

ASSESSING THE RELIABILITY AND VALIDITY CRITICAL THINKING QUESTIONNAIRE (CThQ)

**Zahra Shafira^{1,6}, Sri Tiatri^{2,6}, Shalsa Dea Purnama^{3,6}, Ayesha Desfitrianie^{4,6}
& Jap Tji Beng^{5,6}**

¹ Psychology Undergraduate Program, Universitas Tarumanagara, Jakarta, Indonesia
Email: zahra.705210249@stu.untar.ac.id

² Faculty of Psychology, Universitas Tarumanagara, Jakarta, Indonesia
Email: sri.tiatri@untar.ac.id

³ Psychology Undergraduate Program, Universitas Tarumanagara, Jakarta, Indonesia
Email: shalsa.705210263@stu.untar.ac.id

⁴ Psychology Undergraduate Program, Universitas Tarumanagara, Jakarta, Indonesia
Email: ayesha.705210243@stu.untar.ac.id

⁵ Faculty of Information System, Universitas Tarumanagara, Jakarta, Indonesia
Email: t.jap@untar.ac.id

⁶ Science, Technology and Society Research Center, Universitas Tarumanagara, Jakarta

ABSTRACT

Critical thinking is a logical, rational, and reflective process, making it essential to develop an effective measurement tool. This study has three objectives: (1) to translate the Critical Thinking Questionnaire (CThQ) from English to Indonesian; (2) to ensure the reliability and validity of the Indonesian version of the CThQ as a measure of critical thinking ability; (3) to adjust several items of the measurement tool to better align with specific measurement needs. The Indonesian version of CThQ consists of 25 items to assess a person's ability to evaluate information's validity. The research was conducted using a quantitative method through a questionnaire administered to 416 students. The instrument adaptation process was carried out in three stages: translation using the forward-backward translation technique, expert validation for item adjustment, and empirical testing. The analysis results showed that the Indonesian version of the CThQ has good internal consistency ($\alpha = 0.875$). The validity of the instrument was analyzed by comparing the r -calculated value obtained from the Pearson correlation with the r -table value at a significance level of 0.05. The results of the validity analysis show that the 25 items in CThQ have r -calculated values greater than r -table ($r_{table} = 0.096$), indicating that the items are valid. These findings indicate that the adapted Indonesian version of the CThQ is a valid and reliable instrument for measuring critical thinking skills. From the analysis conducted, the CThQ demonstrated satisfactory results in reliability and validity, making it a highly suitable instrument for use in educational contexts.

Keywords: reliability, validity, critical thinking, university students

1. PREFACE

In the industrial era 4.0, critical thinking has become a crucial skill in various fields, especially in the education sector (Tiatri et al., 2023). Critical thinking is an essential skill in the 21st century era that involves thinking processes with clear and systematically organized goals (Tiatri et al., 2023). Critical thinking according to Chance is the ability to examine information, develop and organize ideas, defend opinions, compare things, draw conclusions, evaluate arguments, and solve problems (Chance, 1986). Critical thinking according to Mertes is a process that is carried out consciously and purposefully to analyze or evaluate information and experiences, relying on reflective attitudes and skills that support the formation of beliefs and making appropriate and wise decisions (Mertes, 1991). In addition, according to Ennis, critical thinking is logical reflective thinking aimed at determining the most appropriate thing to believe or do (Ennis, 2018). In higher education, students face increasingly complex challenges, starting from the rampant spread of hoaxes (Suryanto et al., 2018). The phenomenon of massive information dissemination, especially through social media and other digital platforms, has become one of the main challenges such as hoax news aimed at leading opinions (Tamhidah, 2023). The information circulating is not always verified, so students are often faced with difficulties in

distinguishing facts, opinions, and hoaxes. This phenomenon not only has the potential to mislead, but can also influence decision-making, both in academic contexts and everyday life.

