Analysis of the Model Relating Creativity in Fermi Problems and General Creativity in Psychology and Mathematical Creativity

Hidemichi Okamoto[1*]
[1] University of Education Karlsruhe, Germany, Bismarckstraße 10 76133 Karlsruhe
*Corresponding author. Email: hidemichi.okamoto@ph-karlsruhe.de

Abstract
Recently, creativity in mathematical modelling has been studied. These studies have found a relationship between mathematical modelling and creativity. However, few correlational analyses of those associations have been conducted using quantitative measures. The present study, using Fermi problems, regarded as a type of mathematical modelling, examined whether there are correlations between creativity in the Fermi problem and general creativity in psychology and mathematical creativity by Structural Equation Modeling. The results of a survey of junior high school students (n = 364) in Japan showed a strong correlation (r = .711, p < .01 in the acceptable model between creativity in Fermi problems and mathematical creativity. A moderate correlation (r = .429, p < .01 in the acceptable model was also found between creativity in Fermi problems and general creativity in psychology. In addition, these correlations were shown to vary depending on the content and format of the Fermi problem.

Keywords: Fermi problem, General creativity in psychology, Mathematical creativity, Mathematical modelling, Structural equation modelling.

Introduction

Creativity is required in mathematics education (Mann, 2005; The European Commission, 2019). In addition, the importance of creativity in solving mathematical problems is shown (Silver, 1997; Mann, 2006). Many studies on the relationship between mathematics and creativity exist for these reasons. For example, some studies on mathematics teaching methods and learning environments foster creativity (Meyer, 1970; Nadjafikhah et al., 2011). Additionally, other researchers studied the relationship between mathematical creativity (hereinafter referred to as MC), which is defined as creativity specific to the mathematical domain, and mathematical ability, as well as how to evaluate these abilities (Balka, 1974; Mann, 2005; Kattou et al.,2013). However, due to the lack of insight into the relationship between general creativity in psychology (hereinafter referred to as GC) and MC, it is still not entirely clear how creativity and mathematics are related (Schoevers et al., 2020). Furthermore, the importance of creativity is emphasized, but it is not enough to develop creativity in mathematics education (Silver,1997; Thohari et al., 2020).

There is recent research on the relationship between creativity in mathematical modelling, which is the translation process between the real world and mathematics in both directions, and GC, to overcome this situation (Blum & Borromeo Ferri, 2009; Wessels, 2014; Lu & Kaiser, 2021). These studies evaluated students' solutions using a rubric corresponding to psychology's creativity factor. As a result, a specific relationship between mathematical modelling and creativity has been found. However, few studies have statistically analyzed the relationship and examined the structure. Therefore, the following research questions can be considered. "How is any creativity related to creativity in mathematical modelling?" Furthermore, "Does the content or form of the problem change its relevance to creativity?"

The present study uses Fermi problems as a mathematical modelling problem. Fermi problem is a type of mathematical modelling, and is viewed as a quick approximation for an open problem (Carlson, 1997; Greefrath & Frenken, 2021). In addition, it is required creative thinking (Silver,1997; Goel & Singh, 1998). The purpose of the present study is to analyze the relationship between creativity in Fermi Problem (hereinafter referred to as FC), GC, and MC. using structural equation modeling (hereinafter referred to as SEM). Based on the analysis results, it is discussed how the three types of creativity are related. Furthermore, it is considered how the creativity factor in the Fermi problem is viewed and the impact on correlations between FC, MC, and GC by the content and format of the Fermi problem.

Literature Review

General Creativity in Psychology (GC)

There is a wide range of research on creativity, including research that focuses on so-called genius individuals, methods of measuring creativity, and research on the development of creativity (Yano et al., 2002). From the 1960s, research on creativity using factor analysis by Guilford and others was published, and research on how to analyze and measure the factors of creativity became the most common. The focus has became not on individuals with high creativity but on the structure of creativity and general people's creativity. When thinking about creativity, it is necessary to define it. However, there is no single definition of creativity. Treffinger (2011) collected literature with definitions of creativity up to 2011. He collected more than 100 references, which discuss creativity from different perspectives. Thus, creativity can be studied from a variety of perspectives. The present study focuses on creativity factors and considers the related structure of each creativity that GC, MC, and FC, are consisted of these creative factors. Therefore, it focuses first on the study of Wilson et al. (1954) about factors in creativity. They hypothesized that creativity is composed of several factors. It was hypothesized that creativity has the following factors: sensitivity to problem, fluency, flexibility, originality, penetration, analysis, synthesis, and redefinition (Wilson et al, 1954).

Based on these hypotheses, tests were created to measure those factors (Guilford, 1959). Additionally, Torrance (1963) created a creativity test called the Minnesota Tests of Creative Thinking to measure these factors. Fluency is defined as the total number of meaningful ideas that were generated. Flexibility is considered to be the number of categories of the idea. Originality is weighted to those that are statistically rare per problem. Elaboration is measured by how much decoration was added to the idea. Other creativity tests that have been developed include the Test for Creative Thinking-Drawing Production or the Torrance Test of Creative Thinking (Urban & Jellen, 2010; Kim, 2006).

