generate research questions to assess their creativity in exploring possible problem statements. The context of "space exploration" in the original was replaced with "vertical farm" to make the item more specific to the domain of vertical farming rather than science in general.

For item three, students needed to use their creativity to improve an existing technical product. Instead of the original item, the product was changed from "bicycle" to "table lamp", which was a piece of equipment used in the vertical farming integrated STEM activity. This made the item more domain-specific as light is necessary for photosynthesis, and more task-specific as students would use the equipment during the vertical farming STEM activity.

For item four, students had to consider the effects of living in a world without the sun, thus assessing their creativity related to a scientific phenomenon. Instead of "gravity", "the sun" is more relevant to the domain and task of vertical farming due to the importance of light in farming, as compared to gravity which has limited immediate relevance to the farming activity.

Item five from Hu and Adey's SCT [9] was not modified as the question on dividing a square into four equal shapes is relevant in the context of vertical farming in which the abilities of spatial visualization and creative use of space are tested. The item assesses the creativity of the student in approaching a problem.

For item six, students had to propose scientific methods for comparing two kinds of metal. The item thus assesses the students' creativity in investigating scientific phenomena. "Metal" was chosen to replace "tissue paper" from the original SCT as metal would be a more-relevant material in vertical farming than tissue paper when juxtaposed to real-world applications, thus adapting item six to be more domain- and task-specific.

For item seven, students were tasked to design a water sprinkler system, replacing an apple-picking system in the original SCT. The item thus assesses students' creativity in designing a technical product. Apple-picking is not a topic relevant to vertical farms which involves smaller crops instead of trees, making a sprinkler system a more relevant topic due to the importance of water distribution to crops during farming.

Table 1. Original and adapted Scientific Creativity Test [9] that is domain- and task-specific to a vertical farming integrated STEM learning activity.

Item no.	Original item	Adapted item
1	Please write down as many scientific uses as possible for a piece of glass. For example, to cover the opening of bottles.	Please write down as many scientific uses as possible for a piece of aluminum foil. For example, to cover the opening of bottles.
2	If you had a chance to visit outer space, what scientific questions would you want to research? Please list as many as you can. For example, what is the highest number of crops you can grow per unit area?	If you had a chance to visit a vertical farm in Singapore, what scientific questions would you want to research? Please list as many as you can. For example, what is the highest number of crops you can grow per unit area?
3	Generate as many possible improvements as you can to a bicycle, making it more interesting, more useful, and more beautiful.	Generate as many possible improvements as you can to a table lamp, making it more interesting, more useful, and more beautiful.
4	Suppose there was no gravity. Describe what the world would be like? For example, we will need torches to see.	Suppose there was no sunlight. Describe what the world would be like? For example, we will need torches to see.
5	Divide a square into four equal pieces (same shape). Generate as many possible methods as possible. Draw it on the answer sheet.	Divide a square into four equal pieces (same shape). Generate as many possible methods as possible. Draw it on the answer sheet.
6	There are two kinds of tissue paper. How can you test which is better? Please write down as many possible methods as you can, including the instruments used, principles, and simple procedure.	There are two kinds of metals. How can you test which is better? Please write down as many possible methods as you can, including the instruments used, principles, and simple procedure.
7	Design an apple-picking system. Draw a picture, pointing out the name and function of each part.	Design a water sprinkler system. Draw a picture, pointing out the name and function of each part.

3.1.2. Scoring the SCT

Items one to four were scored based on the fluency, flexibility, and originality displayed, while items five to seven were graded on the flexibility and originality. Fluency scores the number of responses generated, flexibility scores the ability to think in many categories, and originality scores the statistical rarity of the response [9].

Fluency in items one to four was graded simply by counting the number of responses with each response worth one point.

Flexibility in items one to five were graded by counting the number of categories an individual generated responses in with each category worth one point. In item six, each correct method was worth up to three points: one point each for principle, instrument, and procedure. In item seven, each unique function of the water sprinkler was awarded three points.

Originality was scored via the method of calculating the statistical rarity (i.e., frequency, f) of the response and assigning a score to the frequency range following Hu and Adey [9]. For items one to four, two points were awarded if the frequency was less than 5% ($f < 0.05$), one point was awarded if the frequency was between 5% and 10% ($0.05 < f < 0.1$), and zero points were awarded if the frequency was more than 10% ($f > 0.1$). For item five, three points were awarded if the frequency was less than 5% ($f < 0.05$), two points were awarded if the frequency was between 5% and 10% ($0.05 < f < 0.1$), and one point was awarded if the frequency was more than 10% ($f > 0.1$). For items six and seven, four points were awarded if the frequency was less than 5% ($f < 0.05$), two points were awarded if the frequency was between 5% and 10% ($0.05 < f < 0.1$), and one point was awarded if the frequency was more than 10% ($f > 0.1$).

