Linear: Journal of Mathematics Education
Volume 6, No. 2, December 2025
p-ISSN: 2722-6913 e-ISSN: 2722-760X
https://doi.org/10.32332/f9cy8d51

Development of PISA-type questions on change and relationship content from level 1- 6

Kadek Restu Astuti[1)], Khaerani Dwi Saputri[2)], Rendi Kurniawan[3)], Ratu Ilma Indra Putri[4)], Elsa Susanti[5)*], Ruth Helen Simarmata[6]
[1) 2) 3) 4) 5) 6) 7) 8)] Program Studi Pendidikan Matematika, Universitas Sriwijaya, Indralaya, Indonesia
*elsasusanti@fkip.unsri.ac.id

Received: 9 June 2025| Revised: 2 September 2025| Accepted: 8 September 2025| Published Online: 9 October 2025

Abstract

This study aims to develop PISA model math questions on Change and Relationship content that are valid, practical, and potentially effective to improve students' reasoning and mathematical literacy skills. The method used is design research type development studies, which consists of preliminary and formative evaluation stages. The instrument was developed through a series of evaluation stages, namely self-evaluation, expert review, one-to-one, and small group. The research subjects were VIII grade students of SMP Negeri 1 Indralaya. The questions developed included six items with difficulty levels from level 1 to level 6 and varied contexts (personal, general, and scientific). Based on the analysis of student answers, it is obtained that the questions developed by researchers can be said to be valid and practical, this is supported by the results of the validity, reliability, difficulty level, differentiating power tests quantitatively and declared practical based on the results of the analysis of student answers which show students have been able to understand the questions, the sentences used do not have double meanings and do not cause errors caused by the questions. Thus, the questions developed are declared feasible to be used to measure and train mathematical literacy and reasoning of junior high school students.

Keywords: PISA Development; PISA Question; Change and Relationship

Published by Linear: Journal of Mathematics Education

INTRODUCTION

Programme for International Student Assessment (PISA) is an international assessment that aims to measure students' ability to apply their knowledge and skills in real-life contexts, particularly in the areas of reading, mathematics, and science (Mayari, 2020). In mathematics, PISA emphasizes critical thinking, reasoning, and contextual problem-solving over mere mastery of formulas or algorithms (OECD, 2019).

The 2022 PISA results show that although Indonesia's ranking has improved by 5 to 6 positions compared to 2018, the average mathematics literacy score has actually decreased from 379 to 366 (A. Susanta et al., 2023). In addition, almost no Indonesian students reached level 5 or 6 in the 2022 PISA mathematics test, while the OECD average showed that around

9% of students reached that level (OECD, 2022).

Low student performance may be caused by limited exposure to PISA-style contextual mathematics questions (Rawani, 2019). Developing PISA-based questions is a strategic way to measure mathematical literacy while training critical thinking and problem-solving skills (A. Safitri et al., 2023). The Change and Relationship domain emphasizes variable relationships, changes over time, and representations such as graphs and tables (Azzahra, 2021). Functional thinking is essential to understand phenomena like population growth, speed, and data trends. and mastering it improves students' ability to use graphs, tables, and algebraic forms in context-based tasks (Fadlila et al., 2023).

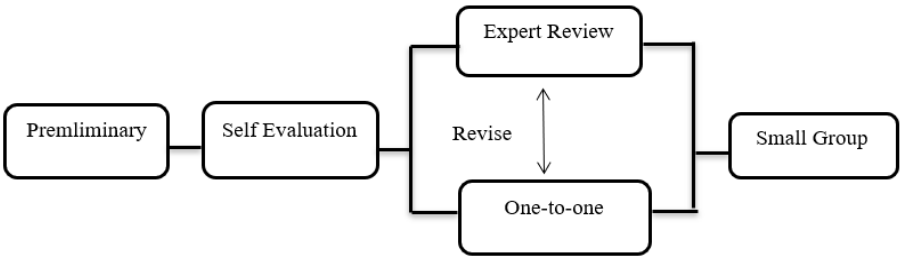
However, many Indonesian students still struggle with Change and Relationship problems requiring higher-order thinking (Murtiyasa, 2023). Many Indonesian students struggle with PISA Change and Relationship tasks, particularly those requiring higher cognitive skills. (Nusantara et al., 2021) It highlights the importance of designing questions that help students shift from procedural to conceptual thinking. Previous studies have developed PISA-based mathematics questions, including ones tailored to students' local and personal contexts. (Hardianti, 2019). Research by (Zulkardi, 2021) Developing PISA-like contextual questions on Change and Relationships through the RME approach can enhance students' higher-order thinking skills. (Lusinda et al., 2025) developing low-level PISA questions (levels 1–2) that use everyday contexts such as markets and transportation. (Mitari et al., 2019) focusing on developing PISA questions in the context of Palembang's local culture, with valid and practical questions for junior high school students. Meanwhile, (Azzahra, 2021) developing challenging high-level questions (levels 4–6) that support students' mathematical reasoning skills.

Research has developed valid and practical PISA-like questions on Change and Relationship to support mathematical literacy (Fadlila et al., 2023) Thus, it is important to design PISA-based questions with broader everyday contexts (Efriani et al., 2021). aiming to foster meaningful learning and align assessments with international standards (Davita et al., 2020).

