

Moodle quizzes as a continuous assessment in Higher Education: an exploratory approach in Physical Chemistry.

Isabel López-Tocón ^{1,*}

¹ Departamento de Química Física. Facultad de Ciencias. Campus de Excelencia Andalucía Tech, Universidad de Málaga, Spain, E-29071 Málaga.

* Correspondence: tocon@uma.es

Statistical Parameters

Psychometric analysis applies statistical principles to determine the suitability of the proposed items based on the responses and their individual relationship with the rest of the answers, thereby detecting whether the proposed items are appropriate to assess the level of knowledge, degree of facility, and degree of discrimination between high and low conceptual skills [1–4].

The statistical parameters collected in the tables were calculated on the basis of the classical test theory and those are directly reported by Moodle platform (https://docs.moodle.org/dev/Quiz_statistics_calculations) in the statistics report section. Equations to represent the Facility Index (FI) and Discrimination Index (DI) adapted from Moodle Statistics are indicated below.

$$FI(\%) = \frac{\bar{x}_p - x_p(\min)}{x_p(\max) - x_p(\min)} 100$$

Facility Index (FI; % correct), which is defined as the average score on the item and it means how easy or difficult an item is, with respect to the rest of the other items inside the quiz. In the Moodle statistics $x_p(\min)$ is always zero, so the above equation is simplified to:

$$FI(\%) = \frac{\bar{x}_p}{x_p(\max)} 100$$

where \bar{x}_p is the mean value from all values obtained for the total users who did every item and $x_p(\max)$ is the maximum value obtained for that item. If the questions could be distributed as correct/incorrect, this parameter would coincide with the percentage of students who responded to the items correctly. The higher the facility index, the easier the question is (for this group of students).

The Discrimination Index (DI) is a coefficient of correlation among the scores of each particular item with respect to the complete quiz. Its mathematical expression is:

$$DI(\%) = 100 \frac{C(x_p, X_p)}{\sqrt{V(x_p)V(X_p)}}$$

where

Citation: López-Tocón I. Moodle quizzes as a continuous assessment in Higher Education: an exploratory approach in Physical Chemistry. *Educ. Sci.* **2021**, *11*, 500. <https://doi.org/10.3390/educsci11090500>

Academic Editor: Diego Vergara, Ana Teresa Ferreira-Oliveira, Sandra Raquel Gonçalves and Marta Abelha

Received: 22 June 2021

Accepted: 30 August 2021

Published: 3 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<http://creativecommons.org/licenses/by/4.0/>).

$$C(x_p, X_p) = \frac{1}{S-1} \sum_{s \in S} (x_p(s) - \bar{x}_p)(X_p(s) - \bar{X}_p)$$

$$V(x_p) = \frac{1}{S-1} \sum_{s \in S} (x_p(s) - \bar{x}_p)^2$$

The lot of students $s \in S$ who have completed at least one attempt on the quiz. The quiz has a number position of position $p \in P$

The weakness of this statistic is that, unless the facility index is 50%, it is impossible for the discrimination index to be 100%, or, to put it another way, if FI is close to 0% or 100%, DI will always be very small. That makes interpreting this statistic difficult.

DI can obtain negative values. Positive values indicate items that discriminate right questions, while those negative values are items that are answered by low-performing students. This means that items with a negative DI are answered incorrectly by students, which penalizes the majority of students.

The Discriminative Efficiency (DE) gets around that weakness in the DI by expressing $C(x_p, X_p)$ as a percentage of the maximum value it could have taken given the scores the students got on this question, and the test as a whole.

$$DE(\%) = 100 \frac{C(x_p, X_p)}{C_{max}(x_p, X_p)}$$

The reference values proposed by the Moodle platform for these indicators in a summative quiz are: for FI, a value between 35-65% is considered correct for the average student. Higher/lower values indicate that the difficulty of the question decreases/increases, respectively. A value between 30-50% is considered adequate for the discrimination index. Upper/lower values indicate good/bad discrimination, respectively. However, these values may vary depending on the data obtained and the aim of the test [5]. Some authors propose a value higher than 20% as a good result in both FI and DI index [6].

