Type of Contribution:
Editorial
Research Paper
Review Paper
Case Study

ENERGY: JURNAL ILMIAH ILMU-ILMU TEKNIK
Vol. 16, No. 2 (2026) pp 431-444
DOI: 10.51747/energy.v16i2.p431-444

E-ISSN: 2962-2565

This article contributes to:

Understanding Cognitive Processes Behind "I Don't Know" Responses in Sustainability Literacy Assessment Using Eye-Tracking and Usability Evaluation

Zidni Ilma Banaina[1]*, Maria Anityasari[1], Ratna Sari Dewi[1]

[1] Department of Industrial and Systems Engineering, Sepuluh Nopember Institute of Technology, ITS Keputih Campus, Surabaya 60111, **Indonesia.**

* zidnibanaina08@gmail.com

Abstract

Assessing sustainability literacy is essential for ensuring that sustainability literacy tests accurately measure individuals' understanding of sustainability concepts. However, conventional evaluation methods primarily rely on test scores and response accuracy, providing limited insight into the cognitive processes underlying participants' decision-making. This study investigates the use of eye-tracking technology to analyze participants' visual attention while completing a Sustainability Literacy Test (SLT). Eye movement data, including scanpaths, were analyzed to identify user interactions with test items and determine which visual elements attracted the greatest attention, particularly when participants selected the "I Don't Know" response option. In addition, system usability was evaluated using the Hybrid System Usability Scale (H-SUS). The results indicate that participants required an average response time of more than 60 seconds for items associated with the "I Don't Know" option, suggesting increased visual attention and cognitive load in specific areas of the interface. The website achieved an average H-SUS score of 63.34, indicating an acceptable level of usability, corresponding to Grade C and an "OK" adjective rating. These findings demonstrate that eye-tracking data can reveal hidden patterns of cognitive effort and user uncertainty that cannot be captured through response accuracy alone. The proposed approach provides valuable insights for improving test design, enhancing question readability, and optimizing the user interface of web-based Sustainability Literacy Tests, thereby supporting more effective assessment of sustainability literacy.

Article Info
Submitted: 2026-06-24
Revised: 2026-08-01
Accepted: 2026-08-07
Published: 2026-08-08

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Publisher
Universitas Panca Marga

Keywords: Eye Tracking, Hybrid System Usability Scale, Sustainability Literacy Test, User Interface, Visual Attention

1. Introduction

Sustainability has become a global priority across higher education institutions, driven by the increasing emphasis on achieving the United Nations Sustainable Development Goals (SDGs). In Indonesia, the Ministry of Higher Education, Science, and Technology (Kemdiktisaintek) adopts the QS World University Rankings (QS WUR) and the Times Higher Education University Impact Rankings (THE Impact Rankings) as key indicators for evaluating universities' contributions to sustainable development. Nevertheless, despite the large number of higher education institutions in Indonesia, institutional commitment to sustainability remains relatively low, as reported by Ardaniah et al. [1]. This condition indirectly reflects the level of engagement of students, academic staff, and faculty members in achieving the SDGs, highlighting the need for individuals to understand global sustainability challenges and their responsibility in addressing them.

To assess this understanding, sustainability literacy is commonly evaluated through web-based assessment platforms. Recognizing the importance of sustainability literacy, Institut Teknologi Sepuluh Nopember (ITS) became the first university in Indonesia to implement a Sustainability Literacy Test (SLT) in 2024 to measure participants' understanding of the Sustainable Development Goals [2]. The assessment was collaboratively developed by the ITS Global Engagement or Directorate of Student Affairs (DKG ITS). The initial implementation of the Sustainability Literacy Test was conducted through the myITS Classroom platform, which functioned as the first prototype for administering the assessment. To improve usability, accessibility, and participant experience, the system was subsequently redesigned into a dedicated Sustainability Literacy Test platform featuring a simplified interface, participant-specific access, and clearer presentation of assessment information, including the number of questions and test duration.

The Sustainability Literacy Test (SLT) includes an "I Don't Know" response option, allowing participants to indicate uncertainty rather than selecting an answer through random guessing. According to the 2024 assessment data, 244 faculty members and administrative staff selected the "I Don't Know" option, while data collected in 2024 and 2025 showed that 7,240 students chose the same response [2]. Although these responses reflect uncertainty, the underlying cognitive processes leading participants to select this option remain poorly understood. Eye-tracking technology offers an effective approach for examining visual attention and behavioral patterns during test completion, particularly when participants choose

the "I Don't Know" response [3], [4]. Compared with conventional process data sources, such as think-aloud protocols, computer-generated log files, and administrator observations, eye-tracking provides objective and fine-grained evidence of visual attention and cognitive processing that cannot be captured through traditional assessment methods [5], [6].

