

Belief vs. “Scientist Answers”: Effect on Assessment

Timothy L. McCaskey, Melissa H. Dancy,
and Andrew Elby



Supported by NSF REC grant #0087519

A Common Anecdote

From *Peer Instruction* (Mazur):

When taking the Force Concept Inventory (FCI) (Hestenes, Halloun, *et al.*), a student said:

“Professor Mazur, how should I answer these questions? According to what you taught us, or by the way I think about these things?”



Research Questions

- 1.) Do students taking standardized physics tests really believe the answers they put down?
- 2.) What happens when students are asked to distinguish between their belief and a “scientist’s answer”?
- 3.) Are there any gender effects?

(answer key) 1.) Often, no 2.) You’ll see 3.) Yes

Our FCI Study

We had students take the FCI once in the standard way. After that, we had them make a *second* pass with the following instruction:

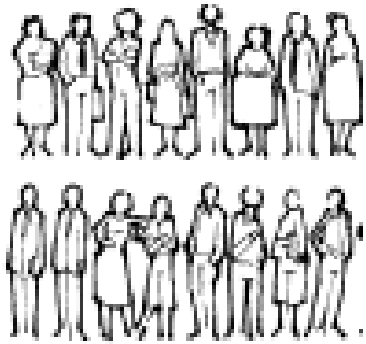
Please **circle** the answer **you really believe**.

Please draw a **square** around the answer **you think scientists would give**.



The Student Task - Splits

4. A large truck collides head-on with a small compact car. During the collision:
- (A) the truck exerts a greater amount of force on the car than the car exerts on the truck
 - (B) the car exerts a greater amount of force on the truck than the truck exerts on the car
 - (C) neither exerts a force on the other, the car gets smashed simply because it gets in the way...
 - (D) the truck exerts a force on the car but the car does not exert a force on the truck
 - (E) the truck exerts the same amount of force on the car as the car exerts on the truck



Our Populations

Students were taking introductory algebra-based courses at the University of Maryland (UMd) and Davidson College (DC). Three classes were tested:

UMd PHYS 121: 1st semester, FCI taken pre-instruction

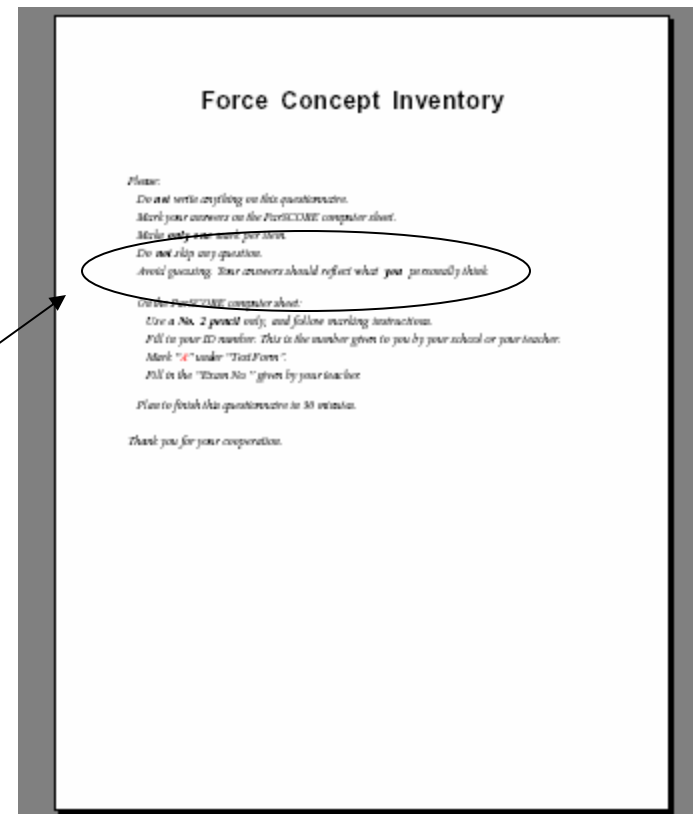
UMd PHYS 122: 2nd semester, FCI taken after one semester of traditional instruction

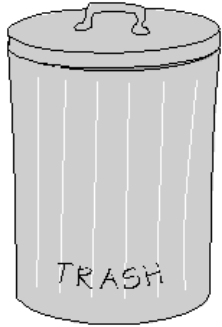
DC: 1st semester, FCI taken *immediately* after one semester of traditional instruction

A Methodological Difference

UMd's test contained the instruction "Avoid guessing. Your answers should reflect what *you* personally think."