References

1. Heck, A.; Van Gastel, L. Mathematics on the threshold. *International Journal of Mathematical Education in Science and Technology*, 2006, 37, 925-945. DOI: <https://doi.org/10.1080/00207390600819003>
2. Revuelta, J.; Ximénez, M. C.; Olea, J. Psychometric and psychological effects of item selection and review on computerized testing. *Educational and Psychological Measurement*, 2003, 63, 791-808. DOI: <https://doi.org/10.1177/0013164403251282>.
3. Gamage, S.H.; Ayres, J.R.; Behrend, M.B.; Smith, E.J. Optimising Moodle quizzes for online assessment. *International Journal of STEM Education*, 2019, 6, 1-14. DOI: <https://doi.org/10.1186/s40594-019-0181-4>.
4. Gómez-Soberón, J.; Gómez-Soberón, M.C.; Corral-Higuera, R.; Arredonde-Rea, S.P.; Almaral-Sánchez, J.L.; Cabrera-Cavarubias, F.G. Calibrating Questionnaires by Psychometric Analysis to Evaluate Knowledge. SAGE, 2013 Open. DOI: [10.1177/2158244013499159](https://doi.org/10.1177/2158244013499159).
5. Lord, F.; Novick, M. *Statistical theories of mental test scores*, 1968, Editorial Addison-Wesley, USA.
6. Guilford, J.P. *Psychometric methods*, 1975, Bombay, Neva Delhi. Editorial Tata McGraw-Hill.