In addition, the world of higher education today no longer focuses only on the ability to memorize or understand theory. Universities are required to produce graduates who are able to think critically, analytically, and solutively in dealing with various problems. This is in line with the demands of the industrial era 4.0 which emphasizes the importance of 21st century skills, such as critical thinking, creativity, collaboration, and communication (Irawan et al., 2020). In this context, critical thinking becomes a core skill that enables students to analyze problems in depth, evaluate arguments logically, and produce innovative solutions. The idea of the importance of critical thinking in education was actually introduced by John Dewey in 1910, and since then, many educational institutions have started to implement it (Hitchcock, 2018).

To be able to measure and assess the extent of a person's critical thinking skills, a valid and reliable instrument is needed. One of the instruments used in measuring critical thinking skills is the Critical Thinking Questionnaire (CThQ) developed by Aleksander Kobylarek, Kamil Blaszczyński, Luba Słosarz, and Martyna (Kobylarek et al., 2022). CThQ is designed to identify and evaluate various aspects of critical thinking skills, such as analysis, synthesis, evaluation, and problem-solving abilities. In this study, the researchers adapted the CThQ instrument into Bahasa Indonesia to ensure that the measurement tool can be used effectively in the Indonesian educational context. This adaptation process involves translation, cultural adjustment, and verification of contextual appropriateness so that the instrument can be interpreted appropriately by respondents. In addition, these adjustments were also important to avoid any distortion of meaning that might occur due to language and cultural differences between the original and adapted instruments.

After the adaptation process, the translated instrument was then tested for reliability and validity to ensure that it retained the consistency and accuracy of the original intended measurement. This evaluation involves testing the level of internal consistency as well as the extent to which the instrument actually measures critical thinking skills in the Indonesian educational context. Research conducted by Skorkova and Gazova-adamkova using the CThQ instrument to measure critical thinking skills in university students, showed an alpha reliability coefficient of 0.766. This shows that CThQ is reliable for measuring critical thinking skills (Skorkova & Gazova-Adamkova, 2024). Other research conducted by Nurwahyuni using CThQ to measure critical thinking skills in students shows that the CThQ measuring instrument is quite good in measuring critical thinking skills (Nurwahyuni, 2023).

Furthermore, research conducted by Barevic iu te et al. using the CThQ instrument shows that Cronbach's Alpha is 0.821, which means that the reliability coefficient and internal consistency assessment measures have been met (Barevičiūtė et al., 2023). Research by Bhuttah et al. adapted one question item from each dimension of the CThQ instrument and had a Cronbach's Alpha value of 0.729 which indicated that the instrument was valid (Bhuttah et al., 2024). Then another study conducted by Gayatri and Sit adapted and modified the CThQ instrument according to the context and showed that the validity and reliability of the CThQ instrument was good (Gayatri & Sit, 2024). These various studies underlie this research to systematically assess and evaluate the reliability and validity of the Indonesian version of CThQ as a credible tool for measuring critical thinking skills.

2. RESEARCH METHOD

Samples

This research was conducted using a quantitative method with purposive sampling technique, which is a method of selecting samples based on specific criteria that have been previously determined. The criteria for participants required in this study are (a) male and female; (b) active students; (c) age 18-25 years; (d) domicile in Jakarta. Data collection in this study was conducted through Google Form. The number of participants in this study was 416 participants. Participants were classified based on gender. For more details can be seen in Table 1.

Table 1
Participant overview by gender

Gender	Frequency	Percentage
Male	127	30.5
Female	289	69.5
Total	416	100.0

In this study, participants were categorized by gender to see the distribution of male and female participants. Of the total participants involved, 127 were male, which when converted into a percentage, equates to 30.5% of the total participants.

Meanwhile, there were 289 female participants compared to male participants. This figure represents 69.5% of the overall participants who participated in the study. Thus, the distribution of participants shows that the number of women is more dominant than men in this study. Meanwhile, a description of participants based on age. For more details can be seen in Table 2.