DC's exam did not have this.





Unusable Data (UMd)

We had to discard data from any student that neglected to do five or more questions on any of the three passes (bubble sheet, circles, and squares).

We were also careful not to count splits if the discrepancy was due to a blank response.

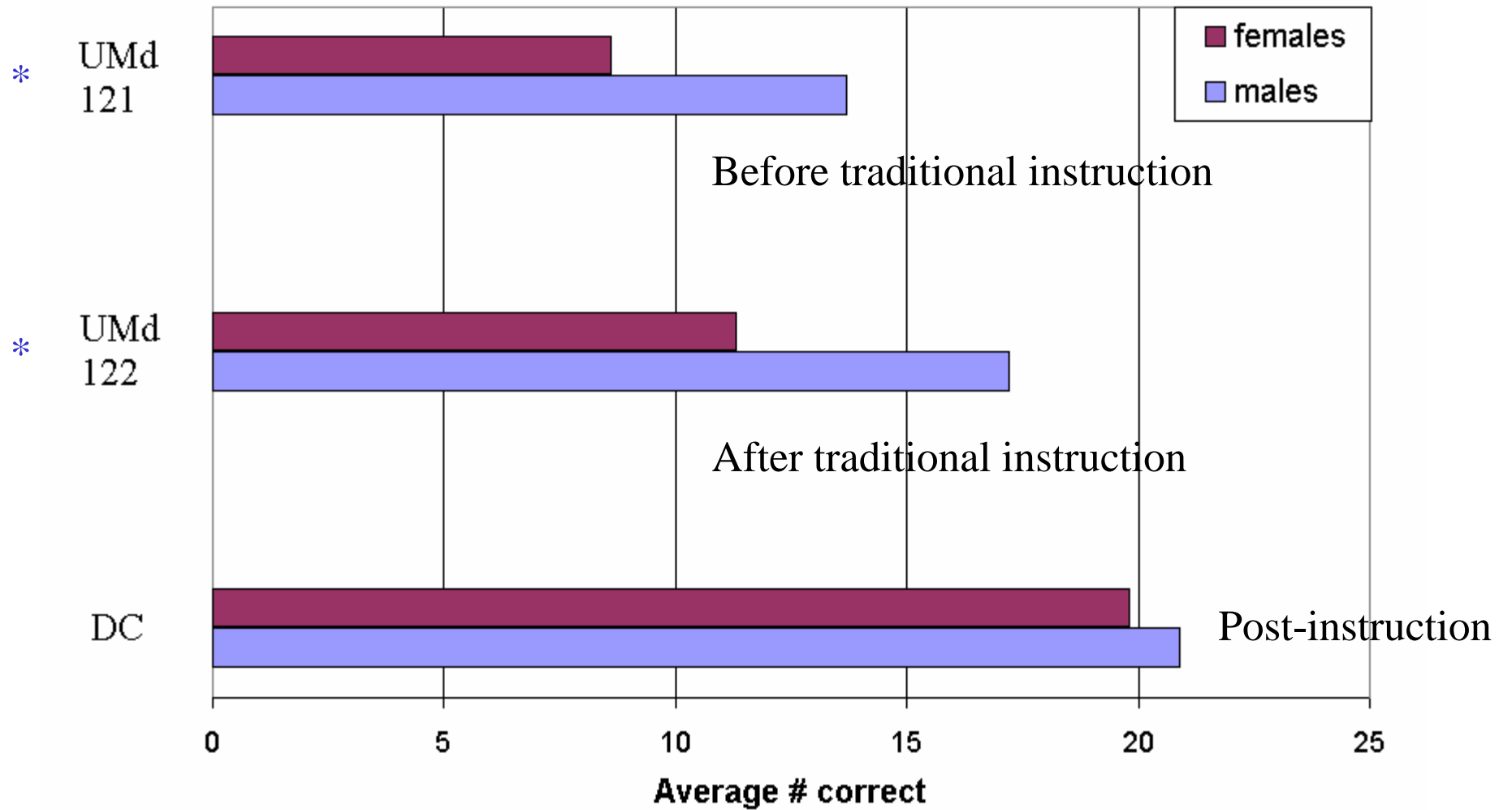


Gender Trends

- Scores are higher for men than women.
- The number of splits is higher for women in all three classes.
- When splits occurred, women were more likely to have bubbled in their “belief” answer on their first pass through the FCI.