However, there are some criticisms of the definition of these factors. For example, some researchers pointed out the similarities between fluency and flexibility and integrated the two factors into a single definition (Hébert et al., 2002; Lu & Kaiser, 2021). Originality was also measured differently in each study. Most studies have been measured using percentage occurrences of ideas (Torrance, 1963; Mann, 2005; Kattou et al., 2013). However, the value of that cut-off point varies from study to study. As mentioned above, how the creativity factor is viewed is still debatable. Thus, it is worth considering how best to interpret the creativity factors in examining the relationship between the three types of creativity, GC, MC, and FC.

Mathematical Creativity (MC)

There is MC, which focuses on the specific domain of mathematics. As with creativity in psychology, its definitions vary widely. For example, Tammadge (1979) defined MC as "includes the ability to see new relationships between techniques and areas of application and to make associations between possibly unrelated ideas" (p. 151). Nadjafikhah et al. (2012) summarized the definition and characteristics of MC and the difficulty of defining it due to the complexity of its structure and characteristics. Because of this variety of ways of perceiving it, MC has also been studied from several perspectives. For example, Sriraman (2004) had a qualitative study in which mathematicians were targeted, and their thought processes were interviewed and analyzed. While some of these studies have focused on specific individuals with the high mathematical ability to find the creativity needed to solve mathematical problems, others have focused on students to measure their creativity in mathematics. Balka (1974) developed six criteria based on a list of items used to measure creativity in psychology and created questions to measure MC. Mann (2005) used research on such assessments to investigate the relationship between MC and mathematical ability among seventh graders attending middle school and found a positive correlation between them. Kattou et al. (2013) also studied the relationship between MC and mathematical ability using a statistical method called structural equation modeling. Furthermore, a study by Kattou et al. (2015) investigated the relationship between MC and creativity in psychology. These studies have revealed a relationship between mathematical ability and MC and between creativity in psychology and MC. Most of those studies have assessed creativity with three creativity factors: fluency, flexibility, and originality, but no consensus has been reached on how to measure these factors (Pitta-Pantazi et al., 2018).

Fermi Problem

"Fermi Problem" is named after the Italian nuclear physicist Enrico Fermi. He had a unique way of raising problems. Probably the most famous of them is "How many piano tuners are there in the city of Chicago?" Fermi problem is a type of mathematical modelling, which is the translation process between the real world and mathematics in both directions (Blum & Borromeo Ferri, 2009) and is viewed as a quick approximation for an open problem (Carlson, 1997; Greefrath & Frenken, 2021). Fermi problem has been used in educational practice as mathematical modelling. For example, Ärlebäck (2009) conducted educational research on a Fermi problem with high school students. Additionally, Peter-Koop (2005) practiced the Fermi problem with students at the primary stage of education. Since those studies have shown the positive effects of the Fermi problem on mathematical ability, it is considered valuable as a learning material for mathematics.

Fermi Problems is focused on from the perspective of creativity. Fermi problem requires fluency in creativity and encourages creative thinking (Silver,1997; Goel & Singh, 1998). In addition, Bennevall (2016) classified Fermi problems as a creative thinking type of problem, as mentioned by Wakefield (1992). Moreover, Greefrath and Frenken (2021) state that "Fermi problems can also typically be broken down into smaller problems that can be solved by estimating" (p. 56). This skill, which breaks down a problem into smaller pieces, requires creativity (Helps & Lunt, 2001).

Several studies have shown the relationship between Fermi problems and creativity. However, few studies have investigated the relationship between creativity in Fermi problems and factors of GC and MC. In addition, there is still a lack of research that examines the changes in associations between different problem contents and formats. There is also not much discussion of creativity, focusing on problems like Fermi problems and mathematical modelling (Wessels, 2014; Lu & Kaiser, 2021).

Research Questions

The main research question:

What are the relationships between FC, MC, and GC?

The following are two subordinate research questions:

1. How are the correlations differ depending on the content and form of the Fermi problem?
2. How are the correlations differ depending on which creativity factor is considered?

Methodology

This study aims to examine the correlations between FC, MC, and GC. Thus, a survey was conducted among junior high school students in Japan. The survey results were analyzed and examined for indexes of model fit of the hypothetical models using lavaan, a package for SEM software R.

Participants

A total of 364 students from a public junior high school in Japan participated. Their ages ranged from 12 to 15 years old and included 195 boys and 169 girls. The student's academic performance was the same as the national average for annual academic achievement surveys conducted throughout Japan. Additionally, mathematical modelling such as the Fermi problem was not used in school lectures, and students had little experience with such problems. It was indicated that the present study required a sample size of at least 347, with an anticipated effect size of 0.3, desired power of 0.8, 9 latent variable and 20 observed variables, and a probability level of 0.05 (Sloper, 2015).