Table 2
Overview of participants by age

Age	Frequency	Percentage
18	5	1.2
19	31	7.5
20	77	18.5
21	162	38.9
22	84	20.2
23	33	7.9
24	18	4.3
25	6	1.4
Total	416	100.0

In this study, the participants involved ranged in age from 18 to 25 years old. Of the total number of participants, 5 participants or about 1.2% were 18 years old. Meanwhile, the 19-year-old age group includes 31 participants, which is equivalent to 7.5% of the total participants. The 20-year-old age group had a larger number of 77 participants or 18.5%.

Participants who were 21 years old accounted for the largest number of participants, with 162 people or about 38.9% of all participants involved in this study. Furthermore, there were 84 participants aged 22, which accounted for 20.2% of the total. Meanwhile, the 23-year-old group consisted of 33 participants or equivalent to 7.9%. Then, there were also 18 participants aged 24, which represented 4.3% of the total participants. Finally, the least number of participants came from the 25-year-old age group, which was 6 people or around 1.4% of all participants who participated in this study. Meanwhile, a description of participants based on domicile. For more details can be seen in Table 3.

Table 2
Overview of participants by domicile

Domicile	Frequency	Percentage
Jakarta Barat	108	26.0
Jakarta Pusat	65	15.6
Jakarta Selatan	100	24.0
Jakarta Timur	63	15.1
Jakarta Utara	80	19.2
Total	416	100.0

In this study, participants were grouped by domicile to understand the geographical distribution of participants involved. Of the total number of participants, 108 were from the West Jakarta area, which is equivalent to 26% of the total participants who took part in this study. Meanwhile, 65 participants were recorded as domiciled in Central Jakarta, with a percentage of 15.6%. Then, there were 100 participants from South Jakarta, which accounted for 24% of the total number of participants.

On the other hand, participants from East Jakarta totaled 63 people, or around 15.1% of the total participants in this study. There were 80 participants who lived in North Jakarta, with a percentage of 19.2%. Thus, the distribution of participants' domicile shows that the majority of participants come from West Jakarta and South Jakarta, while the number of participants from Central Jakarta and East Jakarta is relatively smaller than other regions. Another description of participants is based on faculty. For more details can be seen in Table 4.

Table 4
Participant overview by faculty

Faculty	Frequency	Percentage
Engineering	52	12.5
Psychology	138	33.2
Information System	12	2.9
Communication Science	10	2.4
Law	42	10.1
Administrative Science	6	1.4
Economics and Business	77	18.5
Social and Political Sciences	15	3.6
Education Science	4	1.0
Design and Creative Industry	1	.2
Computer Science	4	1.0
Math and Science	12	2.9
Language and Arts	11	2.6
Medicine	12	2.9
Engineering and Marine Science	1	.2
Fine Art and Design	5	1.2
Agriculture	6	1.4
Nursing Science	6	1.4
Public Health	2	.5
Total	416	100.0

In this study, participants were categorized based on their faculty of origin to understand the academic distribution of the participants involved. Of the total number of participants, 52 people came from the Faculty of Engineering, which when converted into a percentage, is equivalent to

12.5% of the total research participants. Meanwhile, the Faculty of Psychology had the highest number of participants, with 138 people, accounting for 33.2% of the total participants. On the other hand, the Faculty of Information Systems has a smaller number of participants, as many as 12 people or about 2.9% of the total participants who participated.

The Faculty of Communication has 10 participants, with a percentage of 2.4%, while the Faculty of Law has 42 participants or about 10.1% of the total. The Faculty of Administrative Sciences only had 6 participants, with a percentage of 1.4%. The Faculty of Economics and Business also has a significant number of participants, as many as 77 people or around 18.5% of the total participants. Meanwhile, the Faculty of Social and Political Sciences has 15 participants with a percentage of 3.6%.

The Faculty of Education has a smaller number of participants, as many as 4 people or 1% of the total participants. The Faculty of Design and Creative Industries has only 1 participant, which is the smallest number in this study, with a percentage of 0.2%. Meanwhile, the Faculty of Computer Science has the same number of participants as the Faculty of Education, which is 4 people or about 1% of the total participants. The Faculty of Mathematics and Natural Sciences had 12 participants or 2.9%. The Faculty of Language and Arts recorded 11 participants with a percentage of 2.6%.

The Faculty of Medicine also has the same number of participants as the Faculty of Mathematics and Natural Sciences, which is 12 people or about 2.9% of the total participants. The Faculty of Engineering and Marine Science only has 1 participant with a percentage of 0.2%, the same number as the Faculty of Design and Creative Industries. Furthermore, the Faculty of Fine Arts and Design has 5 participants or around 1.2% of the total participants, while the Faculty of Agriculture, Faculty of Nursing Science, and Faculty of Administrative Science each have 6 participants with a percentage of 1.4%. The Faculty of Public Health is one of the faculties with the least number of participants, which is only 2 people or around 0.5% of the total research participants.

Thus, the distribution of participants by faculty shows that the Faculty of Psychology has the largest number of participants, followed by the Faculty of Economics and Business and the Faculty of Engineering. Meanwhile, some other faculties have fewer participants, with the Faculty of Design and Creative Industries and the Faculty of Engineering and Marine Science having the least number of participants in this study.

Measurement

The measuring instrument used in this study is the Critical Thinking Questionnaire (CThQ) developed by Kobylarek et al. (2022). This measuring instrument aims to measure a person's critical thinking ability. CThQ has 6 dimensions, namely analyzing, evaluating, creating, understanding, applying, and remembering. This measuring instrument uses Likert 1-5 which consists of strongly disagreeing to strongly agreeing.

Data collection and analysis

Data collection was carried out by distributing questionnaires via Google Form to students with an age range of 18-25 years in Jakarta. Data processing in this study used SPSS version 25.0.

3. RESULTS AND DISCUSSIONS

The face validity test of the CThQ instrument was conducted to ensure that the items were relevant and appropriate for measuring critical thinking skills. Face validity is determined

through reviews from experts in the field of Psychology who assess whether the items are clear, relevant, and in accordance with the aspects of critical thinking to be measured. Based on the results of the experts' review, the instrument is considered to have good face validity, because the items are related to the dimensions of critical thinking. Based on content validity, the instrument was thoroughly reviewed to determine whether the items comprehensively represented the entire spectrum of critical thinking skills. The CThQ instrument includes six dimensions, namely analyzing, evaluating, creating, understanding, applying, and remembering. Expert judgment was used to assess whether the items covered all of these dimensions.

In this study, the analysis of the validity of instrument items was carried out by comparing the r -count value obtained from the Pearson correlation results with the r -table (0.096) at the 0.05 significance level. Based on the results of the validity test, it was found that all items in the instrument met the validity criteria, as indicated by the r -count value which was greater than the r -table at the predetermined significance level. In other words, each item in this measuring instrument has a strong enough relationship, so it can be said that the instrument has adequate validity.

This finding indicates that the process of adapting the Critical Thinking Questionnaire (CThQ) measuring instrument has been carried out thoroughly and systematically. During the adaptation process, aspects of content and format validity were carefully considered to ensure that each item in the measuring instrument was able to measure critical thinking skills accurately and in accordance with the characteristics of the target population. Thus, the measuring instrument used in this study can be considered a valid and reliable instrument in measuring critical thinking skills.

Then, the reliability test results are carried out to determine the Cronbach Alpha value in each CThQ dimension. Before item elimination, the total number of items in the instrument was 25 with the following details: 4 items in Analysing, 4 items in Evaluating, 6 items in Creating, 4 items in Understanding, 4 items in Applying, and 3 items in Remembering. After eliminating some of the less reliable items, the total number of items in the instrument was reduced to 21, with changes in several dimensions, namely Analyzing was reduced to 3 items, Creating to 5 items, Understanding to 3 items, and Remembering to 2 items, while the Evaluating and Applying dimensions remained with the same number of items.