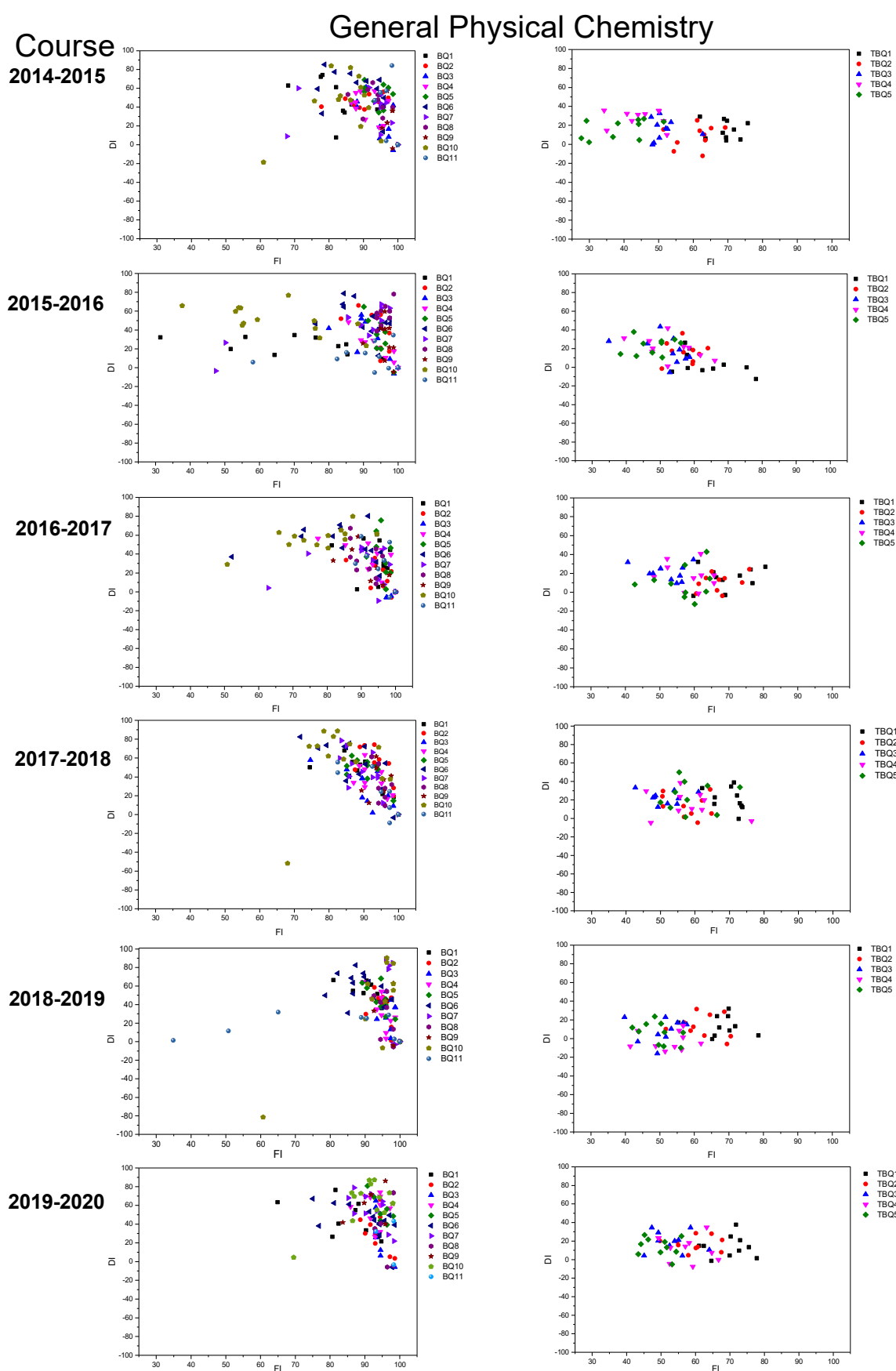


Figure S1. FI-DI diagrams corresponding to the two types of quizzes. Scales axis are in percentage. Own elaboration based on this study.