METHODS

This study uses a design research method (development studies) to address complex learning problems and produce practical and theoretical solutions in education. (Riyani et al., 2017). The research consists of a preliminary stage (needs analysis, theory review, prototype

design) and a formative evaluation stage. (Sri Mertasari et al., 2022). The formative evaluation stage refines the product through self-evaluation, expert review, one-to-one, and small group testing.



The subjects were eighth graders at SMP Negeri 1 Indralaya. Three students of varied abilities participated in the one-to-one phase to evaluate clarity and understanding, while six students joined the small-group phase to assess practicality. Subjects were purposively selected to match each stage of development. (Putri et al., 2024). his research was conducted only up to the small group stage, without progressing to the field test stage.

In the small group stage, data were analyzed through item validity testing, reliability testing, and difficulty level testing (Hadari et al., 2023). The correlation coefficient was used to determine instrument validity, with interpretation based on Table 1 and a minimum validity threshold set for each item $0{,}70 \leq r_{xy} < 90$.

Table 1. Criteria for Correlation Coefficient Validity

Koefisien Korelasi	Korelasi	Interpretasi
$0{,}90 \leq r_{xy} < 1.00$	Very High	Excellent
$0{,}70 \leq r_{xy} < 90$	Height	Good
$0{,}40 \leq r_{xy} < 70$	Moderate	Enough
$0{,}20 \leq r_{xy} < 40$	Low	Bad
$0{,}00 \leq r_{xy} < 20$	Very Low	Very Bad

Furthermore, the reliability analysis of an instrument is determined by the correlation coefficient value between the statement items and the instrument (Fitriani, 2021).Reject measurements to interpret the degree of reliability based on Table 2. The minimum criteria for the reliability correlation coefficient of test items are $0{,}70 \leq r_{xy} < 90$.

Table 2. Reliability Coefficient

Koefisien Korelasi	Korelasi	Interpretasi
$0{,}90 \leq r_{xy} < 1.00$	Very High	Excellent
$0{,}70 \leq r_{xy} < 90$	Height	Good
$0{,}40 \leq r_{xy} < 70$	Moderate	Enough
$0{,}20 \leq r_{xy} < 40$	Low	Bad
$0{,}00 \leq r_{xy} < 20$	Very Low	Very Bad

A question is said to have a good difficulty index when it is neither too easy nor too

difficult (Hadari et al., 2023). Table 3 shows the criteria for the difficulty index of questions, which must be between $0{,}30 \leq IK \leq 0{,}70$.

Table 3. Difficulty Index Criteria

Difficulty Index (DI)	Interpretation of Difficulty Index
$0{,}00 \leq IK \leq 0{,}30$	Difficult
$0{,}30 < IK \leq 0{,}70$	Moderate
$0{,}70 < IK \leq 1{,}00$	Easy

Discriminating power was calculated by comparing the top 27% and bottom 27% of students, indicating how well items distinguish mastery levels. Each item's discrimination index was then determined using established criteria. (Fitriani, 2021) .

Table 4. Discrimination Index Criteria

Difficulty Index (DI)	Interpretation of the Difficulty Index
$0{,}00 - 0{,}19$	Bad
$0{,}20 - 0{,}39$	Enough
$0{,}40 - 0{,}69$	Good
$0{,}70 - 1{,}00$	Excellent

These criteria show how well items distinguish between high- and low-ability students. Items with high discrimination are effective, while those with low or negative values require revision as they may be invalid or misleading.

RESULTS AND DISCUSSION

Researchers developed mathematics questions similar to those used in PISA, with content on Change and Relationships ranging from level 1 to level 6. These questions consisted of several contexts, with three questions in a personal context, two questions in a general context, and one question in a scientific context, to improve students' mathematical comprehension skills in solving PISA questions.

1.1 Preliminary

At this stage, the researcher reviewed literature on PISA question development in Change and Relationships, identified study subjects, and designed prototype questions. A question matrix, question cards, and a scoring rubric aligned with the PISA framework were also developed. (Purwanto et al., 2024).

1.2 Formative Evaluation

This stage consists of two phases, self-evaluation and prototyping.

a. Self Evaluation

In the self-evaluation stage, the researcher analyzed the Change and Relationships content, student characteristics, and material focus before designing PISA-based problems emphasizing content, construct, and language.

The questions were adapted from previous PISA items, paraphrased, and contextualized for Indonesia across six levels: flower sales (level 1), car sales (2), speed (3), cost savings (4), pyramid patterns (5), and climbers (6). Solving them requires students' communication, representation, and reasoning skills. (Machmuda et al., 2024).

b. Prototyping

In the prototyping stage, there are three stages that are passed through in the validation process, namely the expert review stage, one-to-one, and small group.

Expert Review

Expert review was conducted by Mathematics Education peers and a lecturer at Sriwijaya University, with comments and suggestions summarized in Table 5.

Table 5. Comments/Suggestions from Peers at the Expert Review Stage

Question Level	Comments/Suggestions
Level 1	Clarify the meaning of the question
Level 2	The information provided is unclear or insufficiently specific.
Level 3	The topics used should be those that exist in the surrounding environment.
Level 4	Use Indonesian currency
Level 5	The difficulty level of the questions needs to be increased.
Level 6	No revisions are necessary.

