Measuring Creativity in the Fermi Problem, a Type of Mathematical Modelling, Applying Information Theory

Hidemichi Okamoto

University of Education Karlsruhe, Germany

hidemichiokamoto@gmail.com

Abstract: *Many methods of measuring creativity have been studied - mainly in psychology. In recent years, there have been attempts to incorporate such creativity into mathematical modelling, a topic handled in mathematics education. Accordingly, some studies have been trying to assess creativity in it. However, there have been no clear criteria or formulas that can be used for any problem, since a rubric has been created for each problem and evaluated individually. In the present study, to measure creativity in the Fermi problem, a type of mathematical modelling, a formula that applies information theory used in information science is proposed and examined using Structural Equation Modeling (SEM). A survey of Japanese junior high school students (n = 364) was conducted and analyzed, and the results show that the model using the formula proposed in the present study is a good fit. In addition, a moderate positive correlation (r = .41, p < .01) is found between creativity in the Fermi problem and creativity in psychology measured by the Test for Creative Thinking-Drawing Production (TCT-DP).*

Keywords: creativity, information theory, fermi problem, mathematical modelling, self-information, test for creative thinking-drawing production, automate an evaluation

1. Introduction

In recent years, much attention is paid to creativity in many fields, including economics, engineering, education, and so on (Piirto, 2011; The European Commission, 2019). Considered one of the necessities in technological advancement and scientific innovation, it has been measured and visualized. For example, many methods of measuring creativity have been studied in the field of psychology (Guilford & Christensen, 1954; Torrance,1962; Urban & Jellen 2010). Additionally, in mathematics education, there have been attempts to examine the relationship between mathematical modelling, that is the process of translation between the real world and mathematics in both directions (Blum & Borromeo Ferri, 2009), and creativity. Furthermore, it has tried to assess creativity in mathematical modelling (Wessels, 2014; Lu & Kaiser, 2021). However, most of those studies create rubrics for each question or category for the ideas respondents answer (Mann, 2005; Wessels, 2014; Lu & Kaiser, 2021). The creation of rubrics and categories often depends on the idea of the problem creator, making it difficult to apply to other problems. If there were evaluation methods and indicators that could be easily adapted to any problem, it would reduce the burden of creating problem-specific evaluation criteria. It is also presumed that it would be easier to categorize the character of the problem and consider the difficulty level of the problem. In addition, if creativity can be measured not only by creativity tests such as those used in psychology, but also by using problems that are easy to handle in school classes, teachers and students may become more aware of creativity.

Hence, this paper proposes a formula that applies the concept of self-information content in information theory to measure the creativity in the Fermi problem, a type of mathematical modelling. The proposed formula is then examined using Structural Equation Modeling to determine if it is appropriate.

2. Review of Literature

In this section, prior research on creativity, mathematical modelling, the Fermi problem, and information theory relevant to the present study is shown.

2.1. Creativity

Previous studies have expressed various definitions of creativity. Treffinger (2011) collected ample literature with definitions of creativity up to 2011. There, he collected more than 100 references, which discuss creativity from different perspectives. As a result, there were diverse ways of perceiving creativity. For example, Guilford and Christensen (1954), pioneers in creativity research, hypothesized that creativity is composed of several factors. It was hypothesized that creativity seems to have the following factors: sensitivity to problem, fluency, flexibility, originality, penetration, analysis, synthesis, and redefinition. Tests were created to measure each of these factors. Similarly, several studies have taken the position that creativity consists of several factors and created tests to measure each of these factors. These tests defined "fluency" as the number of ideas per problem, and "flexibility" as the number of categories of those ideas. They also measured "originality" by giving scores according to the rate of occurrence (Guilford, 1959; Torrance, 1962; Kim, 2006).

On the other hand, Urban and Jellen (2010) viewed creativity as a "whole" rather than on a factor-by-factor basis. In other words, instead of capturing creativity by breaking it down into factors, they viewed it as one creativity.