Scores are higher...

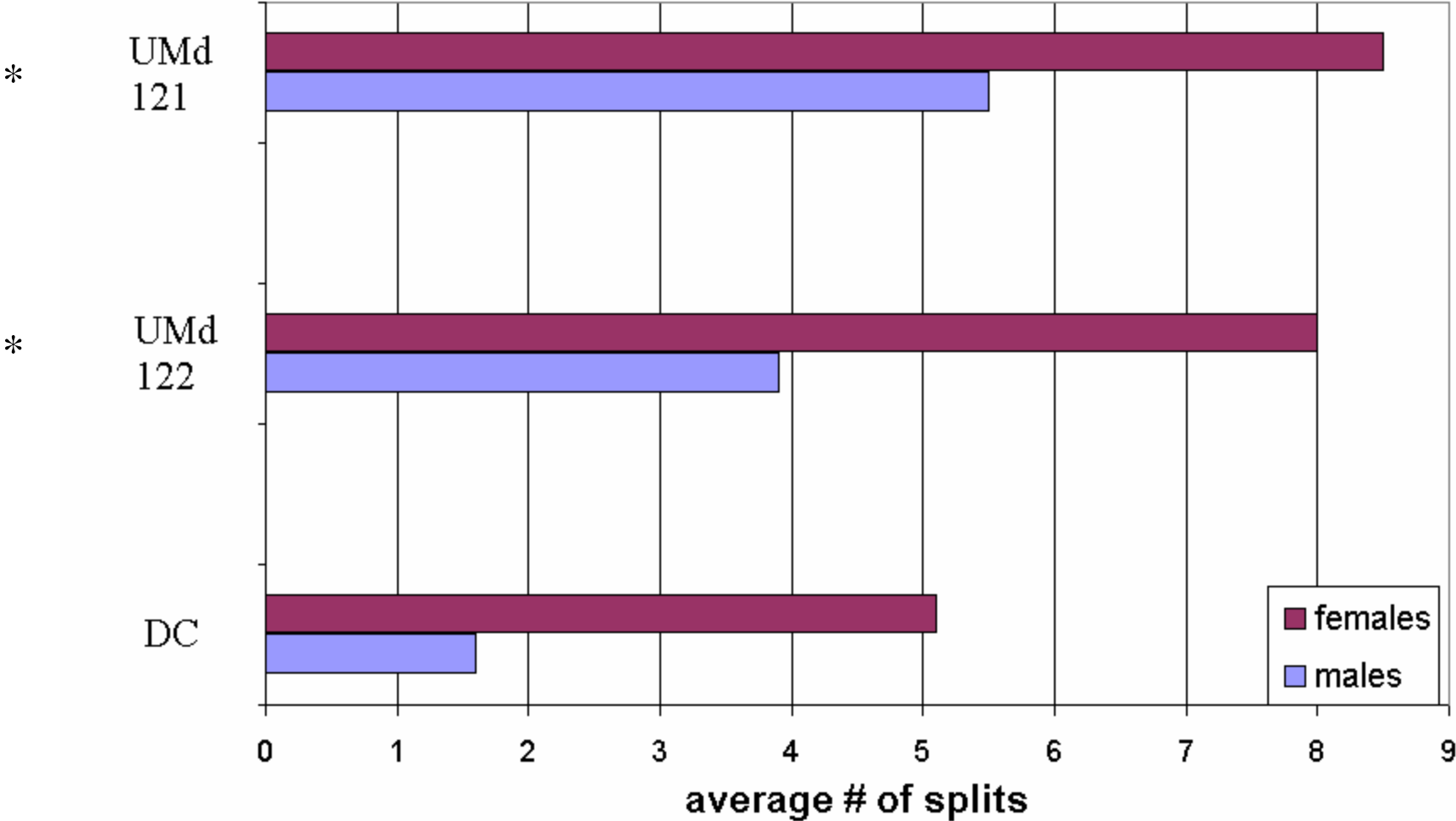
FCI Scores - "First Pass"



*: $p < .001$

The number of splits...