3.2. Micro-longitudinal measurement of creative expressions

The integrated STEM activity session was recorded, and a transcript of the conversations and actions was generated. The transcript and video were analysed against the checklist generated based on Crismond and Adams's [21] Informed Design Teaching and Learning Matrix and aligned to the traits of fluency, flexibility, and originality. The checklist can be found in Table 2.

Fluency being expressed during the integrated STEM activity was measured in terms of the number of ideas generated during the activity (first criterion). Both ideas used and not acted upon are counted to better reflect the individual's dynamic creativity. The second criterion was the ability of the individual to facilitate discussion and hence promote the fluency of others [21]. Furthermore, Crismond and Adams also argued that the ability to set aside one idea for another (third criterion) is a sign of fluency, as clinging onto one idea would limit the potential for generating more ideas.

Flexibility during the integrated STEM activity was the ability to generate ideas in the various categories of the prototyping process, namely problematizing, ideating, representing problems or ideas, and the use of materials [21]. These categories form the four criteria creative behaviors could fall under for flexibility. The more areas that participants provided inputs for, the more categories the participant was able to think in and the more flexible the participant was.

Originality was earlier defined to be the statistical rarity of the responses [9], but this was difficult to determine from the activity and transcript. Rather, the first criteria of originality looked at the ability of participants to generate responses. Second, credit was given to students who could articulate reasons for their response being creative (either novel or useful) as part of the definition of creativity. Third, participants who could generate especially novel and/or useful responses (at the discretion of the assessor following the flowchart in Figure 2) were also credited for originality.

Table 2. Checklist used to analyze the transcript and video recording of the vertical farming integrated STEM learning group task based on the Informed Design Teaching and Learning Matrix [21].

Creative trait	Criteria
Originality	1. Able to raise problems/ideas/methods of representation/ways to use the materials provided during the discussion.
	2. Able to articulate reasons for the problems/ideas/methods of representation/ways to use the materials provided being useful/novel.
	3. Able to add value by generating especially novel and/or useful ideas.
Fluency	1. Able to generate more than the average number of problems/ideas/methods of representation/ways to use the materials provided (including those not acted upon).
	2. Able to facilitate discussion that leads to the group raising more ideas (idea of leading the others toward being more fluent).
	3. Able to set aside a current idea to explore other ideas; willing to consider different options.
Flexibility	1. Able to set aside a current idea to explore other ideas; willing to consider different options.
	2. Able to generate ideas (including those not acted upon).
	3. Able to represent problems/ideas in different modes (at least two of the following: verbal, textual, diagrammatic, demonstration/modeling with materials).
	4. Able to suggest uses for the materials provided.

