We have conducted think-aloud interviews with students as they grappled with questions on the Force Concept Inventory (FCI). Doing so has shown us that the difficulties they have with some questions have nothing to do with their understanding of physics. These difficulties involve diagrams, notations, and vocabulary that make perfect sense to physics teachers but can easily confuse beginning students. Informed by these think-aloud interviews, we modified several questions to improve clarity and administered our revised FCI to two sections of introductory physics students. Here we show how and why we modified four questions and compare the consequent results with several years of archival data generated with the canonical FCI.

In order to demonstrate the effectiveness of our modifications to FCI questions we will use those think-aloud interviews, we modified several questions to improve clarity and administered our revised FCI to two sections of introductory physics students. Here we show how and why we modified four questions and compare the consequent results with several years of archival data generated with the canonical FCI.

**Equivalent Populations**

In order to demonstrate the effectiveness of our modifications to FCI questions we will compare results on the modified FCI to historical data. This kind of comparison obligates us to show that the new student population is similar in ability to past student cohorts.

First we compared the demographics of the two populations:

The historical data set includes four years of responses (N=431) collected in the fall semesters of 2007 - 2010. The modified FCI responses (N=139) were collected fall semester of 2012. All administrations were given during the first week of class to both the introductory algebra-based and introductory calculus-based courses.

The table above shows how the recent and historical populations were divided between gender and type of course. Clearly there is no remarkable difference between the two populations.

Second, we compared the pre-test results of the 15 unmodified questions using two methods.

1. We calculated the correlation coefficient between the percent of students in the new group choosing the correct answer to the percent of students in the historical group who chose the correct answer. The correlation between the modified questions is also shown for comparison.

**Per question**

Standard error was determined for individual questions assuming a binomial distribution: \( \text{SE} = \sqrt{\frac{\text{p}(1-\text{p})}{n}} \)

**Difference in % correct ranges from:** 0.16% to 5.9% which is within the 91% expected distribution range.

**Total score:** An average score on the FCI for the 15 unmodified questions was compared between the historical data and the new data. Standard error was calculated from the standard deviation of the 15-question set of scores in this case.

**Our Thoughts**

Now that the issue of air resistance has been somewhat clarified, student responses have migrated from the air resistance distractor to the very popular misconception that the upward force of the cable must be greater than that due to gravity for the elevator to move at a constant velocity.

**Modification of a Question with Confusing Wording**

**Original Question**

Interviews showed that this diagram was easy to misunderstand; probably because it is not correct. The plane is shown at the time the ball is released, but the ball is shown at the time it hits the ground. In addition, students put excessive and often incorrect value on the wording in the statement. “As observed by a person...”

To clarify the situation we added a drawing of the plane in its position at the time the ball hits the ground. We also removed the wording “as observed by a person...”

**Our Thoughts**

Nearly half of the students who had been choosing option A, now choose the correct answer. We believe both changes, especially that concerning the diagram, were instrumental in bringing about this shift.

**Modification of a Question with an Incorrect Diagram**

**Original Question**

Interviews showed that this diagram was easy to misunderstand; probably because it is not correct. The plane is shown at the time the ball is released, but the ball is shown at the time it hits the ground. In addition, students put excessive and often incorrect value on the wording in the statement. “As observed by a person...”

To clarify the situation we added a drawing of the plane in its position at the time the ball hits the ground. We also removed the wording “as observed by a person...”

**Our Thoughts**

It is also interesting to note that the percent correct on this question (54%) is closer to the percent correctly answering #12, the cannon question (64%), shown below. One might have thought the big difference in the student performance (38% and 64%) between these two questions on the original FCI was due to different physical reasoning. Now it seems that this is not necessarily so.

**Modification of Questions with Poor Representation**

**Original Questions**

Interviews suggested that some students were at a loss as to what the figures of Questions 19 and 20 physically represent. In particular, it was not clear how time and position were related. That the numbers next to the blocks were a proxy for time was a subtle point easily misunderstood.

To remedy this confusion we added a preliminary question and switched the order of the other two. The added question was intended to demonstrate, in an obvious way, how position and time were related in this style of representation.

**Our Thoughts**

The order of the original questions was switched because we wished to present the problems in what the interviews indicated to be situations of increasing complexity.

**Conclusion**

We have explored the consequences of modifying some FCI questions in order to achieve greater clarity and to address the common misconception that the upward force of the cable must be greater than that due to gravity for the elevator to move at a constant velocity.

**Our Thoughts**

The new question 19 is clearly very easy for students to interpret as we intended. Clear improvement in correct responses is seen for both 20 and 21. Our supposition is that adding the easy question caused most of this change but interviews will be needed to support this idea.

**Total score:** An average score on the FCI for the 15 unmodified questions was compared between the historical data and the new data. Standard error was calculated from the standard deviation of the 15-question set of scores in this case.

**Our Thoughts**

Nearly half of the students who had been choosing option A, now choose the correct answer. We believe both changes, especially that concerning the diagram, were instrumental in bringing about this shift.

It is also interesting to note that the percent correct on this question (54%) is closer to the percent correctly answering #12, the cannon question (64%), shown below. One might have thought the big difference in the student performance (38% and 64%) between these two questions on the original FCI was due to different physical reasoning. Now it seems that this is not necessarily so.

**Unmodified Question**

To clarify the situation we added a drawing of the plane in its position at the time the ball hits the ground. We also removed the wording “as observed by a person...”

**Our Thoughts**

It is also interesting to note that the percent correct on this question (54%) is closer to the percent correctly answering #12, the cannon question (64%), shown below. One might have thought the big difference in the student performance (38% and 64%) between these two questions on the original FCI was due to different physical reasoning. Now it seems that this is not necessarily so.

**Unmodified Question**

To clarify the situation we added a drawing of the plane in its position at the time the ball hits the ground. We also removed the wording “as observed by a person...”

**Our Thoughts**

It is also interesting to note that the percent correct on this question (54%) is closer to the percent correctly answering #12, the cannon question (64%), shown below. One might have thought the big difference in the student performance (38% and 64%) between these two questions on the original FCI was due to different physical reasoning. Now it seems that this is not necessarily so.