This study focuses on faculty members and administrative staff as the research participants. This selection was made because the eye-tracking equipment is located in the EPSK Laboratory, where access is limited, and because faculty members and staff are assumed to possess broader knowledge of sustainability concepts, thereby reducing the likelihood of selecting the "I Don't Know" response. Furthermore, this decision aligns with recommendations from the Vice Rector for Academic and Student Affairs of ITS, who emphasized in an interview conducted in December 2025 that faculty members and administrative staff should be prioritized in sustainability literacy assessments because they serve as the primary agents for transferring sustainability knowledge to students. Consequently, establishing a satisfactory level of sustainability literacy among faculty and staff is considered an essential prerequisite before extending the assessment to the student population.

Previous studies have employed eye-tracking to investigate cognitive load and visual attention in multiple-choice assessments, mathematical problem solving involving vectors, digital competency tests, the Programme for the International Assessment of Adult Competencies (PIAAC), interactive analytical problem-solving tasks, and medical education assessments [3]–[9]. However, little attention has been given to understanding the cognitive processes associated with selecting an "I Don't Know" response in sustainability literacy assessments. Addressing this gap, the present study aims to investigate participants' eye movement patterns while answering multiple-choice questions and to determine whether selecting the "I Don't Know" option is associated with increased cognitive load and visual attention during decision-making.

2. Methods

The study was conducted through several sequential stages, including defining the research object, designing the experimental tasks, selecting the participants, collecting the data, analyzing the results, and drawing conclusions. The overall research framework comprised an analysis of the Sustainability Literacy Test, task design, and data collection using an eye-tracking system. **Figure 1** illustrates the workflow of the proposed research methodology.

This study was designed as an exploratory qualitative eye-tracking study to investigate visual attention patterns and cognitive processes associated with participants' "I Don't Know" responses during the Sustainability Literacy Test (SLT).

Rather than producing statistically generalizable findings, the study aimed to provide an in-depth understanding of participants' visual behavior and decision-making processes through eye-tracking observations.

The collected data were analyzed using a descriptive approach to identify patterns in participants' responses under conditions of uncertainty. The analysis specifically focused on instances in which participants selected the "I Don't Know" response option. Eye-tracking metrics, including visual scanpaths and gaze behavior, were examined to investigate participants' visual attention and decision-making processes during these responses. The observed eye movement patterns were subsequently interpreted as indicators of cognitive load associated with uncertainty when answering Sustainability Literacy Test questions.