Implement

Test for Creative Thinking-Drawing Production (TCT-DP)

The TCT-DP is a creativity test created by Urban and Jellen. In this test, creativity is viewed as a "whole," and the subject draws additional pictures on an unfinished drawing and gives it a title. The picture is then rated on 14 items (Urban & Jellen, 2010). The sum of the scores for these items is used as the creativity score. TCT-DP is good inter-rater reliability, with α = .81–.99 for the total score and $\alpha \geq .89$ for the test criteria (Urban & Jellen, 1996). This test in the present study was scored according to the manual (Urban & Jellen, 2010). This study does not consider the TCT-DP results as a single summed value. Because this study is designed from the perspective that creativity is composed of several factors, as Guilford (1959) and Torrance (1963; 1988) considered. Sola et al. (2017) corresponded the Guilford (1959) and Torrance (1988) and so on creativity factors to the 13 variables used to measure creativity in the TCT-DP (the Speed factor did not correspond). This analysis is based on fluency, flexibility, and originality in that study.

Mathematical Creativity Test

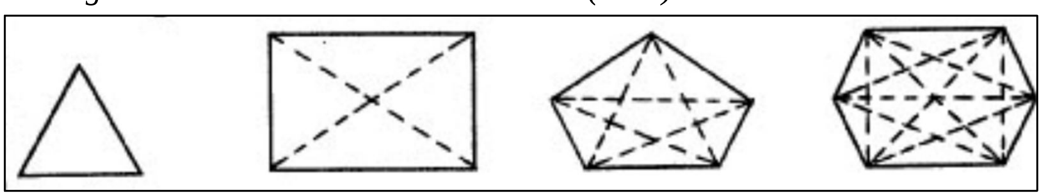
The present study adopted the mathematical creativity tests that were used in previous studies (Balka, 1974; Mann, 2005; Kattou et al., 2013; Pitta-Pantazi et al., 2013). The evaluation was also based on those studies. There are three assessment items: fluency, flexibility, and originality. Scores are assessed fluency (number of correct solutions), flexibility (number of different types or categories of correct solutions), and originality (it is a relative assessment, weighted so that the rarer the answer, the higher the score). The following are tests in this survey:

Problem 1 for MC (Mann, 2005):

> Below are figures of various polygons with all possible diagonals drawn (dotted lines) from each vertex of the polygon. List as many things as you can of what happens when you increase the number of sides on the polygon. For example, the number of diagonal increases. The number of triangles formed by the diagonal increases (p. 107).

Figure 1

The Figure Attached to the Problem in Mann (2005)

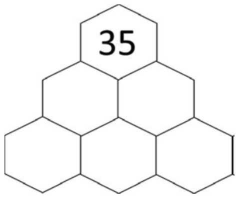


Problem 2 for MC (Kattou et.al, 2013):

> Look at this number pyramid. All the cells must contain one number. Each number in the pyramid can be computed by performing always the same operation with the two numbers that appear underneath it. Fill in the pyramid, by keeping on the top the number 35. Try to find as many solutions as possible (p. 172).

Figure 2

The Figure Attached to the Problem in Kattou et al. (2013)



Problem 3 for MC (Pitta-Pantazi et al., 2013):

Divide a 5cm by 5cm square into five parts of equal area. Find as many different solutions as possible (p. 203).

Fermi Problem Test

The following three problems were used in this survey:

- Fermi problem 1: How many liters of water does one person use in a year?
- Fermi problem 2: If you collected all the smartphones in the world, how many would there be?
- Fermi problem 3: If you collected all the cars in Japan, how many would there be? Think of as many ways as you can to find out how many cars there are. Write down as many ways as you can to find out how many cars there are, and write them down in as much detail as you can, using sentences, formulas, and diagrams. You do not have to calculate how many cars there are in Japan.

The definitions of creativity factor in FC and their measurement methods are given below. The present study focuses on fluency, flexibility, and originality, which are often used in MC and GC. Therefore, These Fermi problems were assessed by those three factors. Fluency is defined as "the richness of aspects solving a Fermi problem." For example, if a student considered "the amount of drinking water" and "the amount of water used in the shower" as the elements needed to solve the problem, thus the fluency is 2. Another response was, "I use 10 liters in the morning, 20 liters in the afternoon, and 30 liters at night." When the day was divided into three parts in chronological order, such as that response, the assumed number was three. Therefore, the fluency is 3. If there was no number to assume or only the answer, it was determined to be 0. Originality is scored by weighting the ideas according to their rate of occurrence regarding Kattou et al. (2013) methods. Flexibility is the number of categories of ideas. This author creates the categories. Because of the possibility that correlations may change depending on the categorization method, two types of categories are created. This way allows for a more detailed discussion for subordinate research question 2. One is divided into seven categories and the other into three (see. Appendix).

Procedure

All tests were conducted at school during regular school hours. At first, students received the Fermi problem tests, which were 10 minutes for each. Then, they took three Mathematical Creativity Tests. Next, students received TCT-DP. The students took the test relaxed, as it was explained to them that the results would not affect their school grades in mathematics or any other subject. Two teachers in this school supervised the survey.