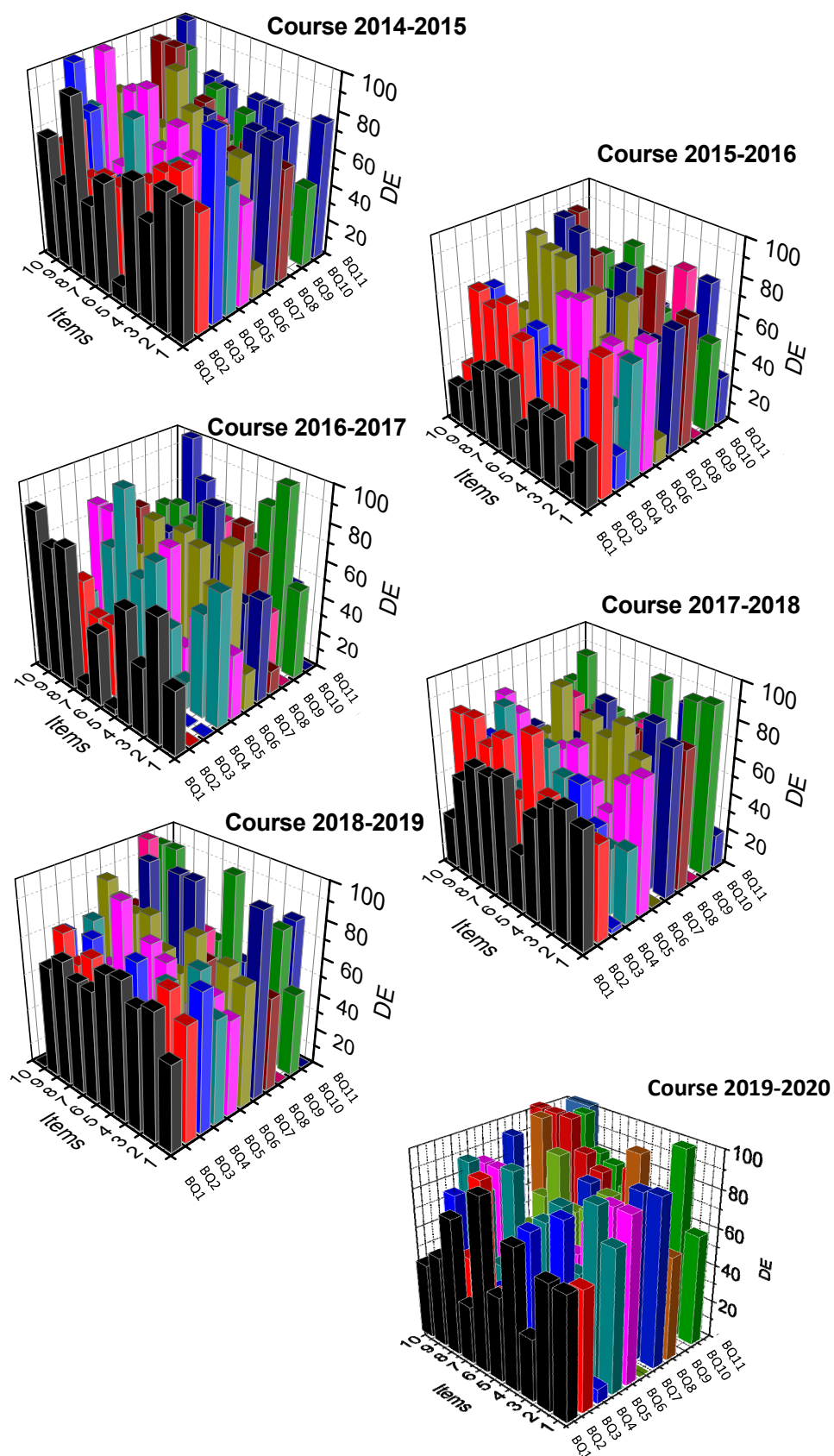


Figure S2. DE diagrams corresponding to the ten items (true/false) of each BQ. Own elaboration based on this study.

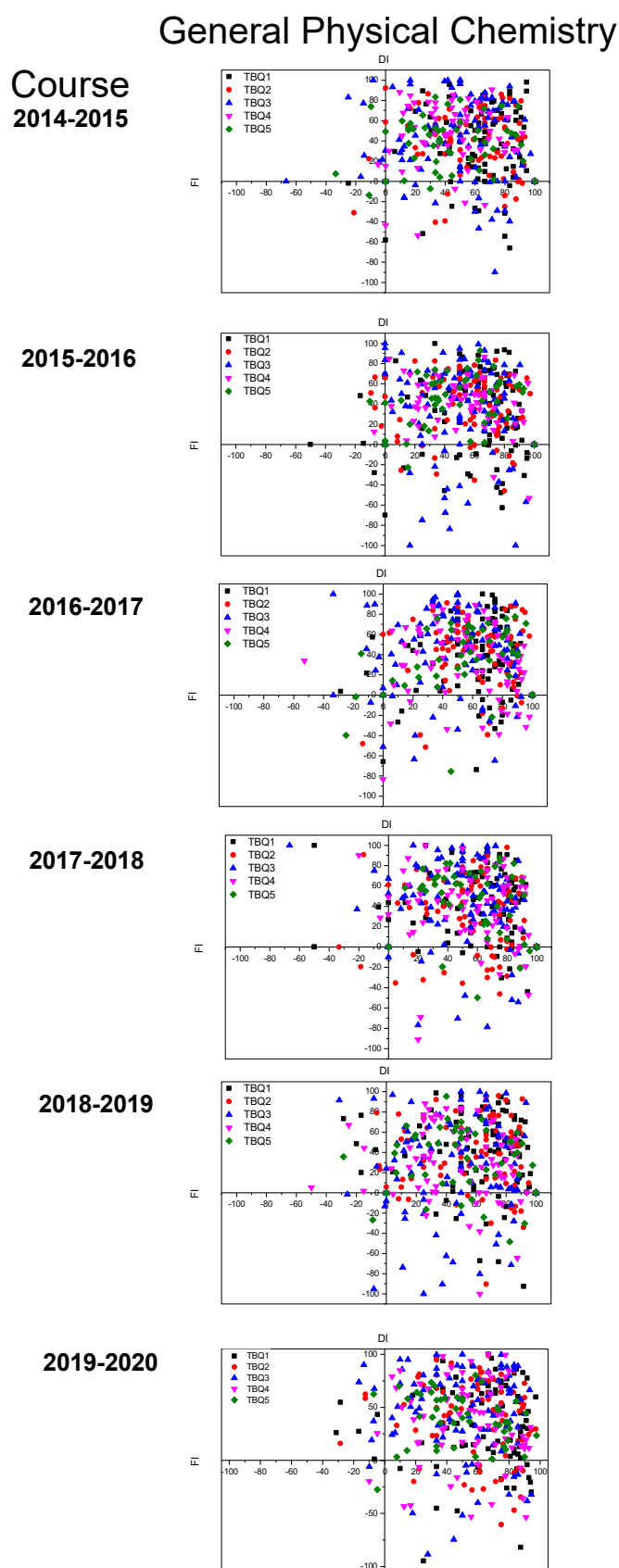


Figure S3. FI-DI diagrams of the different items in each TBQ. Own elaboration based on this study.