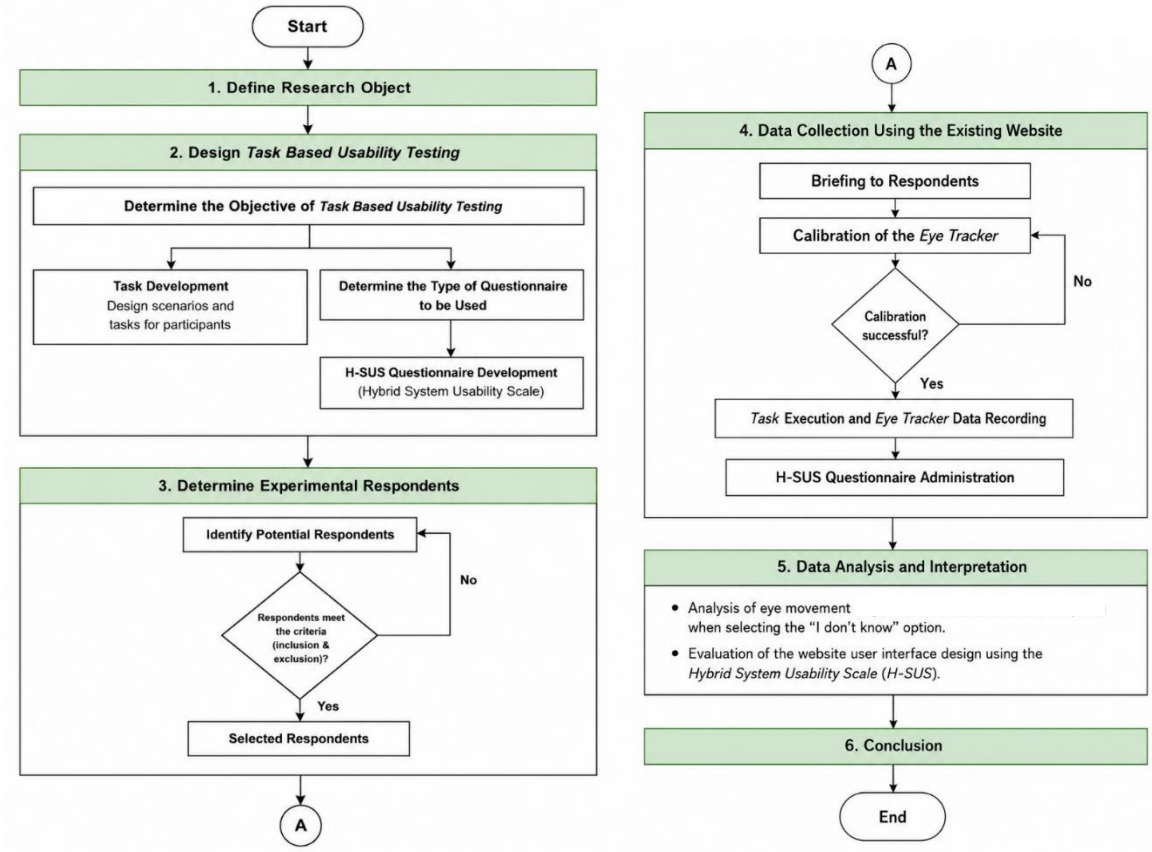

Figure 2. Research Methodology Flowchart

A purposive sample of five faculty members from Institut Teknologi Sepuluh Nopember (ITS) participated in this exploratory qualitative eye-tracking study. The study combined eye-tracking observations with website usability evaluation using the Hybrid System Usability Scale (H-SUS). Following Nielsen's recommendation [10], five participants are generally sufficient to identify most usability issues. In this study, Nielsen's recommendation supports the usability evaluation, while the eye-tracking analysis aims to explore visual attention and cognitive processes rather than produce statistically generalizable findings. All participants had prior experience

with multiple-choice assessments and received a small token of appreciation after completing the experiment.

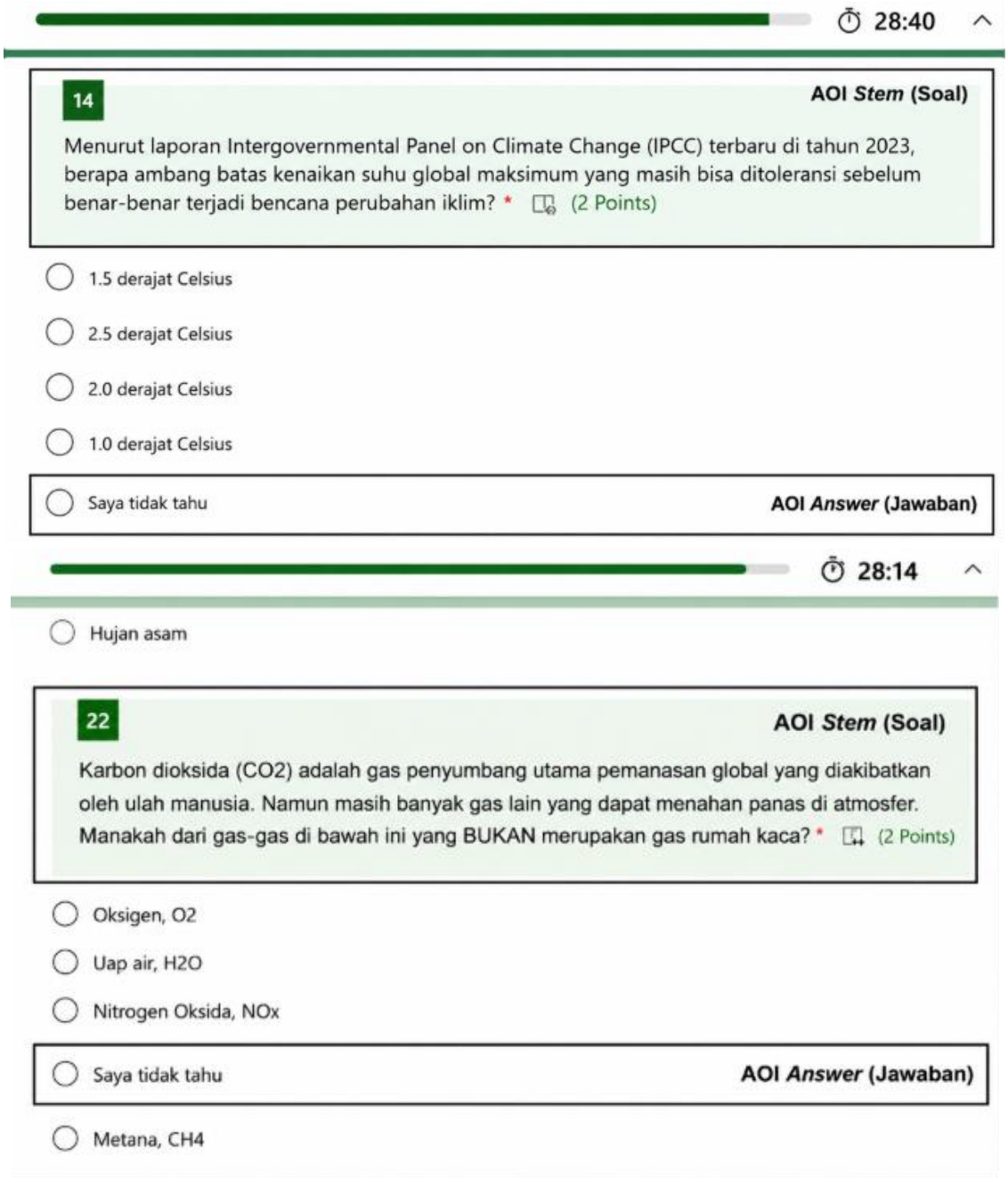

Figure 2. Area of Interest (AOI) on the Question and Answer

Faculty members were selected because they are expected to have greater familiarity with sustainability concepts than students, providing a meaningful perspective on cognitive behavior during the Sustainability Literacy Test (SLT).