For this reason, the Test for Creative Thinking-Drawing Production (TCT-DP) was made by them, which is a test for measuring creativity by creating 14 items to give scores. The total score was then used as an index of creativity.

As mentioned above, there are different ways to view creativity in the measurement of creativity and it is widely discussed.

2.2. Fermi Problem

The Fermi problem comes from the physicist Enrico Fermi. He is said to have posed a typical Fermi problem to students at the University of Chicago: "How many piano tuners are there in the city of Chicago?" This is used in schools as a type of mathematical modelling (Peter-Koop, 2005; Ärlebäck, 2009; Greefrath & Frenken, 2021). So far several studies have examined the link between mathematical modelling and creativity (Wessels, 2014; Lu & Kaiser, 2021). Wessels (2014) defines creativity in mathematical modelling in terms of four elements, which are fluency, flexibility, novelty, and usefulness. It was stated that a "framework with four criteria for the identification of creativity was successfully used to evaluate levels of creativity in the solutions to the MEAs (model-eliciting activities)" (Wessels, 2014, pp. 1). Lu and Kaiser (2021) also stated the connection between mathematical modelling competencies and creativity. They defined the three elements of creativity, which are usefulness, fluency, and originality, in the modelling cycle. They suggested that when assessing modelling competencies, it is better to include the perspective of Usefulness. In both studies, rubrics were created and evaluated for each problem. Additionally, Fermi problems are said to require fluency in creativity and to encourage creative thinking (Silver,1997; Goel & Singh, 1998). Moreover, Marcus (2016) categorized open-ended

problems based on a large body of literature. Then, the Fermi problem was classified as a creative thinking problem. The above leads us to assume that creativity in the Fermi problem is worth considering.

2.3. Information Theory

"Information Theory" was developed by Shannon (1948) and Weaver (1949), who views any sending and receiving of voice, images, text, and so on as communication of information. This theory influences many fields, including modern information science. It was considered different types of data and information in terms of a single measurement: the amount of information. It is called self-information and is defined as follows (Shannon, 1948; Weaver, 1949; Jones, 1979): Let S be a system for events E_1, E_2, ..., E_n. in which $P(E_k) = p_k$ with $0 \leq p_k \leq 1$ and

$$p_1 + p_2 + \ldots + p_n = 1$$

The self-information for event E_k is written as $I(E_k)$ and is defined by

$$I(E_k) = - \log_2 p_k \qquad (1)$$

This means that it is defined in such a way that what is less likely to occur probabilistically has higher information content. The probability of an event occurring 100% of the time is $p_k = 1$. Events that are known to occur with certainty can be considered self-evident phenomena and are viewed as having no value because there is no new information that can be learned. Since the information has no meaning, self-information is zero.

Snyder et al. (2004) discussed the connection between creativity based on Guilford's (1959) or Getzels and Jackson's (1962) creativity theory and information theory. Snyder et al. (2004) attempted to

develop a formula for creativity scores based on information theory and conducted a survey (N = 25) to examine the formula. The formula was expressed as follows:

$$\log_2 \{(1 + u_1)(1 + u_2) \ldots (1 + u_n)\} \quad (2)$$

u_n is the total number of ideas in a category. For example, in the creativity test used in the study to think about the uses of paper, the u_1 category is "Surface Marking" and the u_2 category is "Toy/Game." One subject comes up with the ideas "writing", "painting", and "airplane." Then, "writing" and "painting" are in the u_1 category and "airplane" is in the u_2 category. Therefore, the subject's answer is evaluated as $u_1 = 2$ and $u_2 = 1$.

In this study, the formula to measure creativity was defined so that the more diverse the categories of ideas, the higher the value of creativity. Moreover, this creativity measurement requires the creation of a category for each problem. In addition, it is difficult to guarantee the validity of the categorization criteria. It is also difficult to assess differences in value among categories. For example, the creativity value of a respondent who came up with categories A and B, which are easy to conceive of, and the creativity value of a respondent who came up with categories C and D, which are difficult to think of, are both the same.