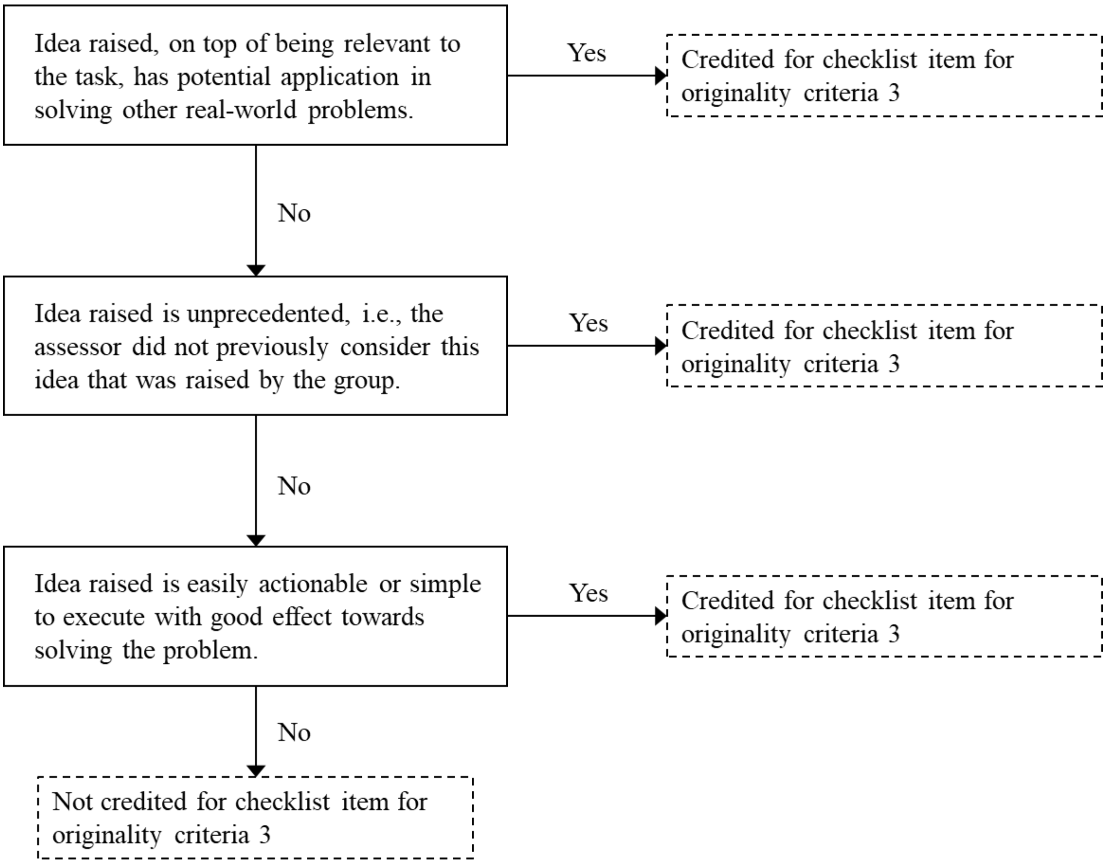

Figure 2. Flowchart used to assess if an idea is considered especially novel and/or useful.

3.3. Data analysis

Responses for both the pre-activity and post-activity SCT were graded. Pre-service teachers' pre-activity and post-activity tests were compared to determine if the (i) total score, (ii) originality, fluency, and frequency scores, and (iii) item scores had changed (change = post-test score – pre-test score). For behavior-monitoring, participants were ranked by the number of creative behaviors identified from the transcript and video recording based on the checklist in Table 2. Based on the rankings, participants were then grouped into high, intermediate, and low levels of creativity with five, seven, and five participants, respectively. The transcript and video recordings of participants in each of the three groups was then reviewed to identify possible trends in the types of creative behaviors associated with the three groups. For both assessment tools, statistical analysis was not conducted due to the small sample size of the study.

4. Results

4.1. Results of the pre- and post-activity SCT

Of the 17 participants, 10 had a lower post-activity SCT score, six had a higher post-activity SCT score, and one had the same post-activity SCT score as the pre-activity SCT score (Figure 3).

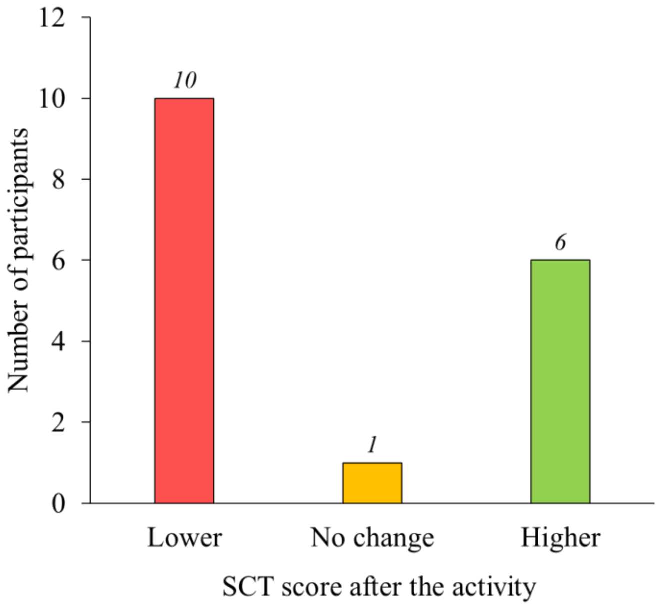

Figure 3. Number of participants (in italics) with a lower, the same, and a higher total score in the adapted Scientific Creativity Test (SCT) after the integrated STEM learning activity.

We then looked at the score breakdown for the three dimensions of originality, fluency, and flexibility (Figure 4). For originality, 11 participants had a lower post-activity score, and six had a higher post-activity score. For fluency, 10 participants had a lower post-activity score, two had no change in scores, and five had a higher post-activity score. For flexibility, six participants had a lower post-activity score, one had no change in scores, and 10 had a higher post-activity score. All in all, more participants experienced a decrease in originality and fluency scores while more participants experienced an increase in flexibility scores.

