Knowledge Test Analysis
10.13.21
Deb Rooney

Methods:

Study. 20 participants from 3 sites completed the web-based module Participants included 2 Ob/Gyne attendings, 2 General Surgery attendings, 3 first-year residents, 1 second-year resident, 1 third-year resident, and 7 medical students. All participating sites were represented (Mbingo, n=3; Soddo, n=6; SIU, n=3; UM, n=8).

Scoring and Statistical analyses.

The identical (but shuffled in presentation) 10-item pre- and post-module quizzes were scored dichotomously (1=correct, 0=incorrect) and summed for each participant. Pre- and post-module summed scores were compared using paired student-test with SPSS Statistics for Windows v.25 (IBM, Armonk, NY) while differences in scores across participants' experience and site were analyzed using a many-facet Rasch model using Facets software v. 3.50 (Winsteps.com, Beaverton, OR) following anchoring on subjects to accommodate for nested design across sites.

RESULTS

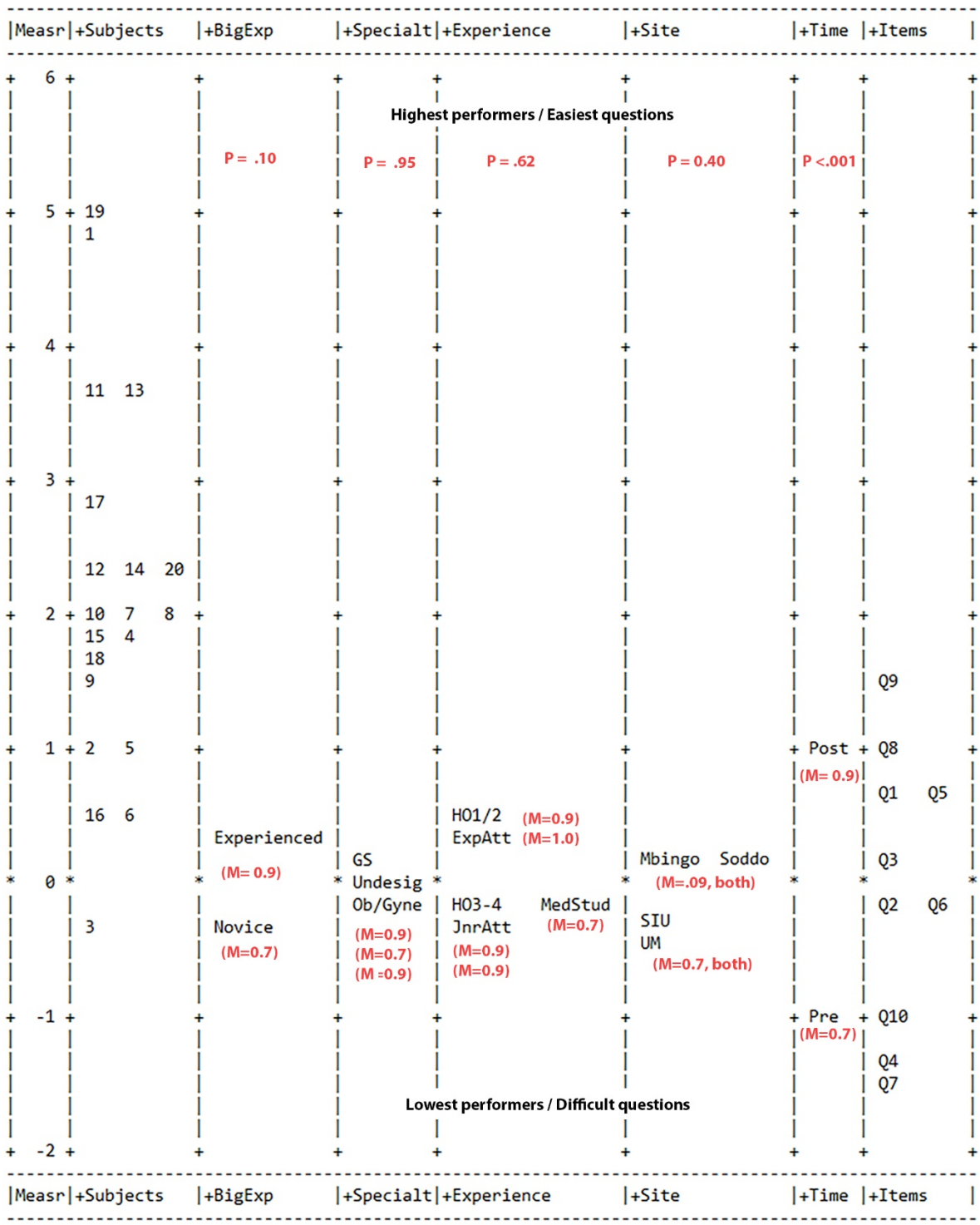
Paired Student T-Test

Comparison of pre- and post-intervention Quiz summed scores from all 20 participants indicated that there was a statistically significant improvement in mean summed scores from Pre (M=6.7, SD=2.2) to Post (M=9.0, SD = 1.5), t(20)= -4.76, P<.000.

Analysis using 7-facet Rasch model

Rasch analysis confirmed statistical improvement across Pre- Post-module Quiz scores, P< .001. Analysis also indicated higher scores for experienced (HO3 and Attendings; M=0.9) over novice (M=0.7) participants, but statistical significance was not reached, P = 0.10 (Figure 1. Variable Map). Further, there were no difference in performance across Specialty (P=.95), or Site (P=0.40).

Figure 1. Rasch Variable Map



Item discrimination

Review of item discrimination showed reasonable distribution of item difficulty for items, with item 7 as the most difficult item (item discrimination=.68), and Q9 as the easiest (item discrimination =1.03) (Table 1).

Item No.	Difficulty	Estimated Discrimination*	Discrimination Power	Suggested Action
Q4	Most difficult	1.40	Highest	Dblcheck language for ambiguity, too difficult?
Q5	Relatively easy	1.36	High	–
Q1	Relatively easy	1.29	High	–
Q6	Moderate	1.04	Good	–
Q9	Easy	1.03	Good	–
Q3	Moderate	0.99	Good	–
Q8	Easy	.92	Good	–
Q10	Moderate	.87	Acceptable	–
Q7	Most difficult	.68	Low	Consider removal/rewriting-too difficult to discriminate
Q2	Moderate	.52	Lowest	Consider removal/rewriting-too difficult to discriminate across participants. Review of response patterns (not shown here) indicate possible guessing from lower performers (higher OutfitMnSq=1.73)

*Values over 1 indicate this item has more discrimination power than expected for its difficulty while values under 1 indicate less discrimination power for its difficulty.

Because of this analysis, we made the following changes to our pre and post tests:

- Question 7: changed to "Which of the following is true of laparoscopic salpingostomy compared to open salpingostomy?"
- Question 12: added descriptors of eponyms