Splits on FCI



*: $p < .01$

The number of splits...

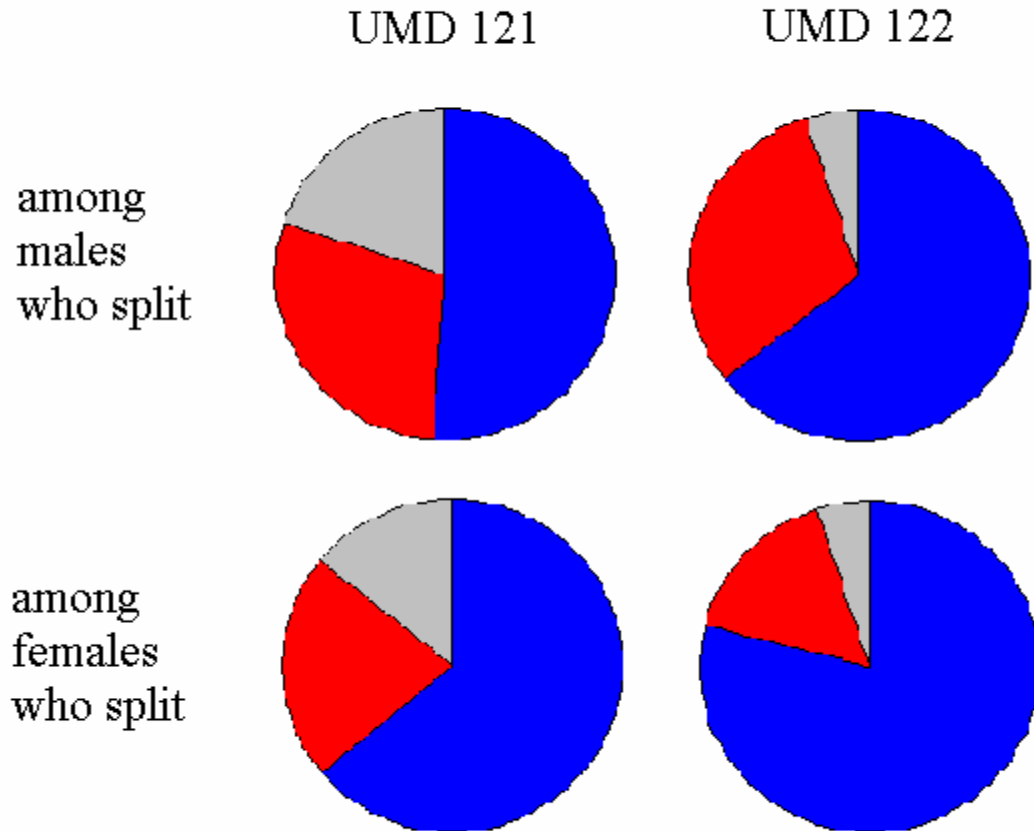
A Possible Explanation?

A comment from a Davidson student:

“Ok, I’m a little confused about the purpose of this. My answers the second go around won’t be different than my answers the first time. Even if my answers are wrong, I wrote them believing that scientists or anyone else for that matter would give the same answer.”

Different distributions...

When students have splits, what did they bubble in the first time through the FCI?

