# Impact of Collaborative Testing on Academic Performance in Pharmacy Education

William A. Prescott Jr., PharmD[1,2]; Jaime Maerten-Rivera, PhD[1]; Ashley Woodruff, PharmD[1]; Nicholas Fusco, PharmD[1]

[1]University at Buffalo School of Pharmacy and Pharmaceutical Sciences

[2]Editorial Board Member, American Journal of Pharmaceutical Education

**University at Buffalo — The State University of New York**

## Bottom-Line

- Collaborative testing (CT), a 2-stage assessment strategy where students take an exam individually and then again as a team, encourages peer-to-peer teaching.
- Students performed better when taking an exam as a team, with significantly better performance across all question levels (recall, skill/concept, and analysis).
- CT was associated with improved learning and retention as demonstrated by student performance on re-examination.
- CT was well-received by students. Notably, 87% of students agreed that CT contributed to their learning and 86% preferred it to traditional 1-stage individual testing (IT).
- A higher proportion of students agreed CT (compared with IT) helped them learn from mistakes and retain what they learned; this was in alignment with exam performance data.
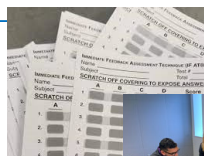
## Objectives

- To determine the impact of CT on academic performance among pharmacy students and secondarily, to characterize their perceptions of CT.

## Methods

|  | Exam 1 | Exam 2 | Exam 3 | Exam 4 | Exam 5 |
|---|---|---|---|---|---|
| Group 1 | CT | IT | CT | CT | IT |
| Group 2 | CT | CT | IT | IT | CT |

Re-examination of high-performing items from previous exams
Final examination of previously validated, related items

IF-AT Form

- Students enrolled in a 2-course patient assessment sequence completed five 80-minute IT exams, with two groups of 13 teams of 4-5 students alternating 40-minute CT exams.
- Learning and retention were assessed by comparing section, re-examination, and final exam performance.
- Student perceptions of CT were assessed through a 12-item survey with 7-point Likert-scale and open-ended questions.
- Data were analyzed statistically using the Fisher's Exact Test, Wilcoxon signed rank test, and Mann Whitney U test.
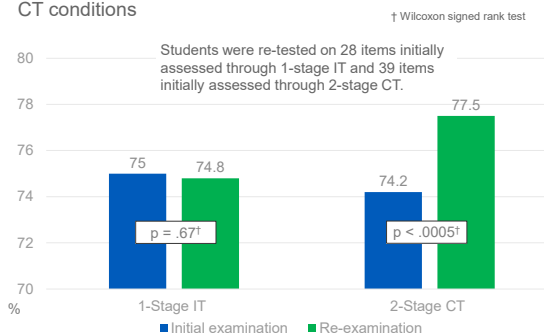
## Results
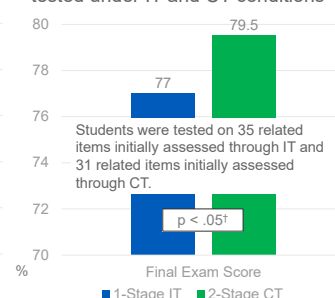
Student performance under individual and team testing

| | Group 1, n=62 | | | Group 2, n=60 | | |
|---|---|---|---|---|---|---|
| | IT mean (sd) | CT mean (sd) | $p$ value | IT mean (sd) | CT mean (sd) | $p$ value |
| Exam 1 (PA1)[a] | 71.5 (11.9) | 94.1 (3.3) | <.001[f] | 73.3 (12.6) | 95.9 (2.9) | <.001[f] |
| Exam 2 (PA1)[b] | 79.8 (11.6) | | | 78.4 (13.0) | 95.6 (3.3) | <.001[f] |
| Exam 3 (PA1)[c] | 78.2 (10.4) | 96.1 (2.9) | <.001[f] | 78.3 (10.4) | | |
| Exam 4 (PA2)[d] | 83.6 (9.3) | 97.5 (2.2) | <.001[f] | 81.1 (11.5) | | |
| Exam 5 (PA2)[e] | 73.2 (11.7) | | | 73.4 (12.6) | 96.4 (2.6) | <.001[f] |

Footnotes: PA1=Patient Assessment 1, PA2=Patient Assessment 2, IT=Individual Test, CT=Collaborative Test. [a]KR20=0.51; [b]KR20=0.67; [c]KR20=0.54; [d]KR20=0.4; [e]KR20=0.51; [f]Wilcoxon signed rank test was used to determine significance, defined as $p$<.05, between IT and CT.
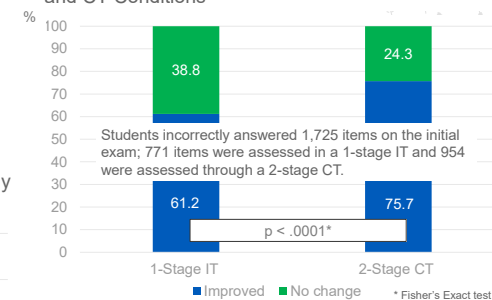
Comparative analysis of performance over time under IT and CT conditions

† Wilcoxon signed rank test

Students were re-tested on 28 items initially assessed through 1-stage IT and 39 items initially assessed through 2-stage CT.

