Context, Gender, and Physics Questions

Laura McCullough
University of Wisconsin-Stout
Physics Department

The Study

- Two versions of the Force Concept Inventory were given to students; the original version and a revised version with stereotypically female contexts*

- Pretest and posttest were analyzed separately

- Analysis focused on whether different context produced different response

*McCullough, L. (2004). Gender, Context and Assessment. Journal of International Women's Studies 5(4). Special Issue: Women in Science.

General Conclusions

- Overall averages between the two versions are the same

- By question there is enormous variation between the two versions

- Pretest differences are larger than posttest differences; due to less focus on context and more focus on content?

- When cued to the test similarities, students had somewhat fewer differences between the tests (averages were the same between versions)

Further Study

- This spring, I gave students the FCI, then the revised version for extra credit, then the next day I offered the FCI again for extra credit. I can compare differences in the two FCI responses to see how much variation is due to guessing or other factors

Testing conditions

- Four groups of data

  - Non-physics classes

  - UW-Stout physics students

    - Matched students (took both tests)

  - UW-Stout physics students cued to test (told it was the same test with different language; not told if answers were the same)

Non-Physics Classes

% correct by gender and version (standard error)

	Original	Revised
Women	21.7 (0.94) N=106	22.3 (0.91) N=79
Men*	33.7 (1.9) N=56	28.5 (1.4) N=71

UW-Stout Physics-Unmatched Sample

% correct by gender and version (pretest)

	Original	Revised
Women*	23.5 (0.95) N=99	29.4 (1.3) N=93
Men*	34.3 (0.99) N=184	39.4 (1.6) N=132

% correct by gender and version (posttest)

	Original	Revised
Women	36.5 (1.4) N=93	38.0 (1.6) N=121
Men	50.1 (1.1) N=247	48.4 (1.6) N=157

UW-Stout Physics- Matched Sample

% correct by gender and version (pretest)

	Original	Revised
Women	25.3 (1.7) N=30	24.9 (1.4) N=30
Men	35.0 (1.6) N=100	33.6 (1.5) N=100

% correct by gender and version (posttest)

	Original	Revised
Women	36.3 (2.7) N=26	35.8 (2.8) N=26
Men	50.0 (1.9) N=87	50.2 (2.0) N=87

UW-Stout Physics-Cued Sample

% correct by gender and version (pretest)

	Original	Revised
Women	23.3 (4.3) N=10	26.3 (2.7) N=10
Men	38.8 (3.6) N=31	38.5 (3.5) N=31

% correct by gender and version (posttest)

	Original	Revised
Women	44.8 (6.5) N=7	45.2 (6.8) N=7
Men	53.1 (4.6) N=30	52.2 (4.2) N=30

UW-Stout Physics-Matched Sample (Pre/Post)

- Average scores are the same; are they answering questions the same? No!

- Overall an average of **13** questions answered differently on the pretests, average of **9** questions different on posttests;
  individuals had between **1** and **21** different answers on the pretests, between **0** and **22** different answers on the posttests

- Women averaged **15** questions different on the pretests, **11** questions different on the posttests;
  individual answers differed between **5** and **21** on pretests, between **3** and **22** on the posttests

- Men averaged **13** questions different on pretests, **9** questions different on the posttests;
  individual answers differed between **1** and **21** on pretests, between **0** and **22** on posttests

UW-Stout Physics-Cued Sample (Pre/Post)

- Did the cued students answer the questions the same? No!

- Overall average of **11** questions answered differently on pretest, **6** on posttest; individuals answered between **4** and **21** questions differently on pretests, between **0** and **22** on posttests

- Women averaged **13** questions different on pretests, **5** on posttests; individual answers between **7** and **21** on pretests, between **0** and **8** on posttests (N=7)

- Men averaged **10** questions different on pretests, **6** on posttests; individual answers between **4** and **17** on pretests, between **0** and **22** on posttests (N=30)

Questions provoking changed responses (Pre/Post)

- Looking at individual questions (not individual persons) the average % of the class changing responses to any particular question was **44%** on the pretest and **31%** on the posttest

- For the pretest, the biggest change was **58%**, the smallest was **22%** (on one question **58%** of the class changed their answer)

- For the posttest, the biggest change was **56%**, the smallest was **16%**

Questions provoking changed responses by gender (Pre/Post)

- Looking at individual questions (not individual persons) the average % of women changing responses to any particular question was **36%** on the pretest and **36%** on the posttest; for men, the average was on the pretest and **30%** on the posttest

- For the pretest, the biggest change for women was **77%**, the smallest was **27%**; for men the biggest change was **60%** and the smallest was **19%**

- For the posttest, the biggest change for women was **58%**, the smallest was **19%**; for men the biggest change was **55%** and the smallest was **15%**

- More women changed answers than men on all but 6 questions