3. Research Question

Can creativity in the Fermi problem be expressed in a formula using information theory, without recourse to more or less arbitrary category systems?

4. Method

The purpose of the present study is to propose and examine a formula for measuring creativity in the Fermi problem with the application of information theory. Thus, a survey of creativity in psychology and the Fermi problem was conducted among junior high school students in Japan. The results of the survey were computed using the proposed formula and existing formula (2). The results of the calculations were analyzed and examined for indexes of model fit of the hypothetical models using lavaan, a package for structural equation modeling of the software R.

4.1. Participants

A total of 364 students from a public junior high school participated in the survey. Their ages ranged from 12 to 15 years old and included 195 boys and 169 girls. The academic performance of the students was about the same as the national average for annual academic achievement surveys conducted throughout Japan. In addition, 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 sample size of at least 100, with an anticipated effect size of 0.3, desired power of 0.8, 1 latent variable and 4 observed variables, and a probability level of 0.05 (Sloper, 2015).

4.2. Procedures

The survey was conducted to measure creativity in the Fermi problem and creativity dealt within psychology. Three Fermi problems and the TCT-DP are given to the participants. All survey questions are conducted on the same day, for a total of approximately 40 minutes. The order in which the surveys are conducted is as follows: first, the three Fermi problems are conducted. After that, the TCT-DP is performed.

4.3. Test for the Survey

Firstly, the Fermi problem is described. The three Fermi problems conducted are as follows;

Problem 1: How many liters of water does one person use in a year?
Problem 2: If you collected all the smartphones in the world, how many would there be?
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 present study views creativity in the Fermi problem as the richness of aspects of solving a Fermi problem. For example, in Problem 2, ideas such as "the number of smartphones a family can have", "age groups that own smartphones", and "wealth differences by country" are given. Based on these ideas, the students are asked to create a mathematical model of how many smartphones there are in the world, calculate it, and come up with a single answer. Problem 1 is similarly a problem of generating ideas, creating a model, performing calculations, and coming up with a single answer. Problem 3, on the other hand, provides an idea of how to solve the problem. This problem does not require calculations, but rather the generation of a solution. Furthermore, categories of ideas are created for each problem (see. Appendix).

Secondly, the TCT-DP is described. In this test, a subject draws additional pictures on an unfinished drawing and gives it a title. The picture is then rated on 14 of these items (Urban & Jellen, 2010). The sum of the scores for all of these items is used as the creativity score. TCT-DP has good inter-rater reliability: a = .81 – .99 for the total score and a ≥ .89 for test criteria (Urban & Jellen, 1996). Desmet et al. (2021) studied the validity and usefulness of the TCT-DP. This study provided evidence for the utility and divergent validity of the TCT-DP when used with a Dutch population (Desmet et al., 2021). This test is scored according to the manual (Urban & Jellen, 2010).

4.4. Proposal for Mathematical Formula

Formula (2) is computed by categories. When evaluating the Fermi problem, a type of mathematical modelling, with the formula (2), is necessary to create a category for each problem. In addition, it is difficult to evaluate the value of each category. Therefore, the following formula is proposed.

The incidence of ideas for solving the Fermi problem is defined as $P(x_1)$, $P(x_2)$,…, $P(x_n)$.

$$\log_2 1/P(x_1) + \log_2 1/P(x_2) +…+ \log_2 1/P(x_n) \quad (3)$$

This formula is weighted by an idea. It is also closer to the definition of the amount of self-information in information theory than formula (2). An example is considered with Fermi problem 1: How many liters of water does one person use in a year? conducted in this survey. A subject considers the water one uses to solve this problem. The subject thinks of it as "drinking water", "bathing", and "laundry." The incidence of "drinking water" is then 90%, the incidence of "bath" is 80%, and the incidence of "laundry" is 10%. The result for the subject is calculated as follows. x_1 is "drinking water." x_2 is "bathing." x_3 is "laundry."

$$\log_2 1/0.9 + \log_2 1/0.8 + \log_2 1/0.1$$
$$= 0.15… + 0.32… + 3.32…$$
$$\approx 3.79$$

Thus, the subject's creativity is rated at about 3.79. As can be observed from this calculation, the lower the rate of occurrence of an idea, the higher the value.