Hypothesis for the research questions

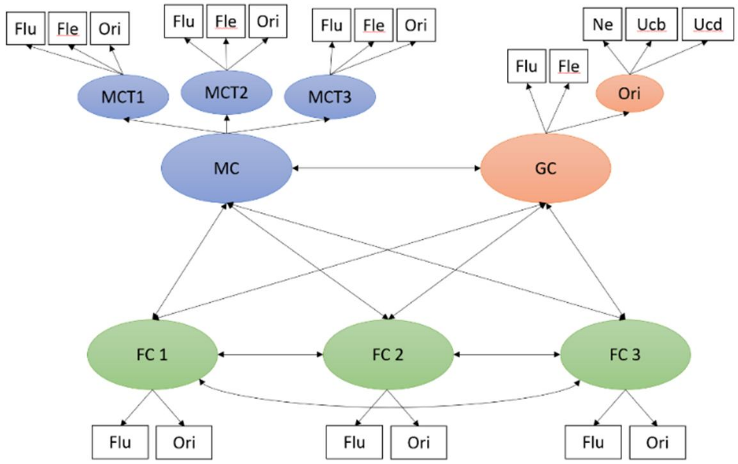
Below are the hypotheses for the research questions of the present study. Firstly, in answer to subordinate research question 1, it is hypothesized that the correlations differ depending on the content and format of the Fermi problem. Fermi problems 1 and 2 are predicted to be strongly associated with MC, while Fermi problem 3 is strongly associated with GC. This hypothesis is based on the idea that the form of the problem affects the expression of creativity. Fermi problem 3 requires only solution ideas and no computation. In other words, the format is similar to a general psychological test of creativity, which facilitates free ideas and does not require mathematical skills. On the other hand, Fermi problems 1 and 2 involve a computational process and require structuring ideas to solve the problem. The format of Fermi problems 1 and 2 is similar in format to tests of MC.

Secondly, in answer to subordinate research question 2, it is hypothesized that the more creativity factors to be measured, the better the model. As described in the Literature Review section, whether fluency and flexibility should be integrated or considered apart when measuring creativity is still discussed. Even in studies of creativity in mathematical modelling, Wessels (2014) defines fluency and flexibility separately, while Lu and Kaiser (2021) remove flexibility and just fluency. Based on the existence of flexibility measured by the number of categories of ideas in psychological and mathematical creativity tests, it is assumed that flexibility would be a better model to position for creativity in the Fermi problem as well. It is supposed that the correlation would change depending on how many layers of idea categories are divided. The difference between high and low flexibility in categories with fewer layers is not likely to be clearly expressed. Therefore, it is hypothesized that flexibility measured with more categories of ideas will be a better model than flexibility measured with fewer categories of ideas.

From now, it is described that hypothetical models should be considered in structural equation modeling based on the above hypothesis. Hypothetical models 1 through 3 are based on creativity in Fermi problem 1 (hereinafter referred to as FC1), creativity in Fermi problem 2 (hereinafter referred to as FC2), creativity in Fermi problem 3 (hereinafter referred to as FC3), for each problem, with latent variables are created (see Figure 3). In hypothetical Model 1, each FC is created by fluency and originality, excluding flexibility. In hypothetical model 2, each FC is created from flexibility measured in more hierarchical categories (7 layers), fluency, and originality. In hypothetical Model 3, each Fc is created from flexibility measured in fewer hierarchical (3 layers) categories, fluency, and originality.

Figure 3

Hypothetical Model 1

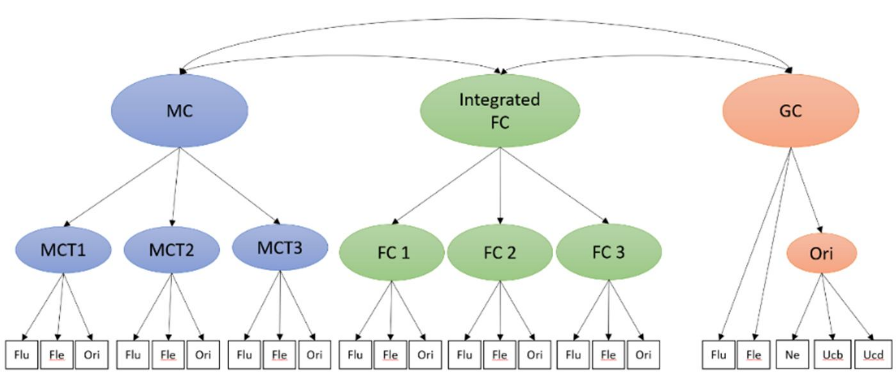


Note. Flu is fluency, Fle is flexibility, Ori is originality, Ne and Ucb and Ucd are sub - factors for originality in TCT-DP (Sola et al., 2017).

In addition to these hypothetical models, models including a latent variable (hereafter referred to as integrated FC), which further combines the latent variables FC1, FC2, and FC3 into one (see Figure 4), are also being considered for the analysis. creativity.

Figure 4

Hypothetical Model 5 and 6



In other words, hypothetical model 4 is a model of the relationship between the integrated FC, which consists of each FC that has fluency and originality and does not have the factor of flexibility, and the other creativity. Hypothetical model 5 is a model of the association between integrated FC, which is flexibility measured by the 7-layer category, fluency and originality, and the other creativity. Hypothetical model 6 is a model of the association between integrated FC, which is flexibility measured by the 3-layer categories, fluency and originality, and the other

Results

Table 1

Fitness Index in SEM for hypothetical models

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
p	.00	.00	.00	.00	.00	.00
χ^2	286.62	496.41	474.55	295.24	505.26	483.42
df	160	220.00	220.00	164.00	224.00	224.00
CFI	.97	.95	.95	.97	.95	.95
RMSEA	.04	.05	.05	.05	.05	.05
SRMR	.05	.06	.06	.05	.06	.06
AIC	19193.76	22665.37	21935.37	19194.371	22666.22	21936.24
BIC	19388.61	22883.61	21975.95	19373.641	22868.88	22138.89

