

The influence of assessment types on students' performance in Machine Learning Education

Madeline El Aissati¹ Supervisor: Gosia Migut¹ Examiner: Mark Neerincx
¹EEMCS, Delft University of Technology

Introduction

- **Motivation**
Growing importance of Machine Learning (ML) → education on ML needs to be improved. Assessment plays a crucial role in improving education and understanding how students learn [1].
- **Research gap**
In [1] the relation between assessment and student performance is investigated, but no pre-test was conducted and assessments were not motivated. Frequent neglect of assessment is highlighted in [2], especially in Engineering education, while it greatly impacts how students learn.
- **Aim**
Investigating the relation between assessment types and students' learning gain in ML education.

Research question

What is the influence of **different assessment types** on students' learning gain in **k-means clustering**?

Methodology

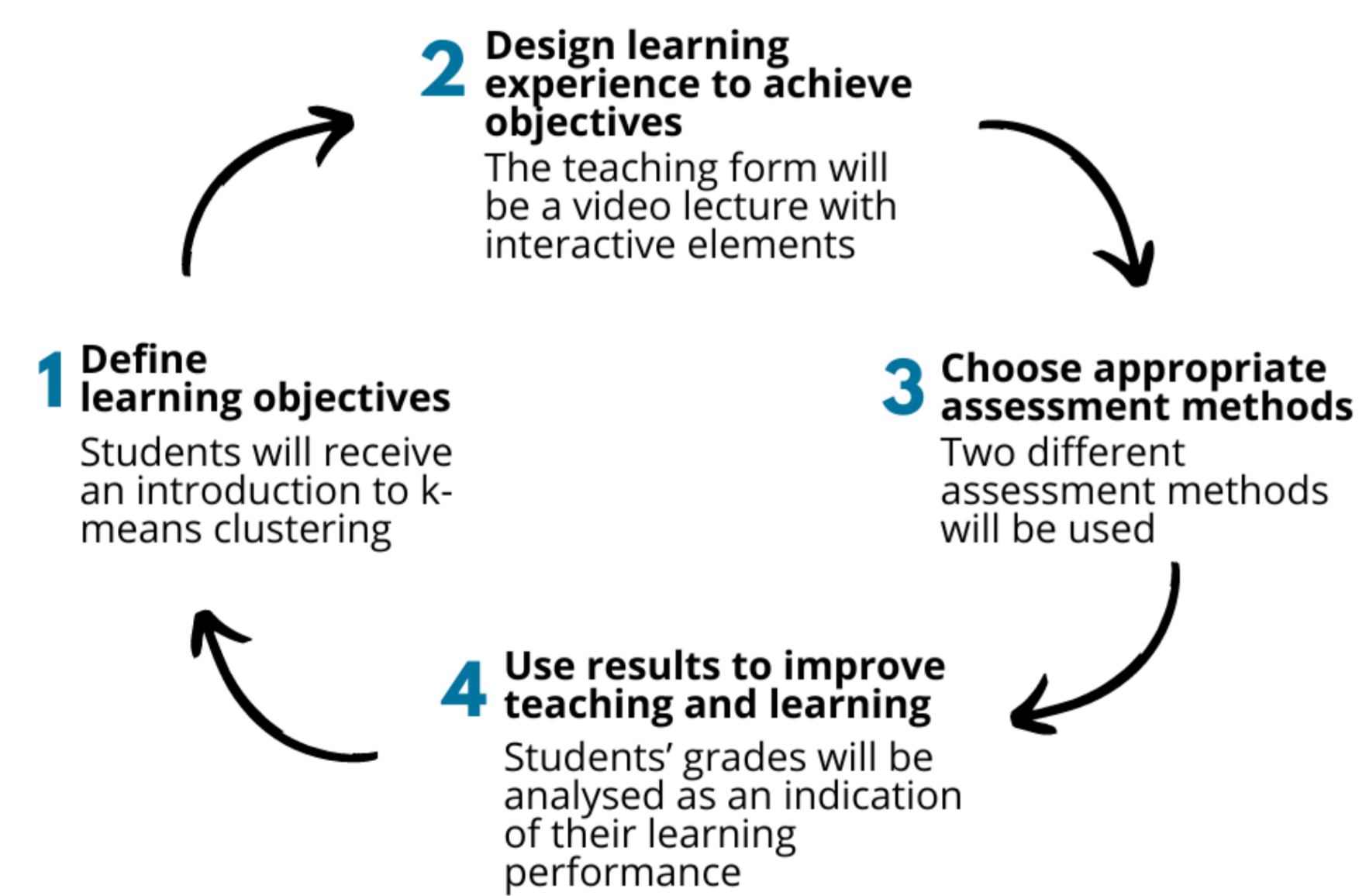


Figure 1. Suskie's Teaching-Learning-Assessment Cycle, adapted from [3]