4.5. Hypothetical model for SEM

Three hypothetical models are presented for consideration in SEM.

Model 1 is a model that calculates the creativity in the Fermi problem using the formula (2). Additionally, the ideas in the Fermi problem are classified into seven categories.

Model 2 is also a model that calculates the creativity in the Fermi problem using the formula (2). Additionally, the ideas in the Fermi problem are classified into three categories.

Model 3 is a model that calculates the creativity in the Fermi problem using the formula (3) proposed in the present study.
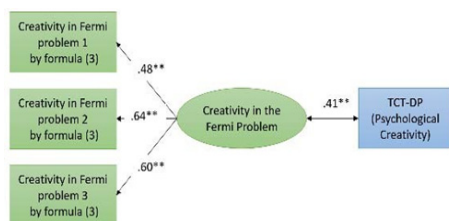
5. Result

Table 1

Fitness Index in SEM for models and Correlation Coefficients

	Model 1	Model 2	Model 3
p	.00	.02	.05
χ^2	.04	8.05	5.86
df	2	2	2
CFI	.93	.93	.96
RMSEA	.12	.11	.09
SRMR	.04	.04	.03
AIC	3808.85	6814.49	4077.26
BIC	3836.60	6842.24	4105.01
r	.36**	.42**	.41**

*$p < .05$;.**$p < .01$

Figure 1.
Hypothetical Model 3

Results of the SEM model fit index show that the best model is Model 3 (see. Table 1). Firstly, the Chi-Square Test shows a significant difference in Models 1 and 2, which was considered unsatisfactory for the model (Joreskog & Surbom, 1996). Secondly, the model is considered good if the CFI is greater than 0.95 (West et al. 2012). In other words, only model 3 was concluded to be a good model. Thirdly, REMSA is ideally less than 0.05. Values above 0.1 are considered poor, values between 0.08 and 0.1 are considered borderline, values in the range of 0.05 to 0.08 are considered acceptable, and values below 0.05 are considered good (MacCallum et al, 1996). In this case, Models 1 and 2 were considered poor, while Model 3 was in the range considered borderline. This result seems to be possibly due to the low degrees of

freedom of these models. Fourth, the SRMR is also considered excellent for values lower than 0.05 (Diamantopoulos & Siguaw, 2000). In this index, all models showed good results.

Therefore, good results were shown for multiple indexes of model fit in Model 3 using the formulas proposed in the present study. In addition, a moderate correlation (r = .41, p < .01) was found between the latent variable considered as the creativity of the Fermi problem, consisting of the three Fermi problem creativity observables created in Model 3, and the TCT-DP values (Hemphill, 2003).

6. Discussion and Conclusion

The results of the hypothetical model comparison by SEM showed that the formula proposed in the present study, which is weighted by the incidence of each idea, is better than the previous study's formula, which views ideas by category (Snyder et al., 2004). A major factor in these results seems to be the difficulty of creating categories and classifying ideas. For example, in the Fermi problem 1 category, the categories "Water used for washing the body" and "Water used to wash things" are created. However, some evaluators can assume that these could be combined into a single category of "things used to wash" Thus, how the categories are created depends largely on the evaluator's perspective. Therefore, the validity of the category is difficult to guarantee. In addition, it is examined whether the model fit is affected by varying the number of categories by the hypothetical models 1 and 2. The results of the analysis show no large differences in model fit indexes. Hence, it is not necessary to create categories of ideas to assess the creativity in the Fermi problem. From a different perspective, ideas with low incidence can also be viewed as being far outside of existing categories. In other words, giving high scores to ideas with low incidence, ideas that are not the way most people come up with, is consequently giving high scores to ideas in a different category.