Results of the SEM model fit index showed that all models are acceptable (see. Table 1). Firstly, the chi-square test shows a significant difference in all models. The chi-square test of exact fit often rejects the null hypothesis, especially in large samples. Therefore, while it would generally be desirable for the null hypothesis not to be rejected, in the present analysis, it was rejected in all models, possibly due to sample size. (Joreskog & Surbom, 1996; Shi et al, 2019). Secondly, the model is considered good if the CFI is more excellent than .95 (West et al. 2012). In other words, all models were concluded to be good models. Thirdly, REMSA is ideally less than .05 (MacCallum et al., 1996). In this case, all models were shown to be good fitting. Fourth, Hu and Bentler (1999) suggested that the value of SRMR less than .06 alone, or .08 in a combination of other SEM indexes within recommended ranges, is a good fit. In this index, all models showed promising results. Furthermore, focusing on the AIC, Models 1 and 4, the FC models with no flexibility factor, have relatively lower values than the other models. Smaller values of AIC are judged to be better models (Akaike, 1974). In addition, the Voung test showed that there were significant differences ($p < .01$) between models without flexibility and with flexibility measured in two types of categories (Genius & Strazzera, 2002).

Then, the correlation coefficients for each hypothetical model are summarized (see. Table 2 – 7). Firstly, the correlation coefficients in hypothetical models 1 through 3 when FC is not integrated are noted. There were no significant differences in the correlation coefficient values for each model. The following were shown from the cut-off value of the correlation coefficient according to Hemphill (2003).

Table 2

Correlation on Model 1: without flexibility factor

	FC1	FC2	FC3	MC	GC
FC1	1				
FC2	.433**	1			
FC3	.375**	.470**	1		
MC	.523**	.446**	.410**	1	
GC	.266**	.247**	.355**	.339**	1

Note. * $p < .05$, ** $p < .01$, FC1 is creativity in Fermi problem 1. FC2 and FC3 as well.

Table 3

Correlation on Model 2: with flexibility factor measured in the more layered category

	FC1	FC2	FC3	MC	GC
FC1	1				
FC2	.433**	1			
FC3	.388**	.490**	1		
MC	.524**	.444**	.426**	1	
GC	.266**	.245**	.367**	.339**	1

Table 4

Correlation on Model 3: with flexibility factor measured in the less layered category

	FC1	FC2	FC3	MC	GC
FC1	1				
FC2	.433**	1			
FC3	.393**	.495**	1		
MC	.524**	.447**	.431**	1	
GC	.266**	.246**	.373**	.340**	1

All were positively correlated. The correlation coefficients between FC1 and MC were strongly correlated. Additionally, FC2 and FC3 had a moderate correlation with MC. In addition, there was also a moderate correlation between GC and FC3. Moreover, FC1 and FC2 are low correlated with GC. Secondly, the correlation coefficients in hypothetical models 4 through 6 when FC is integrated are noted. Then, a strong positive correlation was found between integrated FC and MC, and a moderate positive correlation between integrated FC and GC. There were no significant differences in the correlation coefficient values for each model.

Table 5

Correlation on Model 4: without flexibility factor and with integrated FC

	integrated FC	MC	GC
integrated FC	1		
MC	.711**	1	
GC	.429**	.338**	1

Note. *p < .05, **p < .01, integrated FC is combines the latent variables created for each FC into one.

Table 6

Correlation on Model 5: with flexibility (more layered category) and with integrated FC

	integrated FC	MC	GC
integrated FC	1		
MC	.711**	1	
GC	.427**	.338**	1

Table 7

Correlation on Model 6: with flexibility (less layered category) and with integrated FC

	integrated FC	MC	GC
integrated FC	1		
MC	.713**	1	
GC	.428**	.339**	1

Discussion and Consideration

This analysis allowed us to create a structural model of correlations in FC, GC, and MC. In the models, correlations were found among the three creativities. In particular, in the model that defined latent variables integrating FC1, FC2, and FC3, a strong positive correlation was found between MC and integrated FC, and a moderate positive correlation was found between GC and integrated FC. The results also showed that the flexibility factor had little effect on those correlations. Furthermore, the comparison of the AIC showed better results for the model of FC without the flexibility factor. These results suggest that a correlation between FC and MC and GC can be found without measuring flexibility factor in FC.

Next, subordinate research question 1 is discussed. Firstly, it focuses on the correlations among the latent variables of creativity in each Fermi problem. Looking at the association between FC1, FC2, and FC3, the highest correlation was found between FC2 and FC3. This result seems to be due to the content characteristics of the problems. These two problems require consideration of the broader theme of Japan and the world. In contrast, Fermi problem 1 allows the students to consider more familiar themes than the other two problems. In other words, it is understandable that a relatively strong correlation emerged between FC2 and FC3 when we focus on the familiarity of the content of the problem. The second highest correlation was between FC1 and FC2. It can be assumed that the problem format influences this result; Fermi problem 1 and Fermi problem 2 use the same problem format. These two problem formats were in the form of thinking about the problem-solving ideas, performing calculations, and finally giving a single answer. On the other hand, Fermi problem 3, was in a format that did not require a single answer, only ideas. In addition, FC1 and FC3, which differ in the question format and content type, showed the lowest correlations. From these results, it can be inferred that the content and format of the Fermi problems affect the creativity in the Fermi problems expressed by the students.