- **Measuring learning performance** [4]

$$\text{Gain} = \frac{\text{Post}_i - \text{Pre}_i}{7 - \text{Pre}_i} * 100\% \quad (1)$$

$$\text{Average of Gains} = \frac{\sum_{i=1}^n \text{Gain}_i}{n} \quad (2)$$

Experiment design

- Pre- and post-test consisted of a **concept inventory (CI)** on k-means clustering
- Assessment methods based on a survey among students [5]

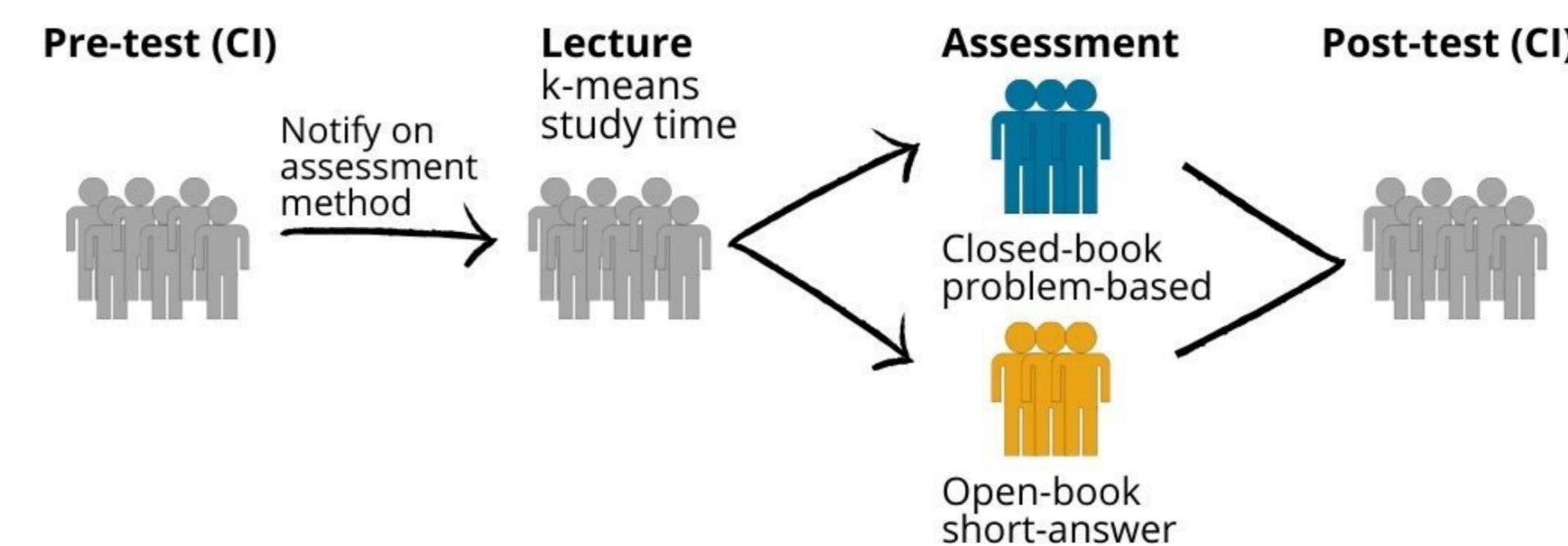


Figure 2. Experiment set-up

Results

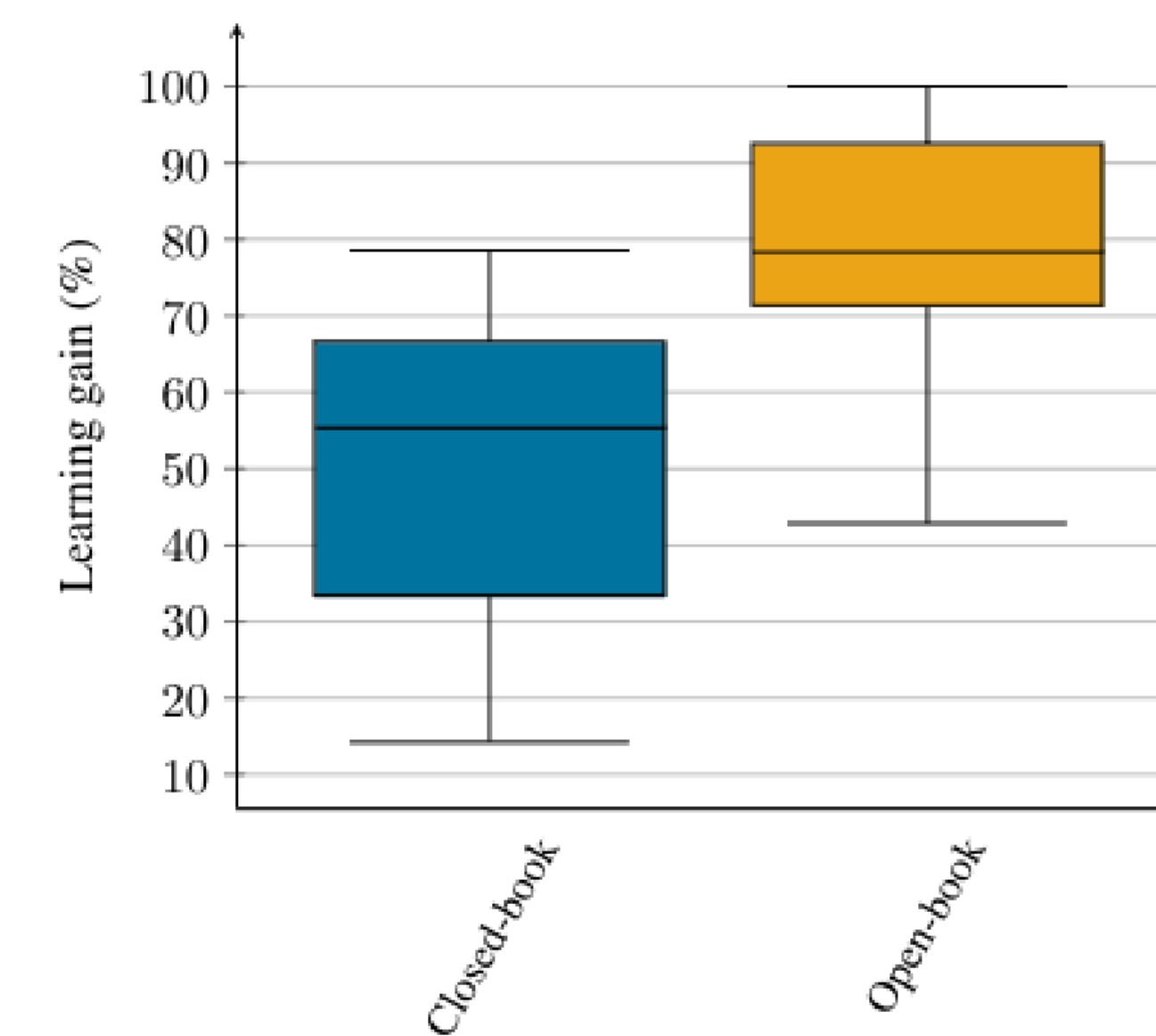


Figure 3. Learning gain for closed- and open-book assessment groups

- Results show a **significantly higher learning gain** for participants who took the open-book assessment ($t(20) = -3.27$; $p = 0.004$) (see Figure 3)
- **Most common mistake** in pre-test was question 3: out of the 10 participants who attempted to answer the question, 7 believed clustering to be a supervised learning method.
- In the post-test, all of the open-book participants were able to conduct a **full iteration of k-means clustering**, whereas only 50% of the closed-book participants managed to do this (see Figures 4 and 5)