The experiment was conducted individually in the EPSK Laboratory. After calibration of the Gazepoint HD3 eye tracker, participants completed a 30-question Sustainability Literacy Test (SLT). Each session lasted approximately 45 minutes. Eye movement data were recorded and analyzed using Gazepoint Analysis software.

For each question, two Areas of Interest (AOIs) were defined: (1) the question area and (2) the answer option area (A–E). To reduce positional bias, the answer options were randomized so that the "I Don't Know" option did not consistently appear in the same position. Figure 3 illustrates the multiple-choice Sustainability Literacy Test used in this study. The assessment was developed collaboratively by the ITS SDGs Center and the Directorate of Global Partnerships (DKG ITS). For each question, two Areas of Interest (AOIs) were defined: (1) the question area and (2) the answer option area (A–E). To minimize positional bias, the locations of answer options (A–E) were randomized for every question, ensuring that the "I Don't Know" option did not consistently appear in the same position, as is commonly observed in conventional paper-based tests.

This study has several limitations. First, the eye-tracking experiment involved only five participants; therefore, the findings should be interpreted as exploratory rather than statistically generalizable. Future studies with larger and more diverse samples are recommended to validate the observed cognitive patterns.

3.Results and Discussion

3.1 Participants' "I Don't Know" Responses

The completion time of the five participants was first calculated to provide an overview of their performance during the Sustainability Literacy Test (SLT). Participants completed the assessment within 560.967–1,096.900 seconds, with an average completion time of 706.398 seconds. The relatively long completion time indicates that several questions required substantial cognitive effort before participants reached a final decision. Following the assessment, participants' "I Don't Know" responses were identified and are presented in **Table 1**.

Table 1. Participants' "I Don't Know" responses in the sustainability literacy test

Quastion No.	D1	D2	D3	D4	D5
14	1	1	1		
22	1				
27	1	1			
29	1		1		1
33		1			
35	1				
36	1	1			1

Note: Di denotes the ith faculty member (e.g., D1 = Faculty Member 1).

Table 1 shows that Questions 14, 22, 27, 29, 33, 35, and 36 received "I Don't Know" responses from at least one participant. Among these questions, Question 14 generated the highest number of uncertain responses, followed by Questions 27, 29, and 36. Because the Sustainability Literacy Test explicitly provided an "I Don't Know" option in addition to the five multiple-choice alternatives, participants were able to express uncertainty rather than relying on guessing. The concentration of

uncertain responses on specific questions suggests that participants' decisions were influenced not only by their level of sustainability knowledge but also by characteristics of the assessment items. Therefore, eye-tracking analysis was conducted to investigate participants' visual behavior while processing these questions.

