



# Administration of Formative Assessments in a Proctored Setting is Associated with Improved Summative Assessment Performance

Yerko Berrocal<sup>1</sup>, Andrew Darr<sup>1</sup>, Jonathan Fisher<sup>1</sup>, Jenna Regan<sup>1</sup>, Amy Lin<sup>2</sup>, Richard Tapping<sup>1</sup>

<sup>1</sup>Department of Health Sciences Education, University of Illinois College of Medicine Peoria, IL 61605

<sup>2</sup>Department of Pathology, University of Illinois College of Medicine Chicago, IL 61612

## Aim

Despite strong evidence supporting the use of formative exams in undergraduate medical education, the potential impact of the setting in which they are administered has not been previously explored.

The goal of this study was to examine whether administering formative assessment in a proctored versus non-proctored setting using multiple choice questions (MCQs) influences student performance on final summative assessment.

## Background on Assessments

- Block 1 (6 weeks of class) includes both non-proctored weekly quizzes and proctored team-based learning (TBL) individual readiness assurance test (IRAT) assessments.
- Assessments are administered online using proprietary software.

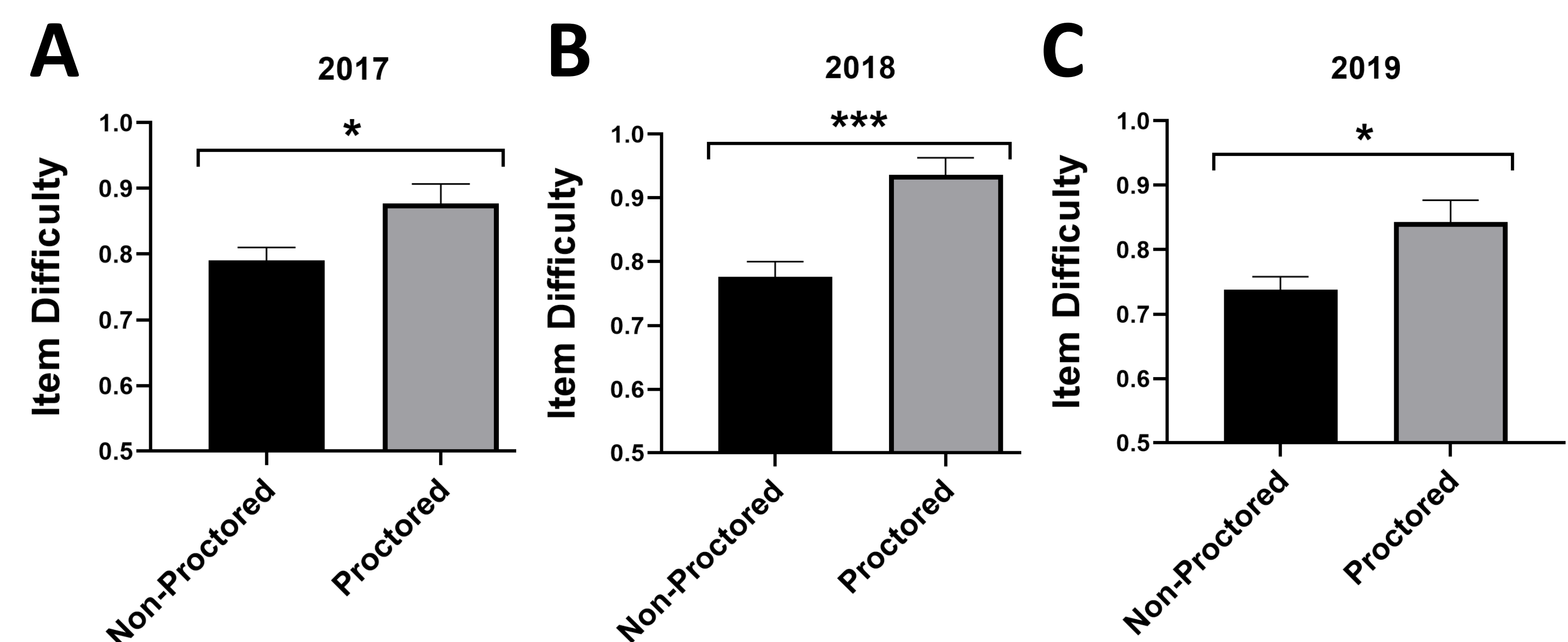
Assessment	Category	Setting	# of Questions	Time per Question
Weekly quiz	Formative	Non-Proctored	20	2 minutes
TBL IRAT	Formative	Proctored	10	1.6 minutes
Final Exam	Summative	Proctored	100	1.6 minutes

- All assessment items are linked to USMLE topics and specified learning objectives (LOs).
- Final exam minimum pass level did not vary significantly between years.