After peer validation, lecturers in Mathematics Education at Sriwijaya University also validated the questions, with their suggestions shown in Table 6.

Table 6. Comments/Suggestions from Expert Review

Question Level	Comments/Suggestions
Level 1	It is already at the appropriate level, no revision is necessary.
Level 2	Increase the difficulty of the questions, try to include simple arithmetic operations.
Level 3	Increase the difficulty of the questions and use more practical words.
Level 4	Pay attention to EYD and writing structure.
Level 5	Note that EYD and the question in the problem sufficiently prove the statement that the percentage of blue cubes is always more than 50%.
Level 6	Provide information on the number of days in a year to clarify the question and prepare additional answer options.

One To One

The one-to-one stage tested three students of varying abilities, focusing on question quality—content, structure, and language—rather than answers, with guidance given as needed.(Oktaviana et al., 2020). After conducting the test questions accompanied by guidance, the researchers also interviewed the students and provided suggestions based on the results of the interviews.

P: Did you have any trouble with the questions from these six levels?

S: Yeah, I had trouble because I wasn't sure what steps to take to solve them, but after getting some guidance, I started to understand what to do.

P: Do you have any suggestions that might make the process easier?

S: Yes, I think the sentences used are a bit too long, and to make it easier to work on, perhaps some additional information could be provided to help guide students through the steps of the solution.

P: For the problems in levels 1 to 4 did you encounter any difficulties with the structure or language used?

S: I only had difficulty determining the correct process for solving them.

The first prototype was revised into a second prototype based on expert review feedback, student input, and one-to-one observations. Figures 2–5 illustrate revisions made to questions at levels 2, 3, 5, and 6.

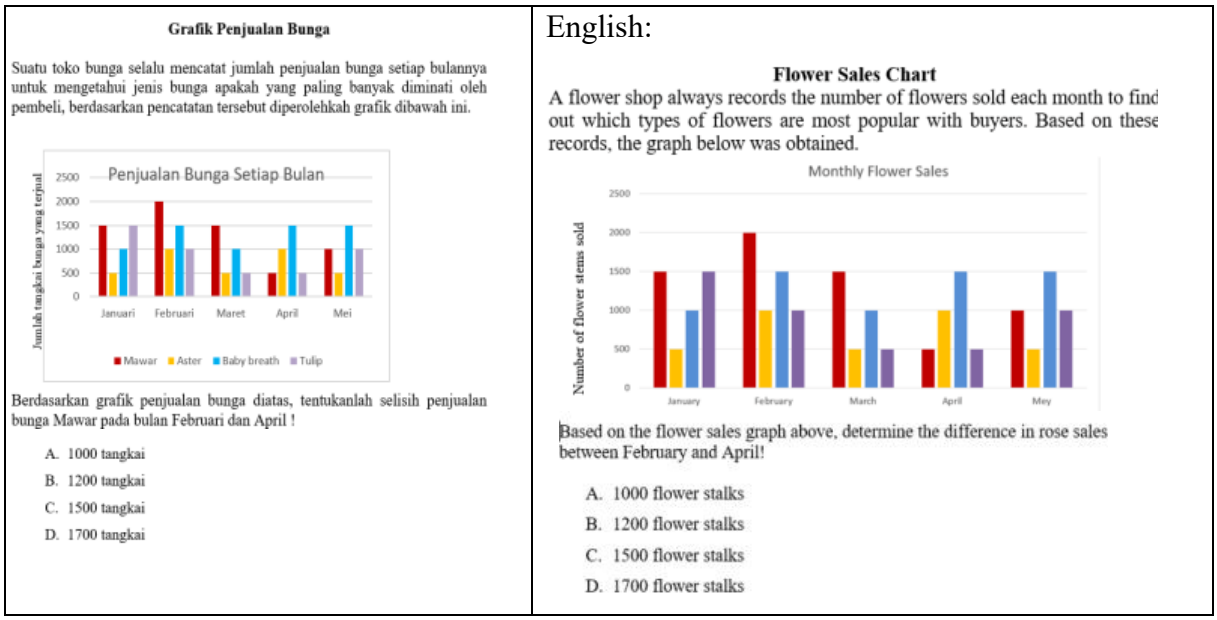


Figure 2. Changes in Questions at Level

Mobil

Terdapat tiga mobil yang memiliki kecepatan yang berbeda sedang melakukan perjalanan panjang. Mobil A, Mobil B, dan Mobil C masing-masing memulai perjalanan dari titik yang sama. Setiap mobil melaju dengan kecepatan tetap sepanjang perjalanan. Tabel berikut menunjukkan jarak yang ditempuh oleh mobil pada setiap waktu tempuhnya.

Nama Kendaraan	Waktu Tempuh				
	2 jam	4 jam	6 jam	8 jam	10 jam
Mobil A	120	240	360	...	...
Mobil B	130	260	390	...	...
Mobil C	140	280	420	...	...

Tentukanlah jarak yang dapat ditempuh masing-masing mobil dalam waktu 8 jam dan 10 jam !

English:

Car

There are three cars with different speeds making a long journey. Car A, Car B, and Car C each start their journey from the same point. Each car moves at a constant speed throughout the trip. The table below shows the distance traveled by each car for different travel times.

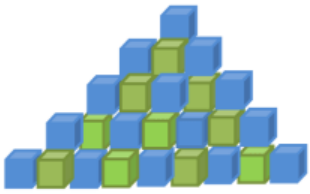
Vehicle Name	Travel Time				
	2 hours	4 hours	6 hours	8 hours	10 hours
Car A	120	240	360	...	...
Car B	130	260	390	...	...
Car C	140	140	420	...	...

Determine the distance each car can travel in 8 hours and 10 hours!

Figure 3. Changes in Questions at Level 3

Piramida Kubus

Dion memiliki beberapa kubus, dimana ia menyusun kubus-kubus tersebut menjadi piramida kubus seperti dibawah ini!

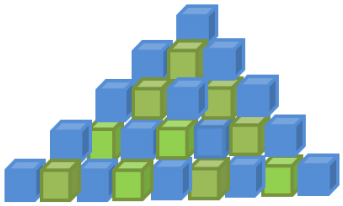


Piramida kubus tersebut akan ditambahkan tingkatanya kearah bawah hingga sepuluh tingkat. Pada tingkat pertama persentase banyak kubus biru adalah 100% sedangkan persentase banyak kubus hijau adalah 0%. Berdasarkan hal tersebut Dion berpendapat bahwa persentase banyak kubus biru akan selalu lebih dari 50% pada setiap tingkat. Apakah pendapat Dion benar? Jelaskan jawabanmu!

English:

Pyramid Cube

Dion has several cubes, which he arranges into a cube pyramid like the one below!



The cube pyramid will be added to ten levels from the bottom. At the first level, the percentage of blue cubes is 100%, while the percentage of green cubes is 0%. Based on this, Dion argues that the percentage of blue cubes will always be more than 50% at each level. Is Dion's opinion correct? Explain your answer!

Figure 4. Changes in Questions at Level 5

Mendaki Gunung Rinjani

Sumber: https://www.trekkingrinjani.com/rinjanitrekking/gunung-rinjani.html

Taman Nasional Gunung Rinjani terletak di Pulau Lombok, Provinsi Nusa Tenggara Barat. Keindahan dan tantangan yang ditawarkan menjadikan Gunung Rinjani sebagai destinasi favorit bagi para pendaki. Setelah terdampak gempa Lombok tahun 2018 dan pandemi Covid-19, kegiatan pendakian di Gunung Rinjani berangsur pulih bahkan terjadi peningkatan terhadap jumlah kunjungan. Berdasarkan data dari Balai Taman Nasional Gunung Rinjani, jumlah pendaki pada tahun 2021 tercatat sebanyak 38.785 orang, sedangkan pada tahun 2022 meningkat menjadi 42.658 orang. Jika pada tahun selanjutnya jumlah pendaki per-tahunnya selalu meningkat dengan persentase yang sama seperti persentase kenaikan tahun 2021 ke 2022, maka pada tahun berapa rata-rata jumlah pendaki per-hari mencapai sekitar 188 orang? (Diasumsikan dalam satu tahun terdapat 365 hari).

Climbing Mount Rinjani

source: https://www.trekkingrinjani.com/rinjanitrekking/gunung-rinjani.html

Rinjani National Park is located on the island of Lombok, West Nusa Tenggara Province. Its beauty and challenges make Mount Rinjani a favorite destination for climbers. After being affected by the 2018 Lombok earthquake and the Covid-19 pandemic, climbing activities on Mount Rinjani gradually recovered and even saw an increase in the number of visitors. According to data from the Gunung Rinjani National Park Office, the number of hikers in 2021 was recorded at 38,785 people, while in 2022 it increased to 42,658 people. If the number of hikers per year continues to increase at the same percentage as the increase from 2021 to 2022, in which year will the average number of hikers per day reach approximately 188 people? (Assuming there are 365 days in a year).

Figure 5. Changes in Questions at Level 6

Small Group

The small group stage, conducted on May 8, 2025 with six grade VIII students (two high, two medium, two low ability), assessed the practicality and readability of the

questions, as well as provided initial statistical analysis of item validity, reliability, difficulty, and discrimination power. (Anshari et al., 2023). This study is limited only to the small group, so the results of statistical analysis are considered as preliminary findings that provide an overview of the quality of the instrument.

The validity test was carried out using the Pearson Product Moment correlation, between the score of each question item and the total score of the students (Ketaren et al., 2024). The results of calculating the validity of each question level are presented in the following

Table 7. Criteria for Correlation Coefficient Validity

Question Level	Correlation Value (r)	Interpretasi
Level 1	0,594	Enough
Level 2	0,594	Enough
Level 3	0,751	Good
Level 4	0,961	Excellent
Level 5	0,973	Excellent
Level 6	0,976	Excellent

Results showed that Level 1–2 questions were valid with minor revisions, while Level 3–6 questions had good to excellent validity and were suitable for use. A reliability test was then conducted to measure the consistency of the PISA-type questions. (Indrawati et al., 2022). In the reliability test process, the researcher used the Cronbach's Alpha and obtained a value of 0.79, which shows that the instrument is classified as reliable because it is in the high category ($0,7 \leq r < 0,9$).The results of the difficulty calculation are presented in the following table:

Table 8. Difficulty Index Criteria

Question Level	Difficulty Level	Interpretasi
Level 1	0,83	Easy
Level 2	0,83	Easy
Level 3	0,86	Easy
Level 4	0,66	Moderate
Level 5	0,50	Moderate
Level 6	0,40	Difficult

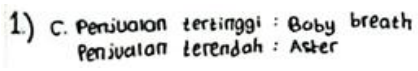
The analysis showed that the instrument covers various difficulty levels and can accommodate differences in students' abilities.

Table 9. Differentiating Power Index Criteria

Question Level	Differentiating Power	Interpretasi
Level 1	0.33	Enough
Level 2	0.33	Enough
Level 3	0.13	Bad
Level 4	0.4	Good

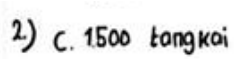
Level 5	0.46	Good
Level 6	0.26	Enough

The analysis showed that Levels 1, 2, and 6 required minor revisions, Level 3 had poor discrimination and needed revision or removal, while Levels 4 and 5 had good discrimination and were suitable for use.



Translate: C. Highest sales: Baby's breath
Lowest sales: Aster

Figure 6. Student Answers to Level 1 Questions



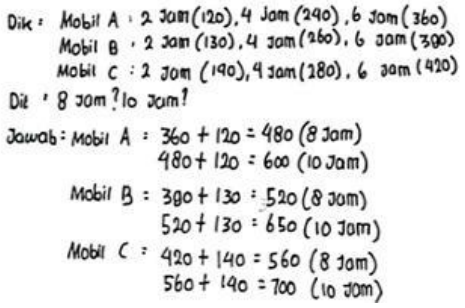
Translate :C. 1500 flower stalks

Figure 7. Student Answers to Level 2 Question

Figure 6 and 7 show that students could analyze flower sales data and estimate differences between months. For Level 1 and 2 questions, 83.33% answered correctly—high and medium ability students answered correctly, while low ability students still made errors.

In the level 3 question, students were asked to determine the distance that could be traveled for 8 hours and 10 hours by three different cars. Here is one of the students' answers to level 3 questions.

Translate:

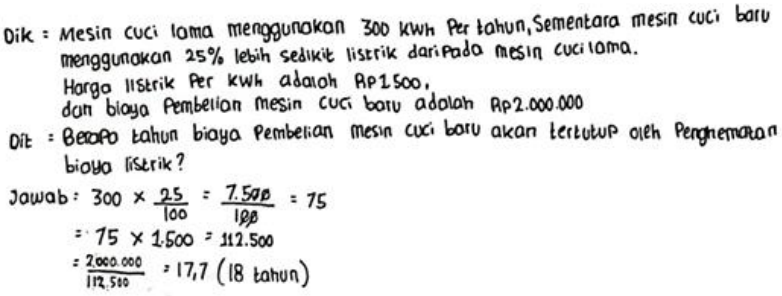


Given:
- Car A : 2 hours (120), 4 hours (240), 6 hours (360)
- Car B : 2 hours (130), 4 hours (260), 6 hours (390)
- Car C : 2 hours (140), 4 hours (280), 6 hours (420)

Question: 8 hours? 10 hours?
Answer:
- Car A :
  360 + 120 = 480 (8 hours)
  480 + 120 = 600 (10 hours)
- Car B :
  390 + 130 = 520 (8 hours)
  520 + 130 = 650 (10 hours)
- Car C :
  420 + 140 = 560 (8 hours)
  560 + 140 = 700 (10 hours)

Figure 8. Student Answers to Level 3 Question

Level 3: Almost all students answered correctly. High-ability students provided complete, systematic, and accurate solutions, while moderate- and low-ability students were also correct but often omitted or misstated information.


1 Dik : Mesin cuci lama menggunakan 300 kWh Per tahun, Sementara mesin cuci baru menggunakan 25% lebih sedikit listrik daripada mesin cuci lama.
Harga listrik Per kWh adalah Rp1500,
dan biaya Pembelian mesin cuci baru adalah Rp2.000.000
Dit : Berapa tahun biaya Pembelian mesin cuci baru akan tertutup oleh Penghematan biaya listrik?
Jawab : $300 \times \frac{25}{100} = \frac{7.500}{100} = 75$
$= 75 \times 1.500 = 112.500$
$= \frac{2.000.000}{112.500} = 17,7$ (18 tahun)

Translate:

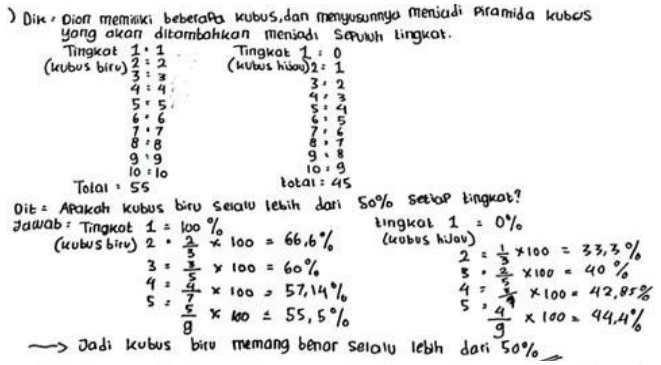
Given:
An old washing machine uses 300 kWh per year, while a new washing machine uses 25% less electricity than the old one.
The price of electricity per kWh is Rp1,500, and the cost of purchasing the new washing machine is Rp2,000,000.