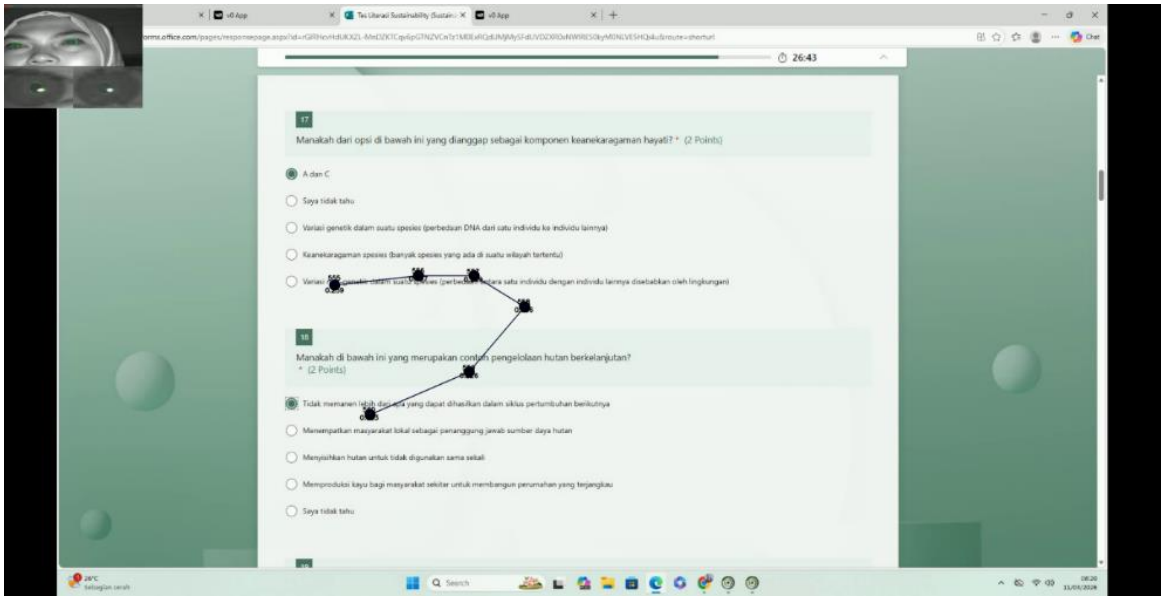

Figure 3. Scanpaths of Participants During the Sustainability Literacy Test (1)

3.2 Scanpath Analysis of Visual Behaviour

Human visual behavior was analyzed using scanpath analysis, which illustrates the sequence of gaze movements while participants interacted with the Sustainability Literacy Test. The scanpaths obtained during the experiment are presented in **Figure 3** and **4**.

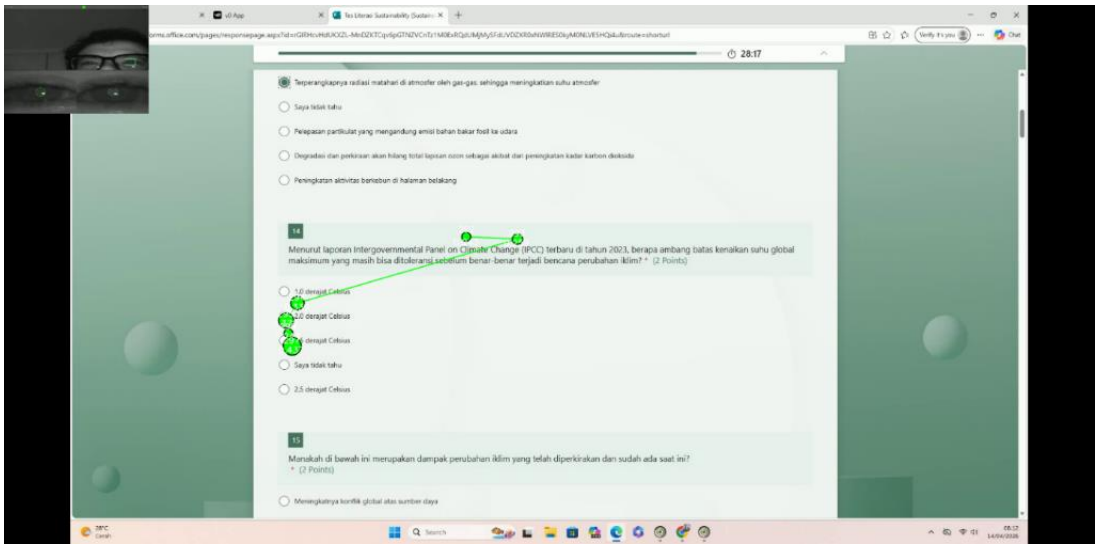

Figure 4. Scanpaths of Participants During the Sustainability Literacy Test (2)

Figure 3 and **4** show that participants generally followed a systematic reading strategy characterized by horizontal eye movements while reading the question text

and vertical eye movements when evaluating the answer options. However, participants who ultimately selected the "I Don't Know" option exhibited repeated gaze transitions between the question and the available answer options before making a final decision. These repeated gaze transitions indicate that participants repeatedly revisited the presented information rather than following a straightforward reading sequence, suggesting greater cognitive effort during information processing.

These findings are consistent with those reported by Yaneva et al. [3], who demonstrated that scanpath analysis can identify information-processing difficulty through repeated visual inspections and gaze regressions. Similarly, the scanpath observations in the present study indicate that participants devoted considerable effort to understanding the linguistic structure of complex questions before selecting a response. Consequently, uncertainty reflected by the "I Don't Know" response appears to be associated not only with limited sustainability knowledge but also with the readability and complexity of the assessment items.

3.3 Readability Barriers Identified from Scanpath Analysis

To identify the factors contributing to participants' uncertainty, the questions associated with "I Don't Know" responses were further examined. The identified readability barriers are summarized in **Table 2**.