A moderate positive correlation was found between creativity in the Fermi problem and creativity by TCT-DP. Based on the result, it is possible to say that creativity in psychology can be expressed by a mathematical formula applying the information theory proposed in the present study as the answer to the research problem. On the other hand, the evaluation of creativity in the Fermi problem is highly dependent on the number of ideas for solving the Fermi problem. There is no evaluation given to the correctness or incorrectness of the computational process performed by the subjects, or to the mathematical models created to solve the problem. In other words, the part of a convergence of ideas, after thinking divergence, is not evaluated. Creativity has a divergent and a convergent part (Runco, 2007), and this evaluation method focuses on the divergence phase of creativity. If the convergent part is appropriately incorporated into the proposed formula, it is expected to show an even higher correlation.

There are several limitations to the present study. Firstly, the scope of the survey is rather narrow, as it is conducted in a single school in Japan with a considerably small age range. Secondly, the formula proposed in the present study is weighted according to the rate of occurrence of ideas, so the creativity values are likely to be affected by the characteristics of the sample. For example, ideas can be biased by country. In Japan, rice is the staple food of the culture. Therefore, when considering Fermi problem 1, not a few subjects considered "water for cooking rice." On the other hand, in cultures where rice is not a staple food, such as in Europe and the United States, this idea is expected to be less likely to

be generated. Therefore, the same idea would have a different rate of occurrence in different groups, and the value of creativity calculated from the formula is changed. Thirdly, it is about the investigation of the Fermi problem. In measuring creativity, solution time and the order in which problems are solved seems to be a potential influence. It is not considered made of the impact of such changes in time or sequence in the present study.

7. Outlook for the Future

The prospects for reducing the limitations of the present study are described in this section. One way to overcome the limitations is to automate the evaluation of the subjects' answers and collect larger data sets. This survey was conducted on a paper basis. If a program is created to allow subjects to answer the survey on a computer, the data analysis can be facilitated. In addition, if a large amount of data can be collected with a diverse sample, the probabilities for weighting the ideas that emerge can be made more accurate. Thus, it is expected that the accuracy of the proposed formula can be increased as well. Moreover, by applying the research of Pla-Castells and García-Fernández (2020) or Okamoto (2021) making the whole thinking process visible in auto, it is possible to evaluate not only the idea but also the whole thinking process. It is assumed that this would make it possible to incorporate variables such as "when the idea was generated", "what computational processes were performed", and "how many times the model was modified" into the formulas proposed in the present study. In the future, methods that can measure creativity in the Fermi problem with high accuracy are to be explored.

Reference

Ärlebäck, J. B. (2009). On the use of realistic Fermi problems for introducing mathematical modelling in school. *The Mathematics Enthusiast, 6*(3), 331–364.

Blum, W., & Borromeo Ferri, R. (2009). Mathematical modelling: Can it be taught and learnt?. *Journal of Mathematical Modelling and Application, 1*(1), 45–58.

Desmet, O. A., van Weerdenburg, M., Poelman, M., Hoogeveen, L., & Yang, Y. (2021). Validity and utility of the test of creative thinking drawing production for Dutch adolescents. *Journal of Advanced Academics, 32*(3), 267–290. https://doi.org/10.1177/1932202X21990099

Diamantopoulos, A. & Siguaw, J. A., (2000). *Introduction to LISREL: A guide for the uninitiated*. London: SAGE Publications, Inc.

Getzels, J. W. & Jackson, P. W. (1962) *Creativity and intelligence*. New York: John Wiley and Sons, Inc

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, 30*(1), 52–73.