Question:
How many years will it take for the savings in electricity costs to cover the purchase cost of the new washing machine?
Answer:

$$= 300 \times \frac{25}{100} = \frac{7500}{100} = 75$$
$$= 75 \times 1500 = 112.500$$
$$= \frac{2.000.000}{112.500} = 17,7 (18\ year)$$

Figure 9. Student Answers to Level 4 Question

At Level 4, 66.67% answered correctly: high-ability students solved systematically and accurately, moderate-ability often omitted details, and low-ability gave incomplete or incorrect answers.


) Dik : Dion memiliki beberapa kubus, dan menyusunnya menjadi piramida kubus yang akan ditambahkan menjadi Sepuluh tingkat.
Tingkat 1 : 1 (kubus biru) 2 : 2, 3 : 3, 4 : 4, 5 : 5, 6 : 6, 7 : 7, 8 : 8, 9 : 9, 10 : 10 Total : 55
Tingkat 1 : 0 (kubus hijau) 2 : 1, 3 : 2, 4 : 3, 5 : 4, 6 : 5, 7 : 6, 8 : 7, 9 : 8, 10 : 9 total : 45
Dit : Apakah kubus biru selalu lebih dari 50% setiap tingkat?
Jawab : Tingkat 1 = 100 %
(kubus biru) 2 = $\frac{2}{3}$ x 100 = 66,6 %
3 = $\frac{3}{5}$ x 100 = 60 %
4 = $\frac{4}{7}$ x 100 = 57,14 %
5 = $\frac{5}{9}$ x 100 = 55,5 %
Tingkat 1 = 0 %
(kubus hijau) 2 = $\frac{1}{3}$ x100 = 33,3 %
3 = $\frac{2}{5}$ x100 = 40 %
4 = $\frac{3}{7}$ x 100 = 42,85%
5 = $\frac{4}{9}$ x 100 = 44,4%
→ Jadi kubus biru memang benar selalu lebih dari 50%

Translate:

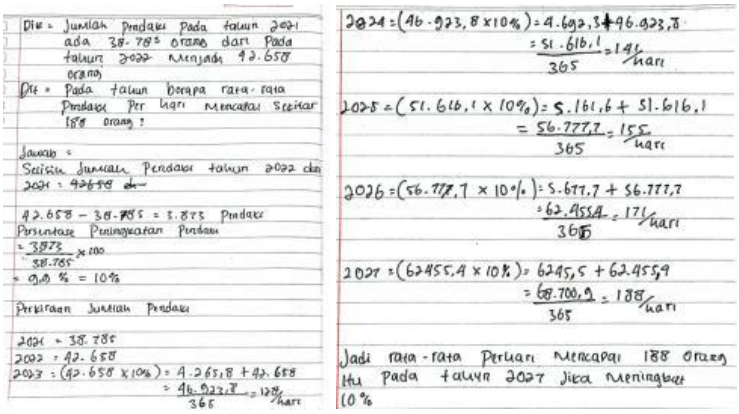
Given: Dion has several cubes, and he arranges them into a cube pyramid consisting of ten levels.

Level 1 : (Blue cubes)	Level 1 : 0 (Green cubes)
2 : 2	2 : 1
3 : 3	3 : 2
4 : 4	4 : 3
5 : 5	5 : 4
6: 6	6 : 5
7 : 7	7 : 6
8 : 8	8 : 7
9 : 9	9 : 8
10 : 10	10 : 9
Total : 55	Total : 45

Question: Is it true that the blue cubes are always more than 50% at each level?
The students calculated the percentages as shown in the figure and concluded that the blue cube was indeed always more than 50%.

Figure 10. Students' Answers to Level 5 Question

For Level 5, 50% of students answered correctly. High-ability students solved the problem but sometimes wrote incomplete or inaccurate information. Moderate-ability students often relied on guesses or incomplete reasoning, while low-ability students failed to state information correctly and produced wrong answers.



Translate and description:

Given:
The number of climbers in 2021 was 38.785 people and in 2022 it increased to 42.658 people.

Question:
In which year will the average number of climbers per day reach around 188 people?
Answer:
Students perform calculations to determine the difference in the number of climbers between 2022 and 2023, then calculate the percentage of climbers and also estimate the number of climbers using the calculations shown in the figure. They then write the conclusion: Therefore, the average number of climbers will reach 188 people in 2027 if there is a 10% increase.

Figure 11. Students' Answers to Level 6 Question

For Level 6, only 33.33% of students answered correctly. High-ability students solved the problem systematically, while some moderate-ability students relied on unstructured reasoning. Most moderate- and low-ability students gave incorrect answers and failed to state the required information.

Based on the research results, the developed PISA-like questions are valid and practical, as shown by students' work in the small group stage. The questions align with the curriculum and the ability level of grade VIII students. (Iswari et al., 2025). The PISA-like questions align with Change and Relationship content and enhance students' cognitive skills in analysis, reflection, conclusion, and real-life mathematical interpretation. (Suharyono et al., 2020).