In terms of reliability, the Cronbach's Alpha value before elimination shows that some dimensions have a low level of reliability, especially in the dimensions of Analyzing ($\alpha = 0.298$), Creating ($\alpha = 0.488$), Understanding ($\alpha = 0.420$), Remembering ($\alpha = 0.413$). Other dimensions have higher values, such as Evaluating ($\alpha = 0.534$) and Applying ($\alpha = 0.557$). After eliminating some items that did not contribute to the reliability of the instrument, there was an increase in most of Cronbach's Alpha values. The Analysing dimension experienced a significant increase from 0.298 to 0.520, indicating an increase in internal consistency after some items were removed. Similarly, the Creating dimension saw an increase from 0.488 to 0.615, and the Understanding dimension increased from 0.420 to 0.530. The Remembering dimension also experienced a slight increase from 0.413 to 0.493. Meanwhile, the Evaluating and Applying dimensions still have the same Cronbach's Alpha value, 0.534 and 0.557 respectively, because there is no change in the number of items.

Overall, these results show that the item elimination process succeeded in improving the reliability on most dimensions in the CThQ instrument, especially on dimensions with low initial

scores. This indicates that after item elimination, the instrument becomes more consistent in measuring critical thinking skills in each of the established dimensions. For more details can be seen in Table 5.

Table 5
Reliabilitas dimensi CThQ

Dimension	Before Item Elimination	After Item Elimination	Cronbach α Before Item Elimination	Cronbach α After Item Elimination
Analysing	4	3	.298	.520
Evaluating	4	4	.534	.534
Creating	6	5	.488	.615
Understanding	4	3	.420	.530
Applying	4	4	.557	.557
Remembering	3	2	.413	.493
Total	25	21		

After testing the reliability of each dimension in the instrument, it was found that there were four items that did not meet the predetermined reliability standards. These items had to be eliminated because their reliability values were not high enough to support the internal consistency of the instrument as a whole. As a next step, reliability testing was conducted again by eliminating the four items that did not meet the standard. This test aims to determine whether the elimination of these items can increase the overall reliability of the instrument used in this study.

The results of the overall reliability test showed that after the removal of the four items, the Cronbach's Alpha value increased and reached 0.875. This value indicates that the instrument used has a very good level of reliability, because in general, Cronbach's Alpha values above 0.8 indicate strong internal consistency. For more details can be seen in Table 6.

Table 6
CThQ reliability

Cronbach's Alpha	N of items
0.875	21

Thus, it can be concluded that the measurement instruments used in this study have high reliability and are able to produce consistent data. This means that each of the remaining items in the instrument is able to measure the intended concept properly without any significant inconsistencies. Therefore, this instrument can be trusted and is suitable for use in further research to measure aspects relevant to the research objectives more accurately.

4. CONCLUSIONS AND RECOMMENDATIONS

This study successfully adapted and tested the validity and reliability of the Critical Thinking Questionnaire (CThQ) instrument that has been translated into Indonesian to measure the critical thinking skills of university students in Indonesia. The validity test results show that all items in this instrument have a strong enough relationship with the dimensions of critical thinking skills measured, so this instrument can be considered valid. The process of eliminating less reliable items succeeded in improving internal consistency in most dimensions, such as the Analysing, Creating, Understanding, and Remembering dimensions. Overall, the Cronbach's Alpha value of 0.875 after elimination indicates that this instrument has excellent reliability and can be trusted

to measure students' critical thinking skills.

Based on the results of this study, it is recommended that the adapted CThQ instrument be used more widely in educational research, especially to measure critical thinking skills among university students in Indonesia. In addition, it is necessary to further test this instrument in various contexts and more diverse populations to ensure that this instrument can be adapted properly. Researchers are also advised to conduct further evaluation and development related to other dimensions that might enrich the measurement of critical thinking skills.

Acknowledgment

The researcher would like to thank all participants who have been willing to take the time and energy to contribute to this study. The researcher would also like to thank those who have helped in this research.