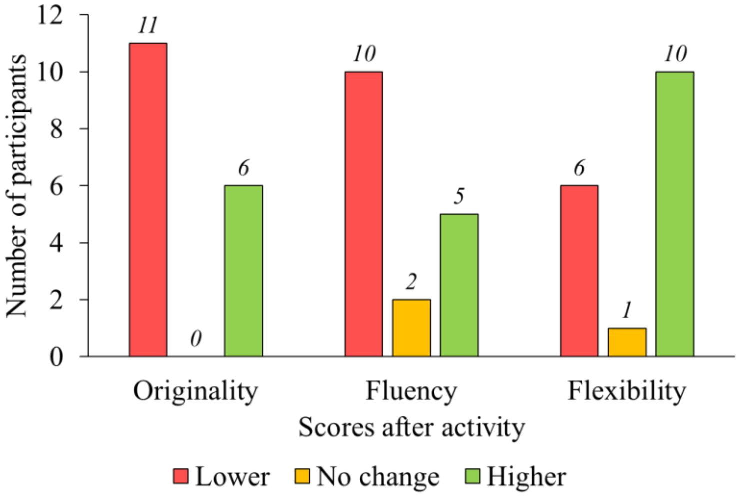

Figure 4. Number of participants (in italics) with a lower, the same, and a higher score for the three dimensions of originality, fluency, and flexibility after the integrated STEM learning activity.

Next, the number of participants with a decrease, increase, or no change in scores was counted for each test item (Figure 5). For item 1, 11 participants had a higher post-activity score, one participant had the same score as before, and five participants had a lower post-activity score. For item 2, nine participants had a higher post-activity score, one participant had the same score as before, and seven participants had a lower post-activity score. For item 3, eight participants had a higher post-activity score, one participant had the same score as before, and eight participants had a lower post-activity score. For item 4, five participants had a higher post-activity score and 12 participants had a lower post-activity score. For item 5, seven participants had a higher post-activity score, three participant had the same score as before, and seven participants had a lower post-activity score. For item 6, four participants had a higher post-activity score and 13 participants had a lower post-activity score. For item 7, nine participants had a higher post-activity score and eight participants had a lower post-activity score.

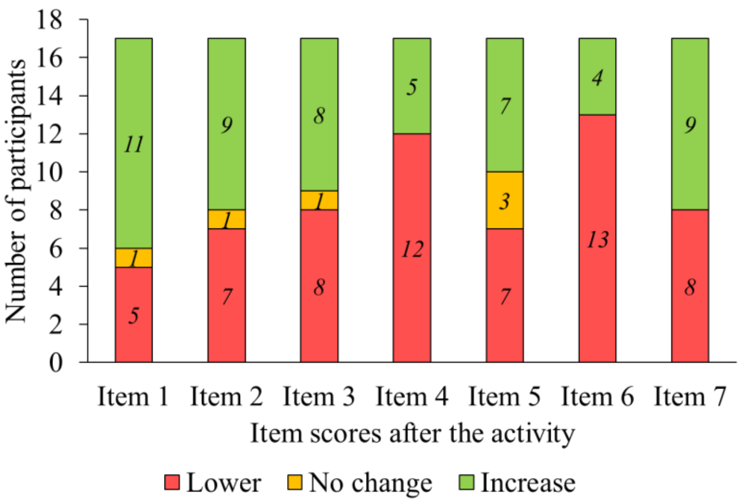

Figure 5. Number of participants (in italics) with a lower, the same, and a higher score for each adapted Scientific Creativity Test (SCT) item after the integrated STEM learning activity.

4.2. Behavior during the integrated STEM activity

Creative behaviors of participants were monitored for the entire session using video recordings of the activity sessions. The following case stories report the behaviors of three students—William, Queenie, and Glenda (pseudonyms)—identified as having high, low, and intermediate levels of creativity, respectively, and identify which of their behaviors are considered "creative behaviors".

4.2.1. William—High level of creativity

First, William was the most creative participant of the 17 during the integrated STEM learning and showed high improvement between the pre- and post-activity SCTs. During the activity, William generated the highest number of ideas of the 17 participants and could generate a wide variety of responses in the problem, ideas, and use of materials categories throughout the activity.

William displayed the ability to lead the group in generating ideas. This was displayed during the discussions at the beginning of the activity and during the problem-solving phase. William's approach was to invite responses from teammates through statements such as "Should we draw the prototype?". Such behavior is congruent with Crismond and Adam's [21] characterizing fluency as facilitating idea generation in others in addition to self-generation of ideas.

On top of generating more novel ideas (36) than the other participants (his average number of ideas generated = 16), William also substantiated ideas that were generated, providing reasons that would make these ideas useful. This is in line with the definition of creativity that ideas are both novel and useful [17]. One example of this trait is as follows:

> William: We can consider the idea of hanging the water bottle [on its side]. Because, if we were to stick the vial [into upright halves of the bottle], it will fall out. As in, if the bottle opens up, then imagine when you spin the bottle—the vials will fall out. So, if you can [instead] suspend the water bottle halves [on their side], the vials will not fall out.

To support the idea of hanging the water bottle in their set-up, William explained the risk of having the contents that were to be placed in the set-up falling out in the alternative scenario in which the bottles were not suspended but attached firmly to the structure. This is an example of William supporting the ideas proposed by his group members by adding value and weight to the idea being proposed.