CONCLUSION

Based on the development stages—including Preliminary and Formative Evaluation (Self-Evaluation, Expert Review, One-to-One, and Small Group)—the PISA-like questions were deemed valid and practical. Quantitatively, they met criteria for validity, reliability, difficulty, discrimination, and bias, while qualitatively, they aligned with the PISA Change and Relationship framework, curriculum, and grade VIII content. The questions were practical, clear, and allowed students to identify information and draw appropriate conclusions. Additionally, the six-level questions enhance students' critical thinking, reasoning, and ability to connect mathematics to real-world contexts. This research is still limited to the small group stage, so further research is needed at the broader field test stage to obtain more in-depth information about the effectiveness of the questions in learning and to improve the questions that have been developed.

REFERENCES

A. Safitri dan Khotimah. (2023). Kemampuan Literasi Matematika Peserta Didik dalam Menyelesaikan Soal PISA Konten Space and Shape Ditinjau dari Gaya Kognitif. *"Peran Soal PISA Dalam Meningkatkan Keterlibatan Siswa," Jurnal Pendidikan Dan Pembelajaran Matematika, 10*(3), 205–213. https://doi.org/10.34312/jmathedu.v4i1.18745

A. Susanta, H. Sumardi, E. Susanto, dan E. R. (2023). Mathematics literacy task on number pattern using Bengkulu context for junior high school students. *"Analisis Hasil PISA 2022 Indonesia," Jurnal Evaluasi Pendidikan, 14*(2), 45–58. https://doi.org/10.22342/JME.V14I1.PP85-102

Adawiyah, Y. N., Zakiah, N. E., & Zamnah, L. N. (2023). Kemampuan Pemahaman Konsep Matematis Siswa pada Materi Persegi dan Persegi Panjang. *Kemampuan Literasi Matematis Siswa Dalam Menyelesaikan Soal Pisa. J-Kip (Jurnal Keguruan Dan Ilmu Pendidikan)*, *4*(1), 198.

Anshari, A. M., Pasani, C. F., & Noorbaiti, R. (2023). Pengembangan Soal Matematika Model Pisa Level 6 Menggunakan Konteks Lahan Basah. *JURMADIKTA*, *3*(1), 56–67. https://doi.org/10.20527/jurmadikta.v3i1.1379

Azzahra. (2021). J urnal Pendidikan Matematika. *Validitas Pengembangan Soal Matematika Model PISA: Konteks Covid-19 Dalam 4 Konten Berbeda.JEMST (Jurnal of Education in Mathematics, Science, and Technology), 4*(1), 29-41.

Davita, P. W. C., & Pujiastuti, H. (2020). Analisis Kemampuan Pemecahan Masalah Matematis Siswa Kelas VIII pada Materi Pola Bilangan. *Nalisis Kemampuan Pemecahan Masalah Matematika Ditinjau Dari Gender. Kreano, Jurnal Matematika Kreatif-Inovatif, 11*(1), 110–117. https://doi.org/10.31949/dm.v4i1.2012

Efriani, A., Febriyanti, Q. A., Zahra, A., & Ramury, F. (2021). Pengembangan Soal Matematika Model PISA Menggunakan Konteks Kalimantan Timur. *Soal Tipe PISA : Konteks Peningkatan Kasus Covid-19.*, *4*(1), 1-10. https://doi.org/10.35974/jpd.v3i1.2230

Fadlila, N., Hayuningrat, S., & Wijaya, A. (2023). Developing mathematics PISA-like problems using change and relationship content. *"Developing Mathematics PISA-like Problems Using Change and Relationship Content," AIP Conference Proceedings, 2733*(1). https://doi.org/10.1063/5.0141064

Fitriani, N. (2021a). Analisis Tingkat Kesukaran, Daya Pembeda, Dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal. *Paedagoria : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan*, *12*(2), 199. https://doi.org/10.31764/paedagoria.v12i2.4956

Fitriani, N. (2021b). Analisis Tingkat Kesukaran, Daya Pembeda, Dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal. *"Analisis Tingkat Kesukaran, Daya Pembeda, Dan Efektivitas Pengecoh Soal Pelatihan Kewaspadaan Kegawatdaruratan Maternal Dan Neonatal," Paedagoria : Jurnal Kajian, Penelitian Dan Pengembangan Kependidikan, 12*(2), 199. https://doi.org/10.31764/paedagoria.v12i2.4956

Hadari Nawawi, J. H. (2023a). *Indonesian Journal of Learning and Instructional Innovation*. *1*(2).

Hadari Nawawi, J. H. (2023b). *Indonesian Journal of Learning and Instructional Innovation*. *1*(2).

