Differences in male/female response patterns on alternative-format versions of FCI items

Laura McCullough, Physics Department, University of Wisconsin–Stout, Menomonie, WI 54751
David E. Meltzer, Department of Physics & Astronomy, Iowa State University, Ames, IA 50011

A modified version of the FCI was created using female and daily-life contexts instead of the male and school-oriented contexts in the original. Both modified and original versions were administered in class. Differences among responses of males and females to both versions are discussed.

An important methodological issue in assessments using a diagnostic instrument is the degree to which slight changes in the instrument may result in altered student response patterns.[1] In view of the widespread use of the Force Concept Inventory (FCI) in physics assessment, exploration of possible context dependence of FCI items is of considerable interest. One issue is that of possible context dependence *in general* of FCI items: do students respond differently when FCI questions are very slightly modified, i.e. when essentially identical questions are posed in a different context? Steinberg and Sabella[2] compared responses on open-ended questions to responses on their FCI equivalents. They found substantial correlation between the results of the two versions with, nonetheless, some notable differences. Rebello and Zollman[3] presented students with four actual FCI questions without including the multiple-choice answer options. They also found, along with general overall agreement, a number of notable differences in response patterns when compared to those observed when multiple-choice responses *were* present. On the other hand, Adams and Slater[4] found that students' written explanations for their FCI responses were, for the most part, in good agreement with the answers they selected. Schecker and Gerdes[5] presented students with a number of FCI items, along with alternate versions of the same items posed in different physical contexts. They found differences in response patterns on some items. For instance, many students who incorrectly responded with an "impetus" model to FCI item 13 (forces acting on a steel ball thrown straight up) gave a correct Newtonian response when an almost identical question was asked with the ball being replaced by a vertical pistol shot.

Another potential issue is whether any possible context dependencies in response patterns are gender dependent. That is, do males and females differ *from each other* in terms of how their responses may change when question context is altered? This issue has been addressed by Rennie and Parker,[6] who suggest that females may be more successful when physics problems are posed in a "real-life" context. Dancy[7] explored this issue in the context of an animated version of the FCI. She found that on questions 3, 5, 14, and 26, females scored significantly better on the animated version than on the original version. Males scored significantly better on the animated version of questions 7, 14, and 26, but worse on item 20. Item 14 is of particular interest because both genders did better on the animated version and because of an unexpected response pattern in our own data.

In order to further explore the issue of possible context dependence, a "Gender" version of the FCI has been developed[8] in which each of the 30 items was rephrased or re-expressed in a slightly different context, or with a new or added diagram. Instead of school- and male-oriented contexts, daily-life- and female-oriented contexts were used in each case (e.g., instead of a cannon shooting a cannonball, a baby knocks a bowl off of her high-chair tray). The physics is identical; only the context has changed.

METHOD

The Original FCI and the Gender FCI were administered to all students enrolled in the first semester course of the algebra-based general physics sequence at Iowa State University during Spring 2001. Only one of the two versions was given to each student, that version being randomly chosen according to the following procedure: Individual piles of question packets, 28 in each, were prepared for each recitation section. In each pile, the Original FCI and Gender FCI were placed alternately, so the sequence was Original, Gender, Original, etc. The recitation instructors were directed to distribute the packets in random order to all the students in their recitation section.

The tests were administered at the start of the recitation session during the second week of class. Students were told the tests would not affect their grade, but would give instructors a better idea of the students' physics background. Instructors were directed to allow at least 30 minutes, and to try to allow all students enough time to finish. Reports indicated good compliance. In one case, the instructor allowed the students to take the exams home and hand them in two days later. Response sheets that contained six or more blank responses were discarded. Three had to be discarded because the "Sex" box was not checked and the names were gender-indeterminate. In the end, the total sample contained 222 students.

We checked the results for every question to see whether there were any significant differences in performance on the two different versions of the exam. Because there are so many comparisons, we adopted $p = 0.01$ as the minimum level required to consider the difference significant. We used a statistical test for comparison of binomial proportions (equivalent in this case to chi-square analysis).