Lastly, William was observed to be constantly engaged in discussion with teammates. It was during such discussions between teammates that creative responses were generated. In fact, many of such discussions were initiated by William, who also prompted other members of the group to provide their inputs. Excerpt 1 shows the exchange of ideas between William, Queenie, and Nadia that were initiated by William.

> <u>Excerpt 1. Alternative use of aluminium foil</u>
> William: I don't think we have a lot of wire, but we can roll up the aluminium foil such that it's very thin, such that you can hold the carriage at four different points and hook it up.
> Nadia: Oh, okay, I understand.
> Queenie: So, it becomes a basket?
> …

William: Yes, yes, yes. But hooked at four points. I think two points of suspending the carriage is not balanced.

To sum up William's engagement, William generated a high number of creative responses in a variety of categories. It was also observed that William was also able to lead and spur other members in being more fluent, i.e., group members providing their own inputs during discussions led by William. Ideas were substantiated with reasoning, and William contributed to discussions actively and offered alternatives to improve promising ideas throughout the process.

4.2.2. Queenie—Low level of creativity

Queenie was the least creative of the 17 participants, having generated the least number of creative responses during the activity (five). She also did not improve between the pre- and post-activity SCT and had deteriorated slightly instead. It was observed that throughout the session, Queenie was generally passive and was less involved in the discussions during idea generation and problem-solving.

One notable display of creativity was when Queenie substantiated ideas as follows: "I feel like the cardboard should be our foundation. It will be a more sturdy base.", explaining why a cardboard base is superior to a base made using ice-cream sticks for the vertical farming structure. Otherwise, Queenie generally had little creative input to the problems, ideas, and use of materials, and was silent throughout the session and less engaged in team discussions. Queenie instead focused more on helping to build the set-up.

4.2.3. Glenda—Intermediate level of creativity

Glenda was moderately creative with an average number of creative responses covering the various categories of problems, ideas, and use of materials. She also showed some improvement in the post-activity SCT scores. Though not to the same extent as William, Glenda was able to substantiate ideas and bounce off creative responses during the discussions. One example is shown in Excerpt 2 below:

Excerpt 2. Substantiation of teammate's ideas

William: What if you just make the stick longer—you connect another stick to the other side of the structure? I was thinking this, so that you can connect the two satay sticks together. Is that possible?

Glenda: I think that would be better. 'Cause then, we will have more space for the carriage and the vial.

In this example, Glenda was explaining the reason that would make William's suggestion useful and shows how Glenda worked with others on the team to improve the ideas suggested. Glenda, in general, generated an average number of creative ideas (16), participated in discussions, and could substantiate ideas that were raised.

4.2.4. Rise above

Having presented the three case stories of participants with the highest, lowest, and intermediate levels of creativity, we will now present an overview of the creative behaviors observed in participants during the vertical farming integrated STEM activity.

Participants who displayed higher levels of creativity typically generated more responses, could substantiate their ideas with reasons, and could generate more responses for each of the three categories of problem, idea, and use of materials. This was contrasted with participants with lower creativity levels, generating fewer responses and in fewer categories, as well as substantiating their ideas to a lesser extent.

One striking observation made was the importance of social interaction during the discussions. The integrated STEM activity involves problem-solving and building a prototype, and participants go through the stages of problem-framing, ideating, building, and testing. In all of these stages, discussion and collaboration are fundamental to generating ideas that can be acted upon to complete the task. Participants with higher levels of creativity were observed to have been more participative during these discussions, which contrasts with participants with lower levels of creativity, who were observed to be more reserved and provided fewer inputs during these discussions. Furthermore, the participants that led discussions and hence influenced others to be more fluent generally had better performance than other participants and are, thus, more creative.

5. Discussion

5.1. Decrease in SCT scores

The decrease in total SCT scores (Figure 3) could indicate a decrease in scientific creativity after participation in the integrated STEM activity. A larger proportion of participants had lower post-activity originality and fluency scores (Figure 4), which suggests a decrease in the novelty and quantity of unique ideas from participants, respectively. Frequency scores showed that more participants had a higher score post-activity, which suggests an increase in the ability to think broadly in the different categories of ideation for the task.

Based on the item analysis in Figure 5, a larger proportion of participants showed increased creativity related to scientific knowledge (assessed by item 1) and decreased creativity related to scientific phenomena (assessed by items 4 and 6) after the STEM activity. The former could suggest that the STEM activity is a useful tool in helping students learn scientific knowledge while the latter could suggest that the STEM activity can limit a students' ability to explore a variety of scientific phenomena. The similar number of participants with higher or lower scores for items assessing creativity related to scientific problems (items 2 and 5) and the technical product (items 3 and 7) suggests that the STEM activity does not help students become more creative in generating scientific problems and improving technical products.