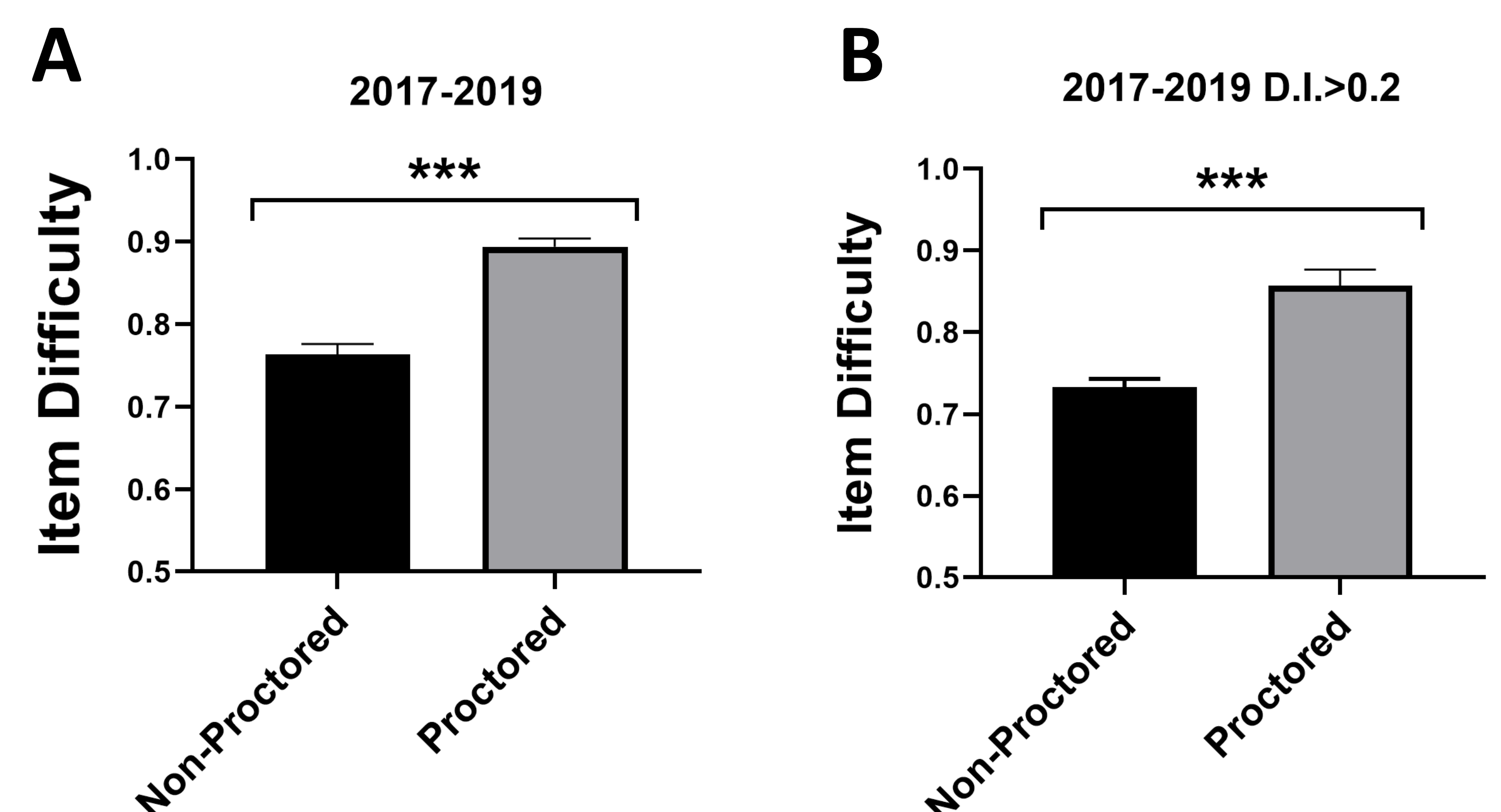
## Data Collection and Analysis

- Data set consisted of final exam items with linked learning objectives which were previously tested on formative assessments.
- Final exam items were categorized as Non-Proctored or Proctored according to whether the linked LO was previously assessed on a quiz or TBL IRAT, respectively.
- Means of item difficulty (percentage of students answering correctly) were compared for Non-Proctored and Proctored final exam items.
- Data was analyzed using an unpaired, two-sample T-test for each of three years (2017, 2018, and 2019) and cumulatively.
- Total number of students completing assessments was N = 187 (2017 N = 58; 2018 N = 67; 2019 N = 62)

## Results



**Figure 1: Exam items on proctored content are more likely to be answered correctly by students.** Mean percentage correct (item difficulty) is shown for final assessment items linked to content assessed in proctored or non-proctored settings in 2017 (A), 2018 (B), and 2019 (C). Data represent mean score +/- SEM. \* p value < 0.05. \*\*\* p value < 0.0001.



**Figure 2: Cumulative data from 2017-19 show higher summative assessment scores on proctored content.** Mean percentage correct (item difficulty) is shown for all instances of final assessment items over three years (A). The difference between Non-Proctored and Proctored is maintained when limiting analysis to only assessment items with a point biserial coefficient greater than 0.2 (B). Data represent mean score +/- SEM. \*\*\* p value < 0.0001.

## Conclusion and Discussion

- Students performed better on summative exam items that were tested previously in a proctored, formative setting versus a non-proctored, formative setting.
- Employing formative assessments in settings that more closely mirror the rigor and environment of high-stakes summative assessments may enhance student success.
- Future work will expand this analysis to additional courses across the undergraduate medical curriculum.
- Additional studies will differentiate effects of concentrated subject-specific formative assessment compared to formative assessment on topics from an entire week.