Table 2. Types of readability barriers identified in the sustainability literacy test

Type of Readability Barrier	Question No.	Question Example	Identified Issue
Long and Complex Sentence Structure	14	*According to the latest 2023 Intergovernmental Panel on Climate Change (IPCC) report, what is the maximum global temperature increase that can still be tolerated before catastrophic climate change occurs?*	*The question contains lengthy introductory information, multiple embedded phrases, and a cause-and-effect relationship, increasing the reading burden and cognitive processing demands.*
	22	*Carbon dioxide (CO_2) is the primary greenhouse gas emitted through human activities. However, several other gases also trap heat in the atmosphere. Which of the following is **NOT** a greenhouse gas?*	*A long introductory statement precedes the core question, requiring participants to process multiple pieces of information before identifying the answer.*
Negative Wording	22	*Which of the following is **NOT** a greenhouse gas?*	*The negative keyword (**NOT**) requires participants to reverse their reasoning process, thereby increasing the likelihood of misinterpretation.*
	33	*Which of the following statements about indigenous communities and sustainability is **incorrect**?*	*The word **incorrect** requires participants to identify the false statement rather than the correct one, increasing cognitive complexity during decision-making.*

Table 2 identifies three major readability barriers: (1) long and complex sentence structures, (2) negative wording, and (3) lengthy answer options. Questions 14 and 22 contained lengthy introductory statements and multiple embedded clauses that required participants to process several pieces of information before reaching the core question. In addition, Questions 22 and 33 used negative wording such as NOT and incorrect, requiring participants to mentally reverse their reasoning process before selecting an answer.

These readability characteristics increase the cognitive resources required to interpret the meaning of the questions before evaluating their conceptual content [13]. Negative wording has also been shown to increase cognitive effort and the likelihood of misinterpretation because respondents must reverse their reasoning process [14]. Although the Microsoft Forms interface provided an orderly presentation, the textual complexity of the assessment items remained the primary source of cognitive difficulty. The scanpath observations support this interpretation by showing repeated gaze transitions between the question and answer options, indicating that participants repeatedly evaluated the available information before making a decision.

3.4 Completion Time Analysis

The relationship between participants' uncertainty and the time required to complete each question is presented in Table 3.

Table 3. Time required by participants to complete questions with "i don't know" responses

Question No.	D1	D2	D3	D4	D5
14	1.79261	0.92381	0.46218	–	–
22	0.66382	–	–	–	–
27	1.4751	1.7651	–	–	–
29	0.8314	–	0.51348	–	0.73912
33	–	0.92381	–	–	–
35	0.89151	–	–	–	–
36	0.94031	1.56793	–	–	0.84968

Note: Di denotes the ith faculty participant. Values represent the time required (minutes) to complete each question.

Source: Authors' analysis based on observation, interview, and maintenance records (2025).

Table 3 shows that participants generally required more than 60 seconds to complete Questions 14, 22, 27, 29, 33, 35, and 36 before selecting the "I Don't Know" response. In contrast, correctly answered questions generally required less than 30 seconds, indicating more efficient information processing and greater confidence during decision-making.

When interpreted together with the scanpath observations, the longer completion times suggest that participants repeatedly revisited both the questions and the available answer options before reaching a final decision. These findings support those of Yaneva et al. [4], who reported that repeated visual inspections reflect increased cognitive effort when respondents encounter difficult assessment items. Similarly, Maddox et al. [6] found that respondents generally process the question before evaluating the available alternatives. Unlike previous studies that focused primarily on incorrect responses, the present study specifically investigates "I Don't Know" responses, indicating that uncertainty results from an interaction between knowledge limitations, question readability, and cognitive load rather than from insufficient sustainability literacy alone.

3.5 Website Usability Evaluation Using H-SUS

Participants' subjective evaluation of the Sustainability Literacy Test website was measured using the Hybrid System Usability Scale (H-SUS). The results are presented in **Table 4**.

Table 4. H-SUS scores of the existing website

Participant	H-SUS Score
D1	48.34
D2	51.68
D3	61.68
D4	83.35
D5	71.68
Average H-SUS Score	**63.34**

Note: Di denotes the ith faculty participant.

The website achieved an average H-SUS score of 63.34, corresponding to an Acceptable usability rating, Grade C, and an adjective rating of "OK". These findings indicate that the website interface was generally usable despite several opportunities for improvement. Participants mainly associated usability difficulties with the presentation of the assessment questions rather than the website interface itself. Long question statements, dense textual information, and variations in question presentation increased the cognitive effort required to complete the assessment. Interestingly, participants D4 and D5 assigned relatively high H-SUS scores despite requiring longer completion times, indicating that the additional time reflected the cognitive demands of the assessment content rather than deficiencies in interface usability.

3.6 Integration of Eye-Tracking and H-SUS Findings

The integration of scanpath observations, completion time, H-SUS evaluation, and participant feedback is summarized in **Table 5**. **Table 5** demonstrates that all three evaluation methods consistently identified the same usability issues, namely lengthy question text, complex sentence structures, and the absence of visual

emphasis on important keywords. While the H-SUS results indicate that the website interface was generally acceptable, scanpath observations and completion time reveal that participants experienced increased cognitive effort when processing lengthy and linguistically complex questions. These findings suggest that the primary usability challenge lies in the presentation of the Sustainability Literacy Test content rather than the website interface itself. Consequently, simplifying sentence structures, reducing unnecessary negative wording, and emphasizing important keywords are expected to reduce cognitive load while improving both usability and the validity of sustainability literacy assessment.