Guilford, J. P., & Christensen, R. P. (1954). A Factor-analytic study of creative-thinking abilities. *Psychometrika, 19*(4), 297–311.

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

Hemphill, J. F. (2003). *Interpreting the magnitudes of correlation coefficients. American Psychologist*, 58(1), 78–79.

Jones, D. S. (1979). *Elementary information theory.* Oxford University Press, USA.

Joreskog, K. G., & Sorbom, D. (1996).

LISREL8 User's reference guide. Mooresville Scientific Software.

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.

Lu, X., & Kaiser, G. (2021). Creativity in students' modelling competencies: conceptualisation and measurement. *Educational Studies in Mathematics*, 1–25.

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.

Mann, E. L. (2005). *Mathematical creativity and school mathematics: Indicators of mathematical creativity in middle school students*. University of Connecticut.

Marcus, B. (2016). Cultivating creativity in the mathematics classroom using open-ended tasks: a systematic review. Retrieved from http://liu.diva-portal.org/smash/get/diva2:909145/FULLTEXT01.pdf

Okamoto, H. (2021). Development of a method for observing the detailed modelling activity and creative factors diagram in a timeline using Fermi problem. In *Proceedings of IJCIME 2021: International Joint Conference on Information, Media and Engineering.* IEEE eXpress Conference Publishing. https://ijcime.com/2021/

Peter-Koop, A. (2005). Fermi problems in primary mathematics classrooms: fostering children's mathematical modelling processes, *Australian Primary Mathematics Classroom, 10*(1), 4–8.

Piirto, J. (2011). *Creativity for 21st century skills.* In Creativity for 21st Century Skills (pp. 1-12). SensePublishers.

Pla-Castells, M., & García-Fernández, I. (2020). TaskTimeTracker: A tool for temporal analysis of the problem solving process. *Investigación en Entornos Tecnológicos en Educación Matemática,* (1).

Runco, M. A. (2007). *Creativity theories and themes: Research, development, and practice.* Burlington, MA: Elsevier Academic Press.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal, 27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *ZDM, the International Journal on Mathematics Education, 29*(3), 75–80.

Sloper, D.S. (2015). A-priori sample size calculator for structural equation models [Software]. Retrieved from http://www.daniel soper.com/statcalc

Snyder, A., Mitchell, J., Bossomaier, T., & Pallier, G. (2004). The creativity quotient: an objective scoring of ideational fluency. *Creativity Research Journal, 16*(4), 415–419. https://doi.org/10.1080/10400410409534552

The European Commission (2019). Key competences for lifelong learning. https://doi.org/10.2766/569540

Torrance, E. P. (1962). *Administration and scoring manual for abbreviated form VII Minnesota Tests Of Creative Thinking.* Minnesota: University of Minnesota.

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 (TCT-DP).* Lisse, The Netherlands: Swets & Zeitlinger.

Urban, K. K. & Jellen, H. G. (2010). *Test for creative thinking – drawing production*

(2nd ed.). Pearson Assessment & Information GmbH.

Weaver, W. (1949). Recent contributions to the mathematical theory of communication. Retrieved from http://waste.informatik.hu-berlin.de/Lehre/ss11/SE_Kybernetik/reader/weaver.pdf

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

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

Acknowledgment

Appendix

1. Fermi problem 1

Categories for Hypothetical Model 1

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

Categories for Hypothetical Model 2

A: Water used for humans

B: Water used for objects

C: Other than categories A and B

2. Fermi problem 2

Categories for Hypothetical Model 1

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

Categories for Hypothetical Model 2

A: Based on a personal smartphone.

B: Based on smartphones owned by other than individuals

C: Other than categories A and B

3. Fermi problem 3

Categories for Hypothetical Model 1

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

Categories for Hypothetical Model 2

A: Based on personal vehicle

B: Based on vehicles owned by other than individuals

C: Other than categories A and B