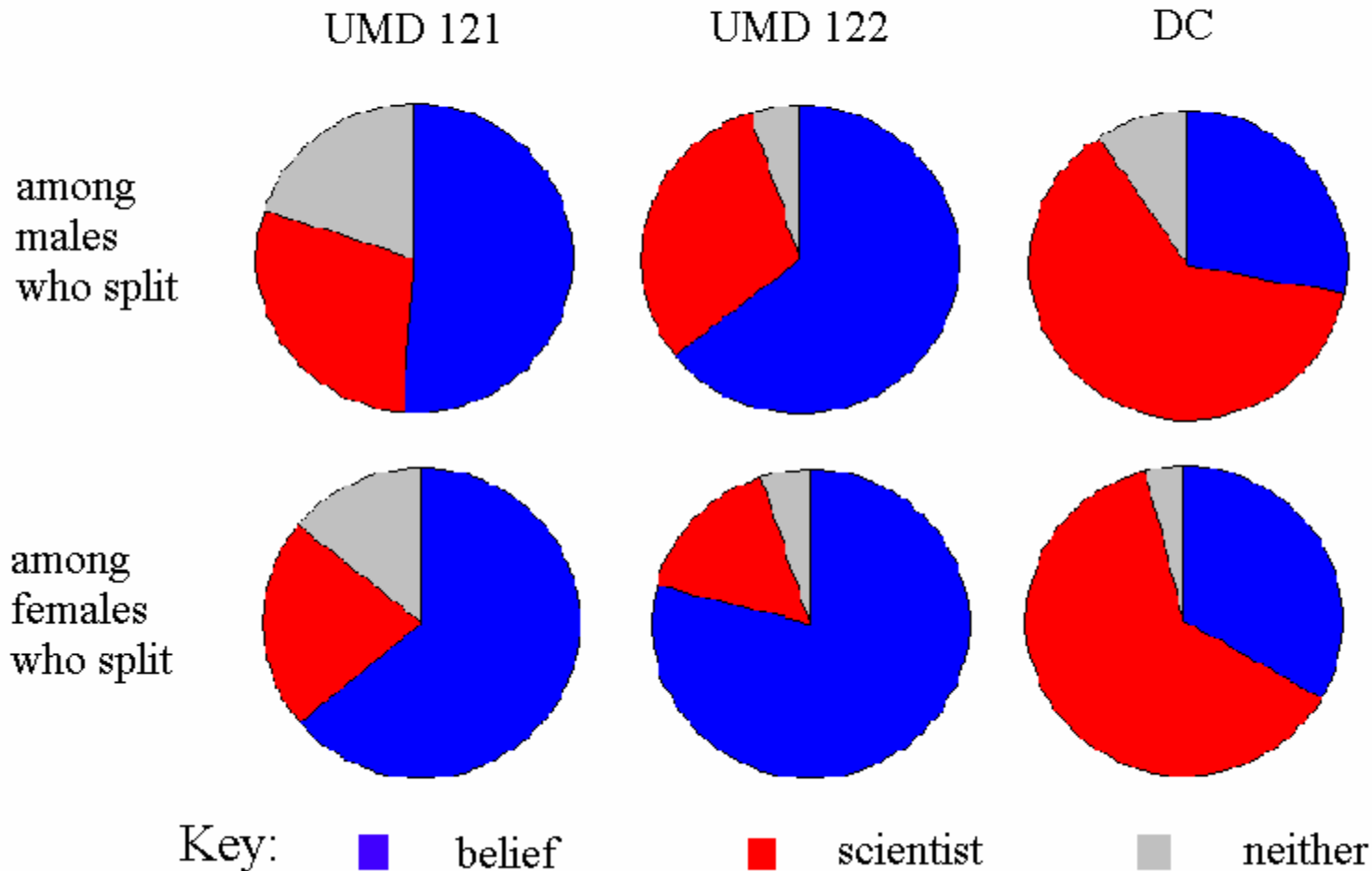


Among UMD students, people seem to be bubbling their belief more often when they have splits.

Key: ■ belief ■ scientist ■ neither

Different distributions...

When students have splits, what did they bubble in the first time through the FCI?





The “Newton 3” subcluster

- We looked at the 4 “Newton 3rd” questions for the larger class (UMd 121).
- 15 of 116 students got the correct “scientist answer” on all N3 questions.
- Of the 15, 12 had at least one split on the N3 items, and 9 had at least two splits.
- Implication: People that get things like N3 right on a standardized test may not always believe it!

Further Questions



- When students split either a lot or very little, what does that mean?
- If a student splits a lot, is reconciliation between “belief” and a “scientist’s answer” possible?
- Why do women split their answers more often, and do men and women split for the same reasons?
- What does all this say about FCI validity?
- Would this have any implications for teachers?

Email: mccaskey@physics.umd.edu

Also, see my poster at the PERC later this week.