Table 5. Usability issues identified in the existing website design

Source of Evidence	Identified Usability Issues
Scanpath Analysis	1. Irregular scanpaths characterized by repeated gaze regressions and prolonged fixations (> 1 minute) when reading lengthy question texts. 2. High cognitive workload caused by long and linguistically complex question statements and answer options.
H-SUS Evaluation	1. Overall usability was rated as "**OK**", although users reported difficulties associated with the Sustainability Literacy Test items. 2. The website was generally acceptable, but the presentation of lengthy textual questions negatively affected the user experience.
Participant Feedback	1. Question statements were excessively long. 2. Important keywords within the questions were not visually emphasized, making them difficult to identify quickly.

4. Conclusion

This study explored the cognitive processes underlying "I Don't Know" responses in the Sustainability Literacy Test (SLT) by integrating scanpath analysis, completion time, and Hybrid System Usability Scale (H-SUS) evaluation. The findings indicate that participants who selected the "I Don't Know" option generally required longer completion times and exhibited repeated gaze transitions between the question and answer options. These behaviors suggest that uncertainty was influenced not only by participants' sustainability knowledge but also by the readability and linguistic complexity of the assessment items.

The main contribution of this study is the integration of objective eye-tracking observations with subjective usability evaluation to provide an exploratory understanding of why participants choose the "I Don't Know" response during sustainability assessment. The findings indicate that scanpath analysis, completion time, and H-SUS complement one another in identifying cognitive difficulties associated with question design, providing practical insights for improving the readability and usability of sustainability literacy assessments.

This study has several limitations. First, the eye-tracking experiment involved only five participants; therefore, the findings should be interpreted as exploratory and should not be generalized to broader populations. Second, the eye-tracking analysis focused on scanpath visualization and completion time, without

incorporating additional quantitative metrics such as fixation duration, fixation count, time to first fixation, or AOI-based statistics. Future research should involve larger and more diverse participant groups and include these quantitative eye-tracking measures to further validate the observed cognitive patterns.

Acknowledgement

The authors would like to thank the Institut Teknologi Sepuluh Nopember (ITS) for providing the facilities required to conduct this research. The authors also sincerely appreciate the faculty members who participated in the eye-tracking experiment and usability evaluation for their valuable time and contributions.

Authors' Declaration

Authors' contributions and responsibilities - The authors made substantial contributions to the conception and design of the study. The authors took responsibility for data analysis, interpretation, and discussion of results. The authors read and approved the final manuscript.

Funding - No funding information from the authors.

Availability of data and materials - All data is available from the authors.

Competing interests - The authors declare no competing interest.

Additional information - No additional information from the authors.

References

[1] V. Ardaniah, Z. D. Billah, L. E. Rahmantyo, and P. Ardhianto, "Embedding Sustainability Literacy with Language Teaching in Higher Education," JL3T (Journal of Linguistics, Literature and Language Teaching), vol. 11, no. 1, pp. 130–148, Jul. 2025, doi: 10.32505/jl3t.v11i1.11027.

[2] F. Anastasya, "Pengembangan Sistem Informasi Tes Literasi SDGs ITS Berbasis Web sebagai Media Pengukuran Literasi Mahasiswa Mengenai Pembangunan Berkelanjutan," Institut Teknologi Sepuluh Nopember, 2026.

[3] V. Yaneva, B. E. Clauser, A. Morales, and M. Paniagua, "Using Eye-Tracking Data as Part of the Validity Argument for Multiple-Choice Questions: A Demonstration," J. Educ. Meas., vol. 58, no. 4, pp. 515–537, Dec. 2021, doi: 10.1111/jedm.12304.

[4] V. Yaneva, B. E. Clauser, A. Morales, and M. Paniagua, "Assessing the validity of test scores using response process data from an eye-tracking study: a new approach," Advances in Health Sciences Education, vol. 27, no. 5, pp. 1401–1422, Dec. 2022, doi: 10.1007/s10459-022-10107-9.

[5] J. Bartolomé, P. Garaizar, and L. Bastida, "Validating item response processes in digital competence assessment through eye-tracking techniques," in ACM International Conference Proceeding Series, Association for Computing Machinery, Oct. 2020, pp. 738–746. doi: 10.1145/3434780.3436641.

[6] B. Maddox, A. P. Bayliss, P. Fleming, P. E. Engelhardt, S. G. Edwards, and F.

Borgonovi, "Observing response processes with eye tracking in international large-scale assessments: evidence from the OECD PIAAC assessment," European Journal of Psychology of Education, vol. 33, no. 3, pp. 543–558, Jul. 2018, doi: 10.1007/s10212-018-0380-2.

[7] M. G. Carneiro Queiroz et al., "Comparison of cognitive workload between very short answer questions and multiple-choice questions: an eye-tracking experiment," Med. Educ. Online, vol. 31, no. 1, 2026, doi: 10.1080/10872981.2026.2621434.