REFERENCES

Barevičiūtė, J., Dadelo, S., & Asakavičiūtė, V. (2023). The skills of critical thinking, creativity, and communication as tools for overcoming social simulation in the context of sustainability: A case study of students' self-assessment of the affective domain of learning. *Sustainability*, 15(14), 10935. <https://doi.org/10.3390/su151410935>

Bhutta, T. M., Xusheng, Q., Abid, M. N., & Sharma, S. (2024). Enhancing student critical thinking and learning outcomes through innovative pedagogical approaches in higher education: The mediating role of inclusive leadership. *Scientific Reports*, 14(1), 24362. <https://doi.org/10.1038/s41598-024-59968-0>

Chance, P. (1986). *Thinking in the classroom: A survey of programs*. Teachers College Press.

Ennis, R. H. (2018). Critical thinking across the curriculum: A vision. *Topoi*, 37, 165–184. <https://doi.org/10.1007/s11245-016-9401-4>

Gayatri, P., & Sit, H. (2024). The adaption of sustainable blended global discussion (SBGD) in English as a foreign language teaching and learning. *Education Sciences*, 14(12), 1279. <https://doi.org/10.3390/educsci14121279>

Hitchcock, D. (2018). Critical thinking. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2018 Edition). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2018/entries/critical-thinking/>

Irawan, E., Arif, S., Hakim, A. R., Fatmahanik, U., Fadly, W., Hadi, S., Pahlevi, F. S., & Aini, S. (2020). *Pendidikan tinggi di masa pandemi: Transformasi, adaptasi, dan metamorfosis menyongsong new normal*. Zahir Publishing.

Kobylarek, A., Błaszczyński, K., Ślósarz, L., & Madej, M. (2022). Critical Thinking Questionnaire (CThQ): Construction and application of critical thinking test tool. *Andragogy Adult Education and Social Marketing*, 2(2), 1–11. <https://doi.org/10.54915/aaesm.v2i2.11>

Mertes, L. M. (1991). Thinking and writing. *Middle School Journal*, 22(5), 24–25. <https://doi.org/10.1080/00940771.1991.11496131>

Nurwahyuni, N. (2023). Student academic stress reviewed from critical thinking skills. *Jurnal Aisyah: Jurnal Ilmu Kesehatan*, 8(2), 250–257. <https://doi.org/10.30604/jika.v8i2.2036>

Skorková, Z., & Gažová-Adamková, H. (2024). Evaluating critical thinking in higher education: An analysis of university students using the CThQ. In *Proceedings of IAC in Budapest 2024* (p. 61). Czech Institute of Academic Education.

Suyanto, T., Prasetyo, K., Isbandono, P., Zain, I. M., Purba, I. P., & Gamaputra, G. (2018). Persepsi mahasiswa terhadap kemunculan berita bohong di media sosial. *Jurnal Civics*:

Media Kajian Kewarganegaraan, 15(1), 52–61.
<https://doi.org/10.21831/civics.v15i1.19137>

Tamhidah, M. A. R. (2023). Pengaruh media sosial terhadap penyebaran informasi palsu dan kejahatan siber. *Innovative: Journal of Social Science Research*, 3(6), 9133–9147.
<https://doi.org/10.31004/innovative.v3i6.13702>

Tiatri, S., Ie, M., Hussy, C., Tirtha, I. B. A. V., Timoria, L. T., Natan, M. T., & Beng, J. T. (2023). Development of the critical thinking categorical syllogism learning model in elementary students. *International Journal of Application on Social Science and Humanities*, 1(3), 1–8. <https://doi.org/10.55559/ijassh.v1i3.15>

Tiatri, S., Veronica, C., Fiscarina, C., Nurkholiza, R., Wakano, V. Y., Ie, M., & Beng, J. T. (2023). Elementary school teachers' perceptions of critical thinking in STEM learning. *International Journal of Application on Social Science and Humanities*, 1(1), 648–658.
<https://doi.org/10.55559/ijassh.v1i1.7>