Secondly, the relationship between creativity in each Fermi problem and MC and GC is discussed. Those relationships were also shown to be influenced by the format and content of the problems. From the analysis results (see. Table 2 - 4), it is clear that there are differences in the correlations between MC and GC and the latent variables created for each Fermi problem. When focusing on the correlation between GC and each FC, the differences in correlations are shown. For example, the highest correlation with GC is FC3. This high correlation is probably due in large part to the format of the problem. Fermi problem 3 is a form of problem that does not involve computational processing and does not specify a single answer at the end of the problem. This format allows for more freedom in generating ideas than the other two Fermi problems.

Additionally, as mentioned in the hypothesis for the research questions section, GC and Fermi problem 3 do not require much mathematical knowledge and skill. This similarity in the characteristics of the problems suggests that there is a difference in correlation. On the contrary, the format possibly limits the generation of ideas for Fermi problems 1 and 2. For example, it can be inferred that even if an idea comes to them in students' minds, students may be conscious that they are expected to formulate a single answer and may not express it on paper. When attention is also paid to the correlation with MC, the results show the highest correlation with Fermi problem 1. In the hypothesis, it was expected that the correlation between Fermi problem 2 and MC would also be relatively larger than that of Fermi problem 3. However, the results of the analysis showed that there was little difference in the correlation magnitude between Fermi problems 2 and 3. Fermi problem 1 had a familiar problem content, while the other two problems had a less familiar problem content. It seems to be considered that the problem content is more likely to influence the relationship with MC than the problem format. These results indicate that when considering creativity in Fermi problems, the relationship can differ depending on the content and format of the problem.

Then, subordinate research question 2 is discussed. As can be seen from Table 1, all models are acceptable. The correlation coefficients also do not change markedly. Furthermore, Models 1 and 4, the FC models with no flexibility factor, have relatively better than the other models with AIC. Therefore, it is also possible that the flexibility factor does not need to be considered for the creativity of the Fermi problem. Intuitively, it was hypothesized that the model would be better if many factors measured it, but the exciting results belied that hypothesis. In other words, the result supports the view of Lu and Kaiser (2021) in their mathematical modelling of the creativity factor concerning flexibility and fluency. In addition, when comparing models with varying layers of categories to measure flexibility, no considerable differences are found in either model fit or correlation coefficients. This result suggests that the number of categories used to measure flexibility does not seriously impact the relationship between FC, MC, and GC.

Finally, from an educational perspective, the results suggest the following. It is desirable to choose the form and content of the problem according to the creativity, a teacher or observer is interested in focusing. For example, an open-ended format, such as Fermi problem 3, is less restrictive in generating ideas and is more related to GC. Thus, it can be concluded that if a teacher or observer wants to know the general creativity in psychology, one can apply a form of the Fermi problem 3. Therefore, the results of this analysis can serve as one potential guideline for a teacher or observer to determine the creativity they want to know and what type of format and content of the problem they want to use. Furthermore, Fermi problems are also used as official test instruments. For example, Germany's so-called comparative tests (VERA) use Fermi problems. In that test, it is used to assess modeling ability (Greefrath & Frenken, 2021). The results of this study suggest that it may be possible to measure both modeling abilities and creativity simultaneously. This study may also contribute to enhancing the value of such existing tests.

On the other hand, the present study has several limitations. Firstly, the survey was conducted at a single school in Japan, and the age range is relatively narrow, so the scope of the survey is quite limited. Secondly, the model created did not incorporate mathematical ability, which is expected to be related to MC and FC. Thirdly, regarding the investigation of the Fermi problem. It appears that solution time and the order in which the problems are solved may be influential in measuring creativity. This study did not examine the effects of such changes in time or order. Moreover, the validity and reliability of the Fermi problem concerning creativity were not examined. Because of these limitations, careful consideration should be given to the application of this study to other studies.

In the future, a diverse sample will be collected, and created model will be analyzed. Additionally, it also will be considered another model. Furthermore, only the student's solved solution was evaluated in this case. Meanwhile, some studies have focused on the process of solving (Runco, 2007; Schindler & Lilienthal, 2020). Such different approaches can also advance the relationship between creativity in Fermi problems and other creativity. Moreover, a test's validity and reliability to measure creativity using the Fermi problem will be examined.