[8] L. Hahn and P. Klein, "Clustering eye-movement data uncovers students' strategies for coordinating equations and diagrams of vector fields," Educational Studies in Mathematics, vol. 118, no. 3, pp. 359–385, Mar. 2025, doi: 10.1007/s10649-023-10243-y.

[9] Y. Hu, B. Wu, and X. Gu, "An Eye Tracking Study of High-and Low-Performing Students in Solving Interactive and Analytical Problems," 2017.

[10] J. Nielsen, "How Many Test Users in a Usability Study?," Nielsen Norman Group (NN/Group). Accessed: Apr. 27, 2026. [Online]. Available: https://www.nngroup.com/articles/how-many-test-users/

[11] A. Lopez-Cardona, P. Emami, S. Idesis, S. Duraisamy, L. A. Leiva, and I. Arapakis, "A Comparative Study of Scanpath Models in Graph-Based Visualization," in Eye Tracking Research and Applications Symposium (ETRA), Association for Computing Machinery, May 2025. doi: 10.1145/3715669.3725882.

[12] A. Mohamed Selim, M. Barz, O. S. Bhatti, H. M. T. Alam, and D. Sonntag, "A review of machine learning in scanpath analysis for passive gaze-based interaction," 2024, Frontiers Media SA. doi: 10.3389/frai.2024.1391745.

[13] K. Brosnan, B. Grün, and S. Dolnicar, "Cognitive load reduction strategies in questionnaire design," International Journal of Market Research, vol. 63, no. 2, 2021.

[14] X. Wang and V. Aryadoust, "An eye-tracking and neuroimaging study of negative wording effects on cognitive load in a metacognitive awareness tool," System, vol. 128, Feb. 2025, doi: 10.1016/j.system.2024.103552.

[15] R. Niu, W. Wu, Q. Li, and Y. Fang, "Evaluating the Effect of Font Size and Line Spacing in Chinese Online Education User Interfaces on Children's Learning Experience," Int. J. Hum. Comput. Interact., 2025, doi: 10.1080/10447318.2025.2524495.

[16] J. Baumgartner, N. Ruettgers, A. Hasler, A. Sonderegger, and J. Sauer, "Questionnaire experience and the hybrid System Usability Scale: Using a novel concept to evaluate a new instrument," International Journal of Human Computer Studies, vol. 147, Mar. 2021, doi: 10.1016/j.ijhcs.2020.102575.

[17] J. Baumgartner, N. Frei, M. Kleinke, J. Sauer, and A. Sonderegger, "Pictorial System Usability Scale (P-SUS): Developing an instrument for measuring perceived usability," in Conference on Human Factors in Computing Systems - Proceedings, Association for Computing Machinery, May 2019. doi: 10.1145/3290605.3300299.

[18] J. Brooke, "SUS: A 'Quick and Dirty' Usability Scale," Usability Evaluation in

Industry, 1996.
[19] S. Sugawara, N. Nangia, A. Warstadt, and S. R. Bowman, "What Makes Reading Comprehension Questions Difficult?," in Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Long Papers, 2022, pp. 6951–6971. [Online]. Available: https://github.
[20] S. Gruhn, E. Segers, J. Keuning, and L. Verhoeven, "Understanding variation in children's reading comprehension: A dynamic approach," J. Comput. Assist. Learn., vol. 40, no. 2, pp. 876–901, Apr. 2024, doi: 10.1111/jcal.12923.
[21] E. Z. Nisa', "Analisis Usability dan Perancangan Perbaikan Interface pada Aplikasi E-Health Menggunakan Eye Tracking Technology (Studi Kasus: E-Health Surabaya)," Institut Teknologi Sepuluh Nopember Surabaya, Surabaya, 2023.
[22] R. A. Purnamasari, "Evaluasi Usabilitas dan Perbaikan Desain Website UPT Bahasa dan Budaya ITS Menggunakan Eye Tracking, Retrospective Think Aloud (RTA) dan System Usability Scale (SUS)," Institut Teknologi Sepuluh Nopember Surabaya, Surabaya, 2023.
[23] I. M. Mahendra, "Analisis Dampak Layout, Warna, dan Time Pressure pada Usabilitas Business Intelligence Dashboard menggunakan Eye Tracking, Retrospective Think Aloud (RTA), dan System Usability Scale (SUS)," 2024.
[24] J. Dobres, B. Wolfe, N. Chahine, and B. Reimer, "The effects of visual crowding, text size, and positional uncertainty on text legibility at a glance," Appl. Ergon., vol. 70, pp. 240–246, Jul. 2018, doi: 10.1016/j.apergo.2018.03.007.
[25] M. G. Zolbin and S. Kujala, "Evaluating the Usability of a Patient Portal: Insights from Older Migrant Adults," Stud. Health Technol. Inform., vol. 336, pp. 2344–2348, May 2026, doi: 10.3233/SHTI260689.