However, these results should not lead us to conclude that the integrated STEM activity had an overall negative effect on scientific creativity considering the low sample size of 17. Furthermore, Barbot et al. [10] reported that studies using TTCT-derived tests with the same task items for the assessment of creativity longitudinally may be invalidated by biases. Participants may recall their responses for their previous attempt and either provide only these responses or only provide new responses, both of which could prevent the SCT from capturing true changes in creativity after participation in an integrated STEM activity. Such "memory" effects can contribute to the decrease observed in SCT performance of the participants in this study.

It may appear that the solution to this challenge in using the SCT for longitudinal assessments of creativity is to change the test items for the second attempt. This, however, runs into another issue – comparing the first and second tests may not be valid as the test items have changed [10]. While it is possible to generate new tasks that are comparable to those in the first test, further studies will need to be conducted to ascertain whether different test items can be said to be comparable and to validate the use of two sets of different SCT test items to monitor changes in scientific creativity.

All in all, the usefulness of the SCT in pre- and post-activity assessment of creativity might be limited, but there is value in the use of DT tests such as the SCT for snapshot measurements of an individual's ability to generate useful and novel ideas.

5.2. Behavior during the integrated STEM activity

Creative behaviors of participants during an integrated STEM activity are more varied and dynamic. Divergent thinking is measured in the generation of creative responses during the activity. However, beyond measuring the divergent thinking ability of participants, monitoring creative behavior during an integrated STEM activity simulates the workings of real life in which creativity is required in social settings. The simulation of the real-world environment in which creative expressions are generated is the major benefit of integrated STEM activities as an assessment tool over the SCT.

Earlier, it was established that creativity is multifaceted, and creative behavior is a product of the various skill sets working together to generate an output [6,10,16,17]. In the social and interactive environment of the integrated STEM activity, other skill sets work in tandem with divergent thinking in producing the creative behavior of participants that are observed. These skill sets include the confidence in expressing your ideas, filtering the numerous ideas and coming to a consensus on which idea to act upon, and the openness to the ideas of others.

These skill sets were reported by Treffinger et al. [22] as three of four ways creativity is expressed: (1) divergent thinking, (2) analytical and convergent thinking, and (3) confidence and receptivity to exploring new ideas. The first expression of divergent thinking has been explained earlier in this paper. Analytical and convergent thinking involves higher-order thinking needed to synthesize and evaluate the usefulness and novelty of ideas. Such skills are needed for a participant to substantiate their own or others' ideas and decide the best idea to act upon during the activity. The ability to substantiate ideas and converge onto a main idea to focus on were behaviors that were observed to be associated with higher levels of creativity in integrated STEM learning.

Confidence and receptivity in exploring new ideas were described by Treffinger et al. [23] as the expression associated with the tolerance of ambiguity and risk, adaptability, and curiosity among other traits. The integrated STEM activity measures this in the form of behaviors like setting aside one idea for another. Furthermore, the openness to take part in discussions can be seen as receptivity to new ideas from others in the team. This was one creative behavior that was evident in the vertical farming integrated STEM learning—participants who were more actively engaged in discussions and ideation could be seen as more confident and tolerant of the risk of ideas being rejected. Participants who were less confident would likely shy away from sharing ideas and participating actively in ideation and discussion, being seen as less creative in the process.

The last form of creative expression reported by Treffinger et al. [23] involves the individual's self-awareness and personality, dealing with one's perseverance, self-determination, and motivation.

Such characteristics are more nuanced and harder to observe through creative behaviors during an integrated STEM activity.

Though the observation of creative behavior during an integrated STEM activity is a powerful tool in the assessment of scientific creativity in students, it is a time- and resource-consuming method of assessment. Close observation of an individual needs to be done to extensively check for creative expressions occurring throughout the activity, and it is not feasible to conduct such observations on multiple students at a time. Large groups may make using this assessment tool tedious, but in small groups, observing students' creative behavior remains a comprehensive and informative assessment tool at our disposal.

Monitoring creative behavior is, however, difficult due to the time-intensive and manpower-intensive nature of monitoring and filtering through video footage and transcripts to identify instances where creative behaviors were displayed. Having identified specific behaviors associated with creativity during integrated STEM learning, it is recommended that the monitoring process be improved by refining the rubrics for easier use by teachers.

5.3. Use of both assessment tools in scientific creativity assessment

The two creativity assessment tools were observed to be non-substitutable for each other. This observation is supported by the lack of congruence between the scores in the SCT and the creative behavior displayed during the activity for the 17 participants. Furthermore, monitoring of the behaviors displayed during integrated STEM activities allows for the assessment of more skill sets related to creativity than just using the SCT alone. It can, thus, be argued that the two tools should not be seen as substitutes for each other as a single measure of creativity assessment.

We should consider when the two assessment tools should be used to measure creativity in students. The use of the SCT alone is justified when there is a large number of participants for which scientific creativity is to be assessed. The administration and scoring of the SCT are significantly less resource-intensive compared to monitoring creative behaviors during an integrated STEM activity. The use of the SCT is also justified when measuring DT ability is the focus of the assessment. Results from the SCT can then be used to give assessors a general idea of the DT abilities of students. DT abilities can be used as a quick estimate of students' creativity, but teachers should bear in mind that DT is just one of the many skills of creativity.

For the case when there are few students for which creativity is to be assessed, or when a better creativity profile is needed, monitoring creative behavior during the integrated STEM activity would be a better assessment tool. The time- and manpower-intensive nature of monitoring student behavior for an extended duration of time limits the use of behavior-monitoring for assessing creativity. This is made more difficult by the various forms the creative behavior can take—verbal, written, diagrammatic, or demonstrations during the activity.

5.4. Limitations

The present study has a small sample size of 17 participants, which prevents the use of statistical analysis for the results of the SCT. Furthermore, the small sample size could limit the number and types of creative behaviors displayed by participants during the integrated STEM task. Trialing the two assessment tools with a larger group of participants would thus allow for statistical analysis to be conducted and for more creative behaviors and group behaviors to be identified during the integrated

STEM task.

While the original SCT has been assessed for its construct validity as a tool that assesses scientific creativity, validity tests were not conducted for the checklist used when monitoring students' creative behaviors during an integrated STEM task. Additionally, inter-rater reliability needs to be established to determine if the checklist can be used by different educators and researchers to provide consistent and reliable measurements of students' creativity.

6. Conclusions

In summary, this study designed two creativity assessment tools—the SCT and behavior-monitoring rubric for integrated STEM activities and trialed the tools on pre–service teachers. Results of the trial suggested a lack of substitutability of the two assessment tools as behavior-monitoring during integrated STEM activities is a more powerful tool that can measure other skill sets associated with creativity beyond DT. Considering the resource-intensive nature of monitoring student behavior during an integrated STEM activity, we recommend that this assessment tool be used only for smaller groups of students. For cases in which the creativity of many students is to be assessed, DT tests like the SCT should be used instead. This study thus provides two task-specific creativity assessment tools for STEM educators and researchers to consider in their future work on STEM learning and assessment of scientific creativity. Future studies should investigate if the use of the SCT in pre- and post-activity testing with different test items is a valid method of assessment for a larger sample of participants. The rubrics used in monitoring the creative behavior of students can also be refined given that specific behaviors associated with creativity during integrated STEM learning have been identified in this trial study, though larger-scale studies would be more informative on the spectrum of behaviors associated with creativity. These studies will help to refine the two assessment tools presented in this study, thereby allowing for a more comprehensive assessment of students' scientific creativity as a proxy measure of the effectiveness of STEM instructional programs.

Author contributions

Hong Liang Lee: Data curation, Formal analysis, Writing – original draft; Aik-Ling Tan: Conceptualization, Supervision. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to acknowledge the funding support from the National Technological University URECA Undergraduate Research Programme and ERFP OER24/19TAL for this research project. The guidance and insight provided by their supervisor A/P Tan Aik Ling was valuable and spurred informative discussion on the topic of creativity in science, helping deepen their learning about the research process and the field of creativity. They would also like to acknowledge the help provided by Toh Nguk Hwee David and Tan Geok Suan, Higher Technical Specialists of the National Institute of Education (NIE) Natural Sciences and Science Education (NSSE) Department, for their support during the data collection process. They would also like to thank the participants for their time and participation.

Conflict of interest

The authors declare no conflict of interest.

Ethics declaration

This work has received the ethics board approval IRB No. IRB-2019-09-004-05.