Appendix

Categories for Fermi problem 1

Categories for Hypothetical Model 2 and 5	Categories for Hypothetical Model 3 and 6
A: Water used for nutrition B: Water used for washing the body C: Water used to wash things D: Water used in an institution E: Age group F: Time, season, and climate G: Other than categories A through F	A: Water used for humans B: Water used for objects C: Other than categories A and B

Categories for Fermi problem 2

Categories for Hypothetical Model 2 and 5	Categories for Hypothetical Model 3 and 6
A: Based on the population B: Based on households C: Based on age group D: Based on country E: Based on location F: Based on the type of phone. G: Other than categories A to F	A: Based on a personal smartphone. B: Based on smartphones owned by other than individuals C: Other than categories A and B

Categories for Fermi problem 3

Categories for Hypothetical Model 2 and 5	Categories for Hypothetical Model 3 and 6
A: Based on the population B: Based on households C: Based on age group D: Based on survey questions E: Based on location F: Based on vehicle type. G: Other than categories A to F	A: Based on personal vehicle B: Based on vehicles owned by other than individuals C: Other than categories A and B

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Ärlebäck, J. B. (2009). On the use of realistic Fermi problems for introducing mathematical modelling in school. *The Mathematics Enthusiast*, *6*(3), 331–364. https://doi.org/10.1007/978-94-007-6271-8_52

Balka, D. S. (1974). Creative ability in mathematics. *The Arithmetic Teacher*, *21*(7), 633–636. https://doi.org/10.5951/AT.21.7.0633

Bennevall, M. (2016). Cultivating creativity in the mathematics classroom using open-ended tasks: A systematic review. Retrieved from https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A909145&dswid=-6476

Blum, W., & Borromeo Ferri, R. (2009). Mathematical modelling: Can it be taught and learnt?. Journal of Mathematical Modelling and Application, 1(1), 45–58. Retrieved from https://proxy.furb.br/ojs/index.php/modelling/article/view/1620/108

Carlson, J. E. (1997). Fermi problems on gasoline consumption. *The Physics Teacher*, *35*(5), 308–309. https://doi.org/10.1119/1.2344696

Genius, M., & Strazzera, E. (2002). A note about model selection and tests for non-nested contingent valuation models. *Economics Letters*, *74*(3), 363-370. https://doi.org/10.1016/S0165-1765(01)00566-3

Goel, P. S., & Singh, N. (1998). Creativity and innovation in durable product development. *Computers & Industrial Engineering*, *35*(1–2), 5–8. https://doi.org/10.1016/S0360-8352(98)00006-0

Greefrath, G., & Frenken, L. (2021). Fermi problems in standardized assessment in grade 8. *Quadrante*, 52–73. https://doi.org/10.48489/QUADRANTE.23587

Guilford, J. P. (1959). Traits of creativity. In H, H Anderson (Ed.), *Creativity and its cultivation*, 142–161. New York: Harper.

Hebert, T. P., Cramond, B., Spiers Neumeister, K. L., Millar, G., & Silvian, A. F. (2002). E. Paul Torrance: His Life, Accomplishments, and Legacy. *National Research Center on the Gifted and Talented*. Retrieved from https://files.eric.ed.gov/fulltext/ED505439.pdf

Helps, C. R., & Lunt, B. (2001). Problem Solving In Engineering: Creativity, Estimation And Critical Thinking Are Essential Skills. *2001 Annual Conference Proceedings*, 6.808.1–6.808.8. https://doi.org/10.18260/1-2--9675

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*(1), 78–79. https://doi.org/10.1037/0003-066X.58.1.78

Hu, L. T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Joreskog, K. G., & Sörbom, D. (1996). LISREL 8: User's Reference Guide (2nd ed.). Chicago: Scientific Software International.

Kattou, M., Kontoyianni, K., Pitta-Pantazi, D., & Christou, C. (2013). Connecting mathematical creativity to mathematical ability. *ZDM*, *45*(2), 167–181. https://doi.org/10.1007/s11858-012-0467-1

Kattou, M., Christou, C., & Pitta-Pantazi, D. (2015). Mathematical creativity or general creativity? In K. Krainer & N. Vondrová (Eds.), Proceedings of the Ninth Conference of the European Society for Research in Mathematics Education (pp. 1016–1023). Charles University in Prague. Retrieved from https://hal.archives-ouvertes.fr/hal-01287305

Kim, K. H. (2006). Can we trust creativity tests? A review of The Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, *18*, 3–14. https://doi.org/10.1207/s15326934crj1801_2

Klavir, R., & Gorodetsky, K. (2011). Features of creativity as expressed in the construction of new analogical problems by intellectually gifted students. Creative Education, 2(03), 164. https://doi.org/10.4236/ce.2011.23023

Lu, X., & Kaiser, G. (2021). Creativity in students' modelling competencies: Conceptualisation and measurement. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-021-10055-y

MacCallum, R.C., Browne, M.W., and Sugawara, H., M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods*, *1* (2), 130-149. http://dx.doi.org/10.1037/1082-989X.1.2.130

Mann, E. L. (2005). *Mathematical Creativity and School Mathematics: Indicators of Mathematical Creativity in Middle School Students*. University of Connecticut.