RESULTS

We found significant discrepancies for four test items (two for females only, two for males only), as follows:

1. Original FCI Item #14 (Gender item #24) [airplane/eagle drops object]

Female:
Original correct: 22%
Gender correct: 55%
$p = 0.002$*

2. Original FCI item #23 (Gender item #27) [rocket/person straight line path]

Female:
Original correct: 10%
Gender correct: 48%
$p = 0.0001$*

3. Original FCI Item #22 (Gender item #26) [rocket/person speed increasing]

Male:
Original correct: 47%
Gender correct: 18%
$p = 0.0003$*

4. Original FCI Item #29 (Gender item #13) [floor force on chair/book]

Male:
Original correct: 30%
Gender correct: 60%
$p = 0.0005$*

(Detailed results and text of "Gender" assessment items follow on the next two pages.)

CONCLUSION

Our data suggest that, in certain cases, slight changes in the context of a conceptual question may affect students' performance. Moreover, it appears that males and females may not be consistent with each other in their response to the contextual changes. More work is needed to better understand how changes in physics assessment instruments may depend on gender in their effect on performance.

ITEMS WITH SHIFTS FOR FEMALES:

Original FCI #14 (Gender item #24)
[plane/bird drops object]
Percentage of total responses each option:

	n	A	B	C	D*	E
Male, Original	*72*	10	8	8	74	0
Male, Gender	*65*	0	6	22	72	0
Female, Original	*41*	49	27	2	22	0
Female, Gender	*44*	9	20	16	55	0

Comment: No significant difference in correct responses for males. On Gender version, females show *drastically* decreased proportion selecting distracter A (which shows "backward" trajectory). Males show decrease for this option on Gender version as well. Net result is increase in correct responses by females from 22% to 55%.

Original FCI #23 (Gender item #27)
[rocket/person straight-line path]
Percentage of total responses each option:

	n	A	B*	C	D	E
Male, Original	*72*	7	47	22	19	4
Male, Gender	*65*	8	52	15	17	8
Female, Original	*41*	27	10	41	17	5
Female, Gender	*44*	5	48	41	5	2

Comment: No significant difference in correct responses for males. On Gender version, females show much higher proportion of correct responses (48% compared to only 10%), mostly due to decrease in number choosing A (sideways path which ignores velocity component due to applied force).

*** indicates correct answer**

ITEMS WITH SHIFTS FOR MALES:

Original FCI #22 (Gender item #26)
[rocket/person speed increasing]
Percentage of total responses each option:

	n	A	B*	C	D	E
Male, Original	*72*	32	47	1	18	1
Male, Gender	*65*	42	18	11	23	6
Female, Original	*41*	29	24	2	39	5
Female, Gender	*44*	34	25	16	18	7

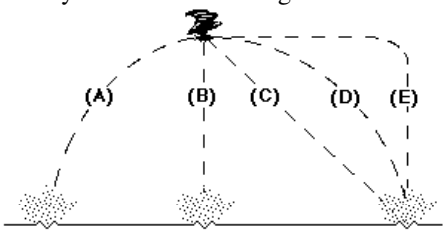
Comment: No significant difference in correct responses for females; however, number of females choosing "decreasing" speed is higher on Gender version. Males also show sharp increase in number choosing "decreasing" speed (and in those who choose "constant" speed). Net result is sharp *decrease* in correct responses by males, 47% to 18%.

Original FCI #29 (Gender item #13)
[surface force on chair/book]
Percentage of total responses each option:

	n	A	B*	C	D	E
Male, Original	*69*	33	30	4	32	0
Male, Gender	*65*	18	60	0	22	0
Female, Original	*39*	28	49	3	21	0
Female, Gender	*44*	30	57	0	11	2

Comment: No significant difference in correct responses for females. Proportion of correct responses for males *doubles* from 30% on Original to 60% on Gender version. Change comes mostly from increase in proportion who now recognize presence of upward force due to surface; also, there is a decrease in number who choose "all three" forces (including air pressure).