References

1. Bryan, L.A., Moore, T.J., Johnson, C.C. and Roehrig, G.H., Integrated STEM education. In *STEM road map: A framework for integrated STEM education*, C.C. Johnson, E.E. Peters-Burton, & T.J. Moore, Eds., 2015, 23–27. Taylor and Francis Inc.
2. English, L.D., Advancing elementary and middle school STEM education. *International Journal of Science and Mathematics Education*, 2017, 15: 5–24. https://doi-org.libproxy.nie.edu.sg/10.1007/s10763-017-9802-x
3. Bybee, R.W., Advancing STEM education: A 2020 vision. *Technology and Engineering Teacher*, 2010, 70(1): 30–35.
4. Ugras, M., The effects of STEM activities on STEM attitudes, scientific creativity and motivation beliefs of the students and their views on STEM education. *International Online Journal of Educational Sciences*, 2018, 10(5): 165–182. https://doi.org/10.15345/iojes.2018.05.012
5. Carruthers, L. and MacLean, R., The dynamic definition of creativity: Implications for creativity assessment, in *Dynamic perspectives on creativity: New directions for theory, research and practice in education*, R.A. Beghetto, & G.E. Corazza, Eds., 2019, 207–223. Springer.
6. Kaufman, J.C. and Beghetto, R.A., Beyond big and little: The four C model of creativity. *Review of General Psychology*, 2009, 13(1): 1–12. https://doi.org/10.1037/a0013688
7. Baer, J., The importance of domain-specific expertise in creativity. *Roeper Review*, 2015, 37(3): 165–178. https://doi.org/10.1080/02783193.2015.1047480
8. Kaufman, J.C., Plucker, J.A. and Baer, J., *Essentials of Creativity Assessment*, 2008. Hoboken, NJ, USA: John Wiley & Sons, Inc.
9. Hu, W. and Adey, P., A scientific creativity test for secondary school students. *International Journal of Science Education*, 2002, 24(4): 389–403. https://doi.org/10.1080/09500690110098912
10. Barbot, B., Lubart, T.I. and Besancon, M., "Peaks, slumps and bumps": Individual differences in the development of creativity in children and adolescents, in *Perspectives on creativity development. New directions for child and adolescent development*, Barbot B. Ed., 2016, 33–45. Hoboken, NJ, USA: John Wiley & Sons, Inc.

11. Beghetto, R.A. and Karwowski, M., Unfreezing creativity: A dynamic micro-longitudinal approach. In *Dynamic perspectives on creativity: New directions for theory, research and practice in education*, R.A. Beghetto, & G.E. Corazza, Eds., 2019, 7–25. Berlin, Germany: Springer.
12. Ministry of Education, Singapore (MOE), *21st century competencies*, 2021, October 18. Ministry of Education, Singapore. Retrieved from: https://www.moe.gov.sg/education-in-sg/21st-century-competencies
13. World Economic Forum, *Schools of the future: Defining new models of education for the fourth industrial revolution*. 2020, World Economic Forum 2020. Available from: https://www3.weforum.org/docs/WEF_Schools_of_the_Future_Report_2019.pdf
14. Kanlı, E., Assessment of creativity: Theories and methods, in *Creativity - a force to innovation*, P. Jain, Ed., 2020. IntechOpen. http://dx.doi.org/10.5772/intechopen.93971
15. Sternberg, R.J., What's wrong with creativity testing? *The Journal of Creative Behaviour*, 2018. https://doi.org/10.1002/jocb.237
16. Corazza, G.E., Potential originality and effectiveness. *Creativity Research Journal*, 2016. https://doi.org/10.1080/10400419.2016.1195627
17. Plucker, J., Beghetto, R.A. and Dow, G., Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity research. *Educational Psychologist*, 2004, 39: 83–96.
18. Amabile, T.M. and Pillemer, J., Perspectives on the social psychology of creativity. *Journal of Creative Behaviour*, 2012, 46(1): 3–15. https://doi.org/10.1002/jocb.001
19. Baer, J., Theory in creativity research: The pernicious impact of domain generality, in *Creativity under duress in education?: Resistive theories, practices, and actions*, C.A. Mullen, Eds., 2019, 119–135. Berlin, Germany: Springer.
20. Mullen, C.A., Creative synthesis: Combining the 4C and systems models of creativity, in *Creativity under duress in education?: Resistive theories, practices, and actions*, C.A. Mullen, Eds., 2019, 3–25. Berlin, Germany: Springer.
21. Crismond, D.P. and Adams, R.S., The informed design teaching and learning matrix. *Journal of Engineering Education*, 2012, 101(4): 738–797. https://doi.org/10.1002/j.2168-9830.2012.tb01127.x
22. Tan, A.L., Ng, Y.S., Koh, J., Ong, Y.S. and Tan, D., Applying concepts of plant nutrition in the real world: Designing vertical farming systems. *Science Activities*, 2022, 60(1): 25–31. https://doi.org/10.1080/00368121.2022.2138249
23. Treffinger, D.J., Young, G.C., Selby, E.C. and Shepardson, C., *The National Research Center on the Gifted and Talented*: *Assessing creativity: A guide for educators*, 2002. University of Connecticut, University of Virginia, & Yale University, USA.

Author's biography

Mr. Hong Liang Lee is a secondary-school science teacher in Singapore. He is a graduate from the Natural Sciences and Science Education faculty of NIE-NTU and is interested in designing science lessons that develop 21st century competencies in students.

A/P Aik-Ling Tan is an Associate Professor of Science Education with the National Institute of Education, Nanyang Technological University, Singapore. She specializes in science education and STEM curriculum design.