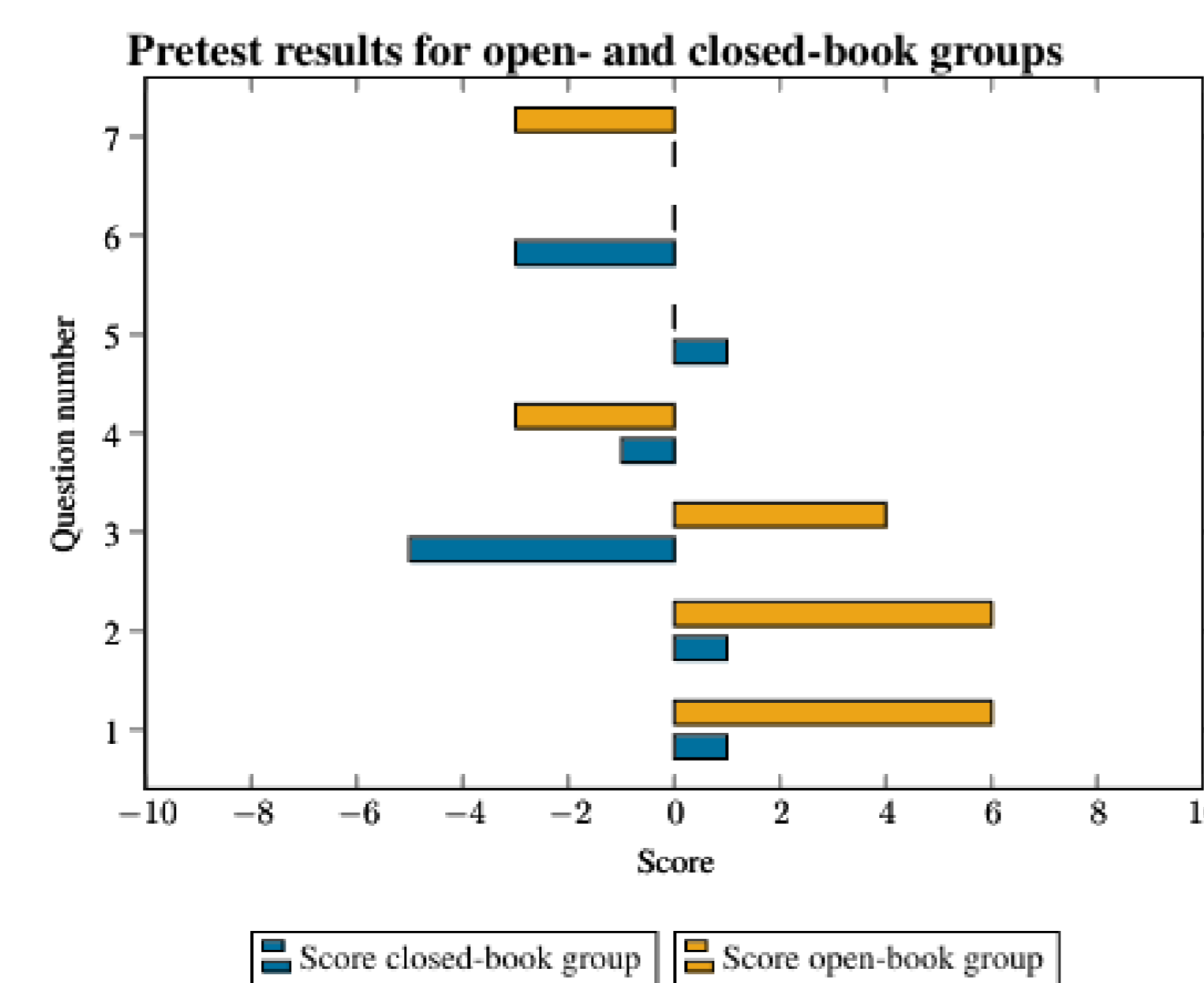


Figure 4. Results of the pre-test per question

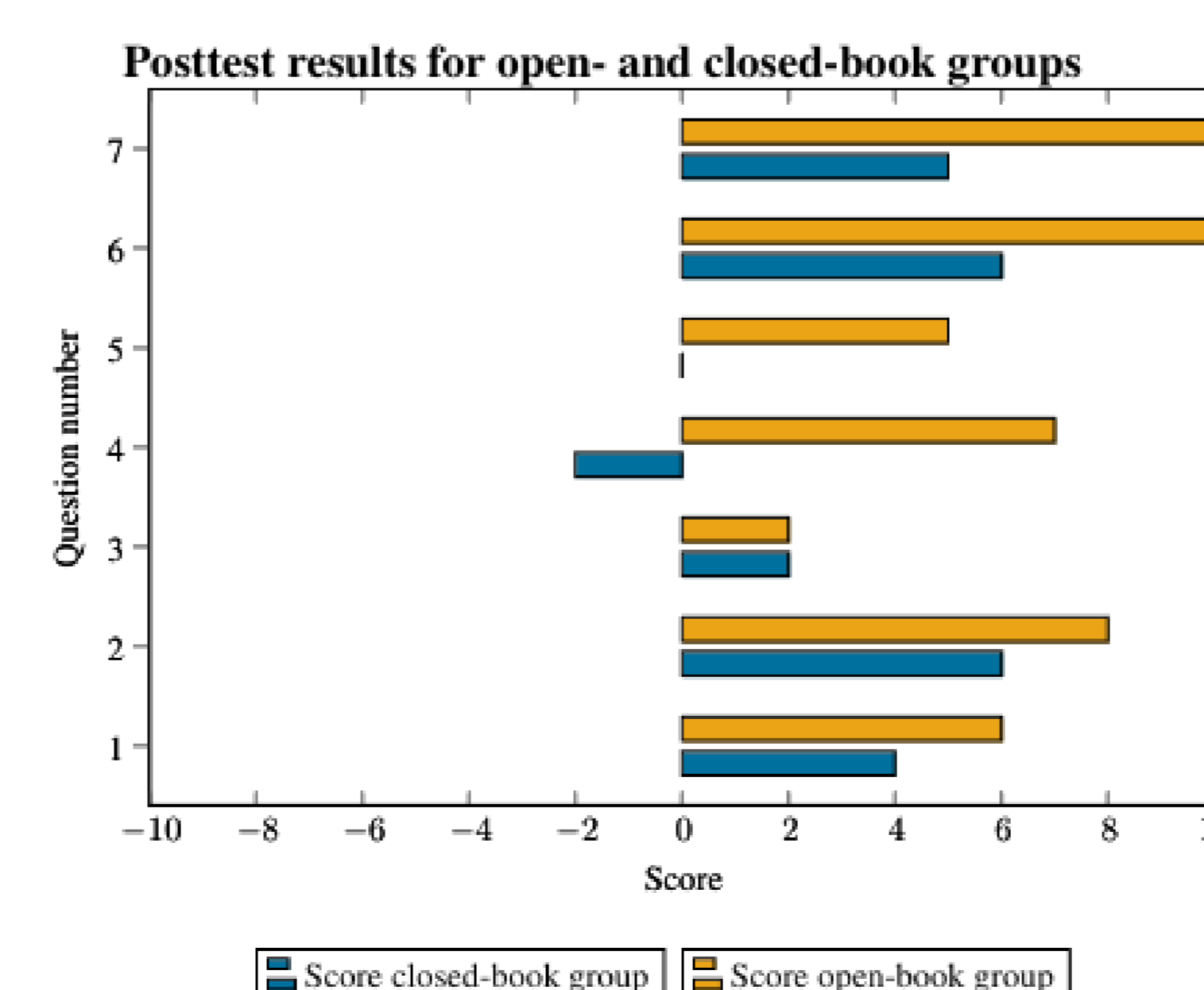


Figure 5. Results of the post-test per question

Research Ethics

The five principles of Research Integrity [6]:

- **Honesty:** full report of results and limitations
- **Scrupulousness & Independence:** personal bias was minimized by discussing with fellow students and experts on k-means clustering
- **Transparency:** study set-up was motivated, removed data was mentioned
- **Responsibility:** benefits of research on assessment ensured relevance of the study

Discussion & limitations

- Some participants had **trouble concentrating** due to noisy locations
- Some participants were able to **communicate with peers** and may have gained advantage
- Some participants gained **more information** than others through questions
- Closed-book group had more trouble understanding **centroids**, a possible explanation for the **experienced difficulty** in conducting a **full iteration of k-means**

Future work

- Investigate **long-term** effects of instruction and assessment methods
- Develop **validated assessments** for Machine Learning topics

Takeaways

- Research on the relation between assessment and student performance could **enhance student learning**
- An open-book short-answer exam is favourable over a closed-book problem-based assignment, as it resulted in **significantly higher learning gain**