A bird is carrying a fish in its claws as it flies along in a horizontal direction above a lake. The bird accidentally drops the fish. As seen from the lakeshore, which path would the fish most closely follow after leaving the bird's claws?



Gender item #24 (Original FCI #14)

A diary is at rest on a nightstand. Which of the following force(s) is (are) acting on the diary?

1. A downward force of gravity.
2. An upward force exerted by the nightstand.
3. A net downward force exerted by the air.

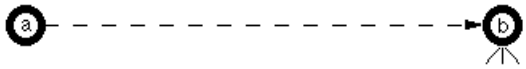
(A) 1 only
(B) 1 and 2
(C) 2 and 3
(D) 1, 2, and 3
(E) none of these.
Since the book is at rest there are no forces acting on it.

Gender item #13 (Original FCI #29)

[Note: there is no diagram included on the Original FCI version of this question.]

USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (25 through 28).

An ice storm has knocked out power in your area and has started a fire. You have grabbed your powerful fire extinguisher and are running to help out. At point "a" you start to slip on a large patch of <u>frictionless</u> ice, sliding across the ice from point "a" to point "b." (Note that this diagram shows a "top view," looking down from above.) At point "b," while trying to keep upright, you accidentally turn on the fire extinguisher. The fire extinguisher produces a constant force on you in a direction at right angles to line "ab," and you slide along the ice toward point "c." When you reach point "c," you are able to turn off the extinguisher, but you continue to slide on the ice.



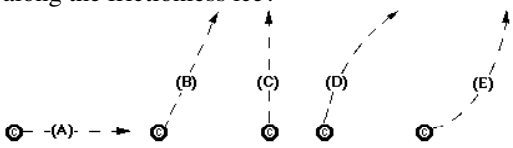
26. As you move from "b" to "c" along the ice, your **<u>speed</u>** is

(A) constant.
(B) continuously increasing.
(C) continuously decreasing.
(D) increasing for a while and constant thereafter.
(E) constant for a while and decreasing thereafter.

Gender item #26 (Original FCI #22)

---

27. At "c" the extinguisher is suddenly turned off completely. Which of the paths below will you follow beyond "c" as you continue to slide along the frictionless ice?



Gender item #27 (Original FCI #23)

---

[1]L. Enderstein and P. Spargo, "The effect of context, culture and learning on the selection of alternative options in similar situations by South African pupils," Int. J. Sci. Ed. **20**, 711 (1998).

[2]R. Steinberg and M. Sabella, "Performance on multiple-choice diagnostics and complementary exam problems," Phys. Teach. **35**, 150 (1997).

[3] N. S. Rebello and D. A. Zollman, "The effect of distracters on student performance on the Force Concept Inventory," *in review;* preprint at: http://www.phys.ksu.edu/perg/papers/

[4]J. P. Adams and T. F. Slater, "Student-supplied rationale for multiple-choice responses on the force concept inventory," *in review*; preprint at: http://www.physics.montana.edu/physed/

[5]H. Schecker and J. Gerdes, "Messung von Konzeptualisierungsfähigkeit in der Mechanik: Zur Aussagekraft des FCI," Zeitschrift für Didaktik der Naturwissenschaften **5**(1), 75-89 (1998).

[6]L. J. Rennie and L. H. Parker, "Equitable measurement of achievement in physics: high school students' responses to assessment tasks in different formats and contexts," J. Women and Minorities in Sci. Eng. **4**(2-3), 113-127 (1998).

[7]M. H. Dancy, *Investigating Animations for Assessment with an Animated Version of the Force Concept Inventory.* Ph.D. dissertation, N.C. State University (2000).

[8]L.E. McCullough and T. Foster, "A Gender Context for the Force Concept Inventory," AAPT Announcer **30**(4), 105 (2000).