1-Stage IT: Initial examination 75, Re-examination 74.8, p = .67†
2-Stage CT: Initial examination 74.2, Re-examination 77.5, p < .0005†
■ Initial examination ■ Re-examination

Comparative analysis of final exam performance on related items initially tested under IT and CT conditions

Students were tested on 35 related items initially assessed through IT and 31 related items initially assessed through CT.

Final Exam Score: 1-Stage IT 77, 2-Stage CT 79.5, p < .05†
■ 1-Stage IT ■ 2-Stage CT

Analysis of re-examination performance on items initially answered incorrectly on exams 1-5 under IT and CT Conditions

1-Stage IT: Improved 61.2, No change 38.8
2-Stage CT: Improved 75.7, No change 24.3
p < .0001*

Students incorrectly answered 1,725 items on the initial exam; 771 items were assessed in a 1-stage IT and 954 were assessed through a 2-stage CT.

■ Improved ■ No change   * Fisher's Exact test

### Academic performance

Students performed better on recall (p<.05) and skill/concept (p<.0001) items initially missed on an individual exam under CT conditions; there was no difference in performance on analysis items by testing method (p=.41).
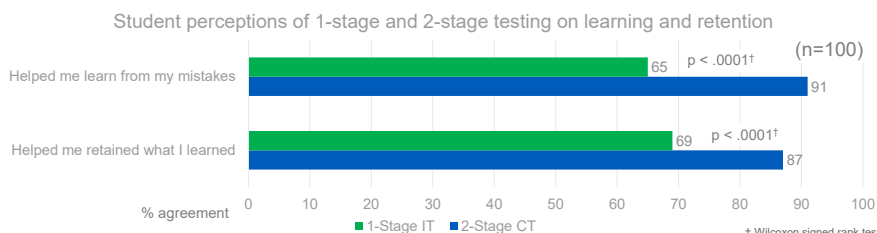
### Student Perceptions of CT

Survey response rate: 84%

| Item | Disagree, No. (%) | Neutral, No. (%) | Agree, No. (%) |
|---|---|---|---|
| Contributed in a positive way to my learning (n=102) | 5 (4.9) | 8 (7.8) | 89 (87.3) |
| Helped me develop my analytical skills (n=102) | 4 (3.9) | 7 (6.9) | 91 (89.2) |
| Helped me develop my critical thinking skills (n=102) | 3 (2.9) | 6 (5.9) | 93 (91.2) |
| Helped me learn from my mistakes (n=100) | 4 (4.0) | 5 (5.0) | 91 (91.0) |
| Helped me retain what I learned (n=100) | 7 (7.0) | 6 (6.0) | 87 (87.0) |

### Student perceptions of CT

86% of students preferred 2-stage CT to traditional testing methods.

Student perceptions of 1-stage and 2-stage testing on learning and retention (n=100)

Helped me learn from my mistakes: 1-Stage IT 65, 2-Stage CT 91, p < .0001†
Helped me retained what I learned: 1-Stage IT 69, 2-Stage CT 87, p < .0001†

% agreement
■ 1-Stage IT ■ 2-Stage CT
† Wilcoxon signed rank test

### How student-groups approached CT

Most indicated they "usually discussed each question until all members agreed on an answer" (47.1%), followed by "voted and if unanimous moved on, otherwise discussed the question until all members agreed on an answer" (32.4%).

### Student-identified strengths of CT

Promoted peer-peer learning
Provided insight into how other students think
Reinforced correct answers
Clarified misunderstandings in a timely manner
Helped students retain information.

References:
- Rivaz M et al. Effect of collaborative testing… *J Adv Med Educ Prof* 2015; 3(4): 178-82.
- Lindsley JE, et al. The two-stage examination… *Acad Med* 2016; 91: 1384-1387.
- Rieger GW, et al. Examinations that support collaborative… *J Coll Sci Teach* 2014; 43(4): 41-47.
- Hannah K, et al. Collaborative testing… *Nurse Educ* 2016; 41(4): 171-174.
- Levine RE, et al. High-stakes collaborative testing: Why not? *Teach Learn Med* 2018; 30(2): 133-140.