Mann, E. L. (2006). Creativity: The Essence of Mathematics. *Journal for the Education of the Gifted*, *30*(2), 236–260. https://doi.org/10.4219/jeg-2006-264

Meyer, R. W. (1970). *The Identification and Encouragement of Mathematical Creativity in First Grade Students (Part1)*. Retrieved from https://eric.ed.gov/?id=ED038292

Nadjafikhah, M., Bakhshalizadeh, S., & Yaftian, N. (2011). Fostering mathematical creativity in learning environments. *The 3rd National Conference on Education*, 18–19. Retrieved from https://www.academia.edu/52554325/Fostering_mathematical_creativity_in_learning_environments

Nadjafikhah, M., Yaftian, N., & Bakhshalizadeh, S. (2012). Mathematical creativity: Some definitions and characteristics. *Procedia - Social and Behavioral Sciences*, *31*, 285–291. Retrieved from https://doi.org/10.1016/j.sbspro.2011.12.056

Peter-Koop, A. (2005). Fermi Problems in Primary Mathematics Classrooms: Fostering children's mathematical modelling processes, *Australian Primary Mathematics Classroom*, *10* (1), 4–8. Retrieved from https://search.informit.org/doi/epdf/10.3316/informit.193795255079569

Pitta-Pantazi, D., Sophocleous, P., & Christou, C. (2013). Spatial visualizers, object visualizers and verbalizers: Their mathematical creative abilities. *ZDM*, *45*(2), 199–213. https://doi.org/10.1007/s11858-012-0475-1

Pitta-Pantazi, D., Kattou, M., & Christou, C. (2018). Mathematical creativity: Product, person, process and press. In F. M. Singer (Ed.), *Mathematical creativity and mathematical giftedness: Enhancing creative capacities in mathematically promising students* (pp. 27–53). Springer. https://doi.org/10.1007/978-3-319-73156-8_2

Runco, M. A. (2007). *Creativity: Theories and themes research, development, and practice*. Elsevier Academic Press.

Schindler, M., & Lilienthal, A. J. (2020). Students' creative process in mathematics: insights from eye-tracking-stimulated recall interview on students' work on multiple solution tasks. *International Journal of Science and Mathematics Education*, *18*(8), 1565–1586. https://doi.org/10.1007/s10763-019-10033-0

Schoevers, E. M., Kroesbergen, E. H., & Kattou, M. (2020). Mathematical Creativity: A Combination of Domain□general Creative and Domain□specific Mathematical Skills. *The Journal of Creative Behavior*, *54*(2), 242–252. https://doi.org/10.1002/jocb.361

Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the Model Size Effect on SEM Fit Indices. Educational and Psychological Measurement, 79(2), 310–334. https://doi.org/10.1177/0013164418783530

Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *Zentralblatt Für Didaktik Der Mathematik*, *29*(3), 75–80. https://doi.org/10.1007/s11858-997-0003-x

Sola, E., Hoekstra, R., Fiore, S. and McCauley, P. (2017) An Investigation of the State of Creativity and Critical Thinking in Engineering Undergraduates. *Creative Education*, *8*(9), 1495–1522. https://doi.org/10.4236/ce.2017.89105

Sriraman, B. (2004). The characteristics of mathematical creativity. *The mathematics educator*, *14*(1). Retrieved from https://ojs01.galib.uga.edu/tme/article/view/1868

Tammadge, A. (1979). Creativity: Presidential Address to the Mathematical Association at the Annual Conference, April 1979. *The Mathematical Gazette*, *63*(425), 145–163. https://doi.org/10.2307/3617885

The European Commission. (2019). Key competences for lifelong learning. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/297a33c8-a1f3-11e9-9d01-01aa75ed71a1/language-en

Thohari, C. A. F., Budiyono, & Pratiwi, H. (2020). Students' creativity level on solving mathematics problem. *Journal of Physics: Conference Series*, *1613*, 012051. https://doi.org/10.1088/1742-6596/1613/1/012051

Torrance, E. P. (1963). *Administration And Scoring Manual For Abbreviated Form Ⅶ Minnesota Tests Of Creative Thinking*. Minnesota:University of Minnesota. https://archive.org/details/dli.ernet.18689/page/n1/mode/2up

Torrance, E. (1988). The Nature of Creativity as Manifest in Its Testing. In R. Sternberg (Ed.), The Nature of Creativity (pp. 43-73). New York: Cambridge University Press.

Treffinger, D. J. (2011). *Creativity, Creative Thinking, and Critical Thinking: In search of definitions*. Sarasota: Center for Creative Learning, Inc.

Urban, K. K., & Jellen, H. G. (1996). Test for Creative Thinking – Drawing Production. Swets Test Services

Urban, K. K. & Jellen, H. G. (2010). *Test for Creative Thinking – Drawing Production* (2nd ed.). Pearson Assessment & Information GmbH.

Wakefield, J. F. (1992). Creative thinking: Problem solving skills and the arts orientation. Praeger.

Wessels, H. M.(2014). Levels of mathematical creativity in model-eliciting activities. *Journal of Mathematical Modelling and Aplication*,*1*(9). 22–40.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of Personality and Social Psychology*, *103*(3), 506–519. https://doi.org/10.1037/a0028857

Wilson, R. C., Guilford, J. P., Christensen, P. R., & Lewis, D. J. (1954). A Factor-Analytic Study Of Creative-Thinking Abilities. *Psychometrika*, *19*(4), 297-311. https://doi.org/10.1007/BF02289230

Yano, M., Shibayama, M., Sun, Y., Nishizawa, M., & Fukuda, M. (2002). Concepts and theories of creativity. National Institute of Informatics, Japan. Retrieved from https://www.nii.ac.jp/TechReports/public_html/02-001J.html