- **Concept Inventories** are not widely applied yet in Computer Science but could provide a standardized assessment format for specific topics

References

- [1] K. Islam, P. Ahmadi, and S. Yousaf, "Assessment Formats and Student Learning Performance: What is the Relation?,"
- [2] S. Jimaa, "The impact of assessment on students learning," *Procedia - Social and Behavioral Sciences*, vol. 28, pp. 718-721, 2011.
- [3] L. Suskie, "UNDERSTANDING THE NATURE AND PURPOSE OF ASSESSMENT," in *Designing Better Engineering Education Through Assessment: A Practical Resource for Faculty and Department Chairs on Using Assessment and ABET Criteria to Improve Student Learning*, pp. 3-22, 2023.
- [4] R. R. Hake, "Evaluating conceptual gains in mechanics: A six thousand student survey of test data," in *The Changing Role of Physics Departments in Modern Universities: Proceedings of ICUPE*, vol. 399, (College Park, Maryland (USA)), pp. 595-604, AIP Conference Proceedings, AIP, March 1997.
- [5] A. Karim, S. Fawzia, and M. M. Islam, "Factors affecting deep learning of engineering students," in *Proceedings of the 26th Annual Conference of the Australasian Association for Engineering Education (AAEE2015)*, pp. 1-8, Australia: School of Engineering, Deakin University, 2015.
- [6] KNAW, NFU, NWO, TO2-Federatie, Vereniging Hogescholen, and VSNU, "Nederlandse gedragscode wetenschappelijke integriteit," 2018.