Hardianti. (2019). Pengembangan Soal Matematika Model PISA Konteks Kain Cual Bangka Belitung. *"Pengembangan Soal PISA Konteks Pribadi," Jurnal Inovasi Pendidikan Matematika, 7*(2), 89–97. https://doi.org/10.31004/cendekia.v3i2.114

Indrawati, I., Yuniawatika, Y., & Suminah, S. (2022). Pengembangan Instrumen Tes Setipe PISA untuk Melatih Kemampuan Literasi Matematika Siswa Kelas V SD LAB UM. *Jurnal Pembelajaran, Bimbingan, Dan Pengelolaan Pendidikan*, *2*(7), 629–639. https://doi.org/10.17977/um065v2i72022p629-639

Iswari, I., Pasaribu, F. T., Gustiningsi, T., & Nusantara, D. S. (2025). Pengembangan Soal Matematika Tipe Pisa Level 6 Menggunakan Konteks Benda Bersejarah. *Numeracy*, *12*(1), 19–36. https://doi.org/10.46244/numeracy.v12i1.3134

Ketaren, M. A., Girsang, K., Manurung, M., & Riauati, E. (2024). *96.+Production_Maya+Alemina+Ketaren (1)*. *1*(11), 3278–3283.

Lusinda, U., Pasaribu, F. T., Gustiningsi, T., & Nusantara, D. S. (2025). Pengembangan Soal Matematika Tipe Pisa Konten Uncertainty and Data Menggunakan Konteks Museum Siginjai Jambi. *Numeracy*, *12*(1), 99–117. https://doi.org/10.46244/numeracy.v12i1.3128

Machmuda, R., Edy, S., & Suryanti, S. (2024). Analisis Kemampuan Komunikasi Matematis Siswa dalam Menyelesaikan Soal PISA. *Ideguru: Jurnal Karya Ilmiah Guru*, *9*(2), 883–892. https://doi.org/10.51169/ideguru.v9i2.984

Mayari, R. P. (2020). Pengembangan Soal Matematika Model PISA dengan Menggunakan Konteks Budaya Palembang. *Penggembangan Soal Matematiika Model PISA Menggunakan Konteks Palembang*, *10*(2), 2. https://doi.org/10.25273/jems.v10i2.11761

Mitari, O., & Zulkardi. (2019). Pengembangan Soal Matematika Model PISA dengan Konteks Wisata Jakabaring Sport City. *Seminar Nasional Pendidikan Matematika Ahmad Dahlan*, *6*, 172–178.

Murtiyasa, A. dan. (2023). PISA Problem-Solving of Students' Mathematical Literacy in Adversity Quotient Perspective. *The Numerical Ability of Students in Solving PISA-like Problems for Change and Relationship Content.*, *8*(2), 128–140. https://doi.org/10.26740/jrpipm.v8n2.p128-140

Nusantara, D. S., Zulkardi, & Putri, R. I. I. (2021). Designing pisa-like mathematics task using a COVID-19 context (Pisacomat). *Journal on Mathematics Education*, *12*(2), 349–364. https://doi.org/10.22342/JME.12.2.13181.349-364

OECD. (2019). Locating inclusion within the OECD's assessment of global competence: An inclusive future through PISA 2018? *PISA 2018 Assessment and Analytical Framework, Paris: OECD Publishing, 20*(1), 56–72. https://doi.org/10.1177/14782103211006636

OECD. (2022). Pisa 2022 Results. In *Factsheets: Vol. I*.

Oktaviana, D., & Clark, B. (2020). * Corresponding Author * Corresponding Author. *Scientific African*, *114*(June), e00146.

Purwanto, A. H., Yogyakarta, U. N., Jl, A., No, C., Malang, K., Depok, K., Sleman, K., & Yogyakarta, I. (2024). *Pengembangan Soal Matematika Model AKM Menggunakan Konteks Budaya Yogyakarta untuk Mengetahui Profil Literasi Matematika Siswa Sekolah Menengah Atas PISA (Programme for International Student Assessment) adalah suatu studi internasional di bidang pendidi*. *6*.

Putri, D., & Retnosari, D. S. (2024). *Analisis Kualitas Soal Penilaian Akhir Semester Mata Pelajaran Dasar Program Keahlian Akuntansi dan Keuangan Lembaga di SMK Ketintang Surabaya*. *7*(April), 8–17.

Rawani, D. (2019). Pengembangan Soal Matematika Tipe PISA Konten Space and Shape. *PISA-like Mathematics Problems: Using Taekwondo Context of Asian Games. Journal on Mathematics Education.*, *10*(2), 277–288. https://doi.org/10.35194/jp.v10i2.1570

Riyani, R., Maizora, S., & Hanifah, H. (2017). Uji Validitas Pengembangan Tes Untuk Mengukur Kemampuan Pemahaman Relasional Pada Materi Persamaan Kuadrat Siswa Kelas Viii Smp. *Jurnal Penelitian Pembelajaran Matematika Sekolah (JP2MS)*, *1*(1), 60–65. https://doi.org/10.33369/jp2ms.1.1.60-65

Sri Mertasari, N. M., & Candiasa, I. M. (2022). Formative Evaluation of Digital Learning Materials. *Journal of Education Technology*, *6*(3), 507–514. https://doi.org/10.23887/jet.v6i3.44165

Suharyono, E., & Rosnawati, R. (2020). Analisis Buku Teks Pelajaran Matematika SMP ditinjau dari Literasi Matematika. *Mosharafa: Jurnal Pendidikan Matematika*, *9*(3), 451–462. https://doi.org/10.31980/mosharafa.v9i3.628

Zulkardi. (2021). Developing A Learning Environment On Realistic Mathematics Education For Indonesian Student Teachers. *PrintPartners Ipskamp*, *January 2002*, 218.