Assessing the Simplified Force Concept Inventory as Adaptation for English Language Learners

Daniel Doucette
International School of Latvia, Pinki, Latvia

Abstract:
Physics teachers often assume student understanding of physical concepts can be readily assessed via written problems. However, for non-native English speakers, this assumption may be unfounded. Will a simplified vocabulary make physics problems more approachable for English language learners? This study attempts to establish a method by which this question might be answered. Additionally, results from a small sample are analyzed and found to suggest that a simplified test is not easier for students.

Introduction:
The Force Concept Inventory (FCI) is an important tool in physics education research [1]. The 30-question multiple-choice exam aims to provide an indication of a student's ability to think “Newtonially”; that is, to correctly apply the concepts usually associated with a high school- or undergraduate-level course in mechanics. The FCI has been administered to thousands of students, and has been used to demonstrate that traditional instruction is often insufficient for instilling “Newtonian” understandings [2].

In 1995, a debate about the applicability of the FCI resulted in a revised version [3]. The central issue of contention was the (now accepted) view that the FCI assesses understanding across 6 subdivisions of mechanics, but does not meaningfully distinguish between these domains [4].

The FCI has been critically analyzed for a variety of factors, including gender [5], question order [6], and reliability [7].

A linguistically-simplified version of the FCI, written in 2009 and updated in 2013, is the Simplified Force Concept Inventory (SFCI) [8]. The SFCI has been shown to provide comparable results to the FCI for students in grades 11-12, while providing accommodation for students in grade 9, who averaged two points better (out of 30) on the SFCI [8]. It has been proposed that the SFCI might be useful for English Language Learners (ELLs) [8].

Contemporary classrooms typically include a diversity of linguistic backgrounds. Adapting curricular and assessment materials be to accessible to ELLs is sometimes proposed as a useful and empirically-validated inclusive measure [9].

The goal of this study is to establish a procedure to determine the effectiveness of the SFCI as a replacement for the FCI for ELLs.

Methodology:
The participants in this study were eight students enrolled in grade 11/12 physics at the author's school. The participants spoke languages other than English as their home or native language, but were currently studying at an English-medium school. Three students were studying in grade 11, and five in grade 12.

The FCI and SFCI were both administered to all subjects one week apart. Half of the participants were randomly selected to write the FCI first, while the other half started with the SFCI. Participants were permitted up to 60 minutes to complete the tests, and typically took 45 minutes to finish.

To compare two sets of test results, physics education researchers typically use a t-test. However, this presumes that the FCI scores follow a normal distribution. There is no a priori reason to believe this to be true, nor is there (to the author's knowledge) any study that has demonstrated such a phenomenon. Instead, a non-parametric test must be used. For this study, a Wilcoxon signed-rank test is used [10].

Figure 3: Results of FCI/SFCI (max: 30)
Results & Discussion:

Figure 3 plots the FCI and SFCI scores of the subjects. The scores show little variation from one test to the next. The data are linearly correlated with a Pearson product-moment coefficient of $r = 0.833$, indicating that most of the variance in test scores is accounted for by some property that is measured by both tests (ie: “Newtonian” thinking).

The results from the eight participants are displayed in table 1. Also calculated are the gains of the SFCI, and the corresponding ranks.

<table>
<thead>
<tr>
<th>Subject</th>
<th>FCI</th>
<th>SFCI</th>
<th>gain</th>
<th>rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>8</td>
<td>-1</td>
<td>1.5</td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>9</td>
<td>-2</td>
<td>3.5</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>10</td>
<td>-4</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>18</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>14</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>19</td>
<td>21</td>
<td>2</td>
<td>3.5</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>14</td>
<td>-1</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 1: Results of FCI/SFCI (max: 30)

The Wilcoxon test statistic $W$ is the sum of the signed ranks, so $W = -2$ with 6 (non-zero) samples. Since this does not exceed the critical value at 95% confidence of $W_{crit} = 17$ [10], we cannot claim a difference between the results on the two tests. Thus, we see no benefit from the SFCI.

Of course, a sample size of eight cannot reasonably motivate a conclusion. Using the standard method, required sample sizes can be calculated. Using the pre-test standard deviation of 3.7 points from a 600 student sample [1], approximately 80 students would be required to see the same 1.67-point increase seen in the work with grade 9 students [8] at a 95% confidence level. 1-point discrimination at 95% confidence would require 200 students.

An uncontrolled variable in this experiment was the English proficiency of the participants. While some began speaking the language in earnest only months prior, others were functionally fluent because of years at English-medium schools. Compounding this problem is the difficult interplay between conversational and academic English skills. A further direction is to account for this factor by also testing academic English reading ability and conducting an analysis of variables.

Conclusion:

This study showed no increase in SFCI scores over FCI scores for a sample of eight high school physics students. A study using this methodology, but involving a larger sample size, would be required to compare the effectiveness of the SFCI for ELLs to the effectiveness of the SFCI for grade 9 students.

If the preliminary results of this study are upheld, it suggests that adapting the FCI to ELLs should be done in other ways, if the goal is to evaluate physics understandings, rather than ability to parse physics syntax. This might be accomplished via abundant use of cartoons or sketches, videos, or animations [11].

References:
