

Preliminary results on a video-based force concept inventory

David Rosengrant

*Department of Curriculum, Instruction and Learning, College of Education, University of South Florida, 140 7th Avenue S., Harold Heller Hall
Room 207, St. Petersburg, Florida, 33701*

Garrett Matthews

Department of Physics, College of Arts and Sciences, ISA 2019, University of South Florida, Tampa, FL 33620

Allan Feldman

Department of Curriculum, Instruction and Learning, College of Education EDU 105, University of South Florida, Tampa, FL 33620

Yi-Hsin Chen

Department of Educational and Psychological Studies, College of Education, University of South Florida, Tampa, FL 33620

Jawaher Alsultan

Department of Curriculum, Instruction and Learning, College of Education EDU 105, University of South Florida, Tampa, FL 33620

The Force Concept Inventory (FCI), one of the most widely used tools in the physics community, is commonly used as a pre and posttest assessment to gauge the effectiveness of various teaching strategies. Over the years there have been various studies on the FCI itself, and different formats have been created. In this preliminary study we are looking at the incorporation of videos that illustrate the problem statements as a way to see if these additions have positive effects on students' responses to FCI questions. We took a subset of four questions from the FCI and varied the format of the questions and videos in quizzes administered to multiple introductory algebra-based physics courses. Though there were no quantitatively significant differences in students' performance on questions between those who answered the as-written FCI questions versus the questions with the videos, we found significant differences on some questions with repeated responses to the same questions with and without video. Furthermore, students overall felt the videos helped them visualize the question scenarios.

I. INTRODUCTION

Assessment is necessary in any classroom and is a key tool for educators. It can not only provide information regarding what students know or do not know, but also inform instructors if they reach their teaching objectives and further determine if they need to modify their teaching approach in the classroom. There are many different styles of formative assessments [1-2] and assessments that focus on students' scientific abilities [3]. Sometimes, tools such as representations are used to aid in assessments [4-5]. Currently, there are many validated assessments ready to use for educators on a large number of topics that range from attitudes towards science [6] to those focusing strictly on content [7]. The Force Concept Inventory (FCI) [8] is arguably the most popular assessment available in physics.

A main reason for its popularity is the ability to use this as a pre- and post- assessment to compare learning gains among individuals because giving the pretest will not affect posttest results [9]. Researchers widely use this tool to investigate the effects of various instructional methods and how it impacts student achievement [10]. Even the inventory itself has been studied intently. Researchers [11] have turned parts of the FCI into open ended questions and found much agreement between the open-ended and multiple-choice versions of the test.

One variation of the FCI involves the use of computer animations [12]. The animations added to the FCI helped the students have a better understanding of the intent of the questions on the exam. This resulted in the students giving an answer that better reflected what the students actually believed to be the answer. Another variation focuses on context sensitivity [13]. Though there was a slight gain in scores, it was not enough to show that changing the context of the FCI warranted researchers and educators change from using the FCI as a diagnostic instrument.

This preliminary study investigates the use of videos from the real world as a possible modification to the FCI. We believed that the addition of a video segment that contextualizes the FCI question would help students make a connection between their real-world experiences and their school science experiences [14-15]. As previous work [12] has already shown that some variations such as animations can better reflect what students are thinking, we hypothesized believed that incorporating the use of videos improves the ability of the FCI to better uncover what students truly think. Thus, we conducted this preliminary study to answer these research questions: (1) *Will the supplemental videos improve student performance on the FCI?* and (2) *How do students perceive the helpfulness of videos while solving questions from the FCI?* To investigate

these questions, we selected four FCI questions to supplement with videos and looked at how the videos affected the students' responses, and if the students thought the videos helped them to answer the questions.

II. METHODS

A. Sample and setting

We conducted this study at large metropolitan Research One University. The students involved were enrolled in an introductory Algebra-based Physics 1 course. Two of the sections were large enrollment courses while a third was a smaller section. It is important to note that this study happened during the Covid-19 pandemic, thus the two large enrollment courses ($N = 168$ for class one and $N = 224$ for class two) were strictly online, while the third course ($N = 58$) was a hybrid version. The students were able to select which section they wanted to attend, but the hybrid section filled up quickly during registration. We presented the FCI questions as an assignment that served as an attendance check-in for the first week of class.

We used the revised FCI [16] for our study. As a pilot study we focused on just four questions. We chose two pairs of questions that were related in terms of their underlying physics concepts. The first pair related to Newton's third law: Question 4 – “A large truck collides head-on with a small compact car” and Question 28 - Student A places his bare feet on the knees of Student B as shown in a figure in the question. Student A then pushes off”. The second pair focused on projectile motion: Question 12 – “A ball is fired from a cannon at the top of a cliff as shown in a figure in the question” and Question 14 – “A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction”.

We searched the Internet for videos that contextualizes each of the questions. We found videos that aligned very closely to questions 4 [17] and 12 [18]. For question 14, we modified the language of the question slightly so that it was a pumpkin and not a bowling ball because the video we used showed a pumpkin being dropped out of an airplane [19]. We edited the videos so they would just frame the context of the question without providing students with information beyond what was provided in the written forms of the questions. Since we did not find any videos that related to question 28, we made our own.

Each class received four variations of the assessment, but students only filled out one variation. Table I shows how we varied the format of the questions. In each test form, the first question was the traditional written FCI without the video. The next three questions were the other FCI questions but with a video displayed at a different location in the text. The video was either before any of the text, embedded in the text, or the video followed the text.

TABLE I. Variations in assessment forms for questions 4, 12, 14, & 28.

Form	Traditional Question	Video then Question	Question then Video	Part of Question Video Rest of Question	Traditional Question with the Video as in the Previous Column
A	Q 4	Q 12	Q 14	Q 28	Q 4
B	Q 12	Q 14	Q 28	Q 4	Q 12
C	Q 14	Q 28	Q 4	Q 12	Q 14
D	Q 28	Q 4	Q 12	Q 14	Q 28

Note. Q = Question

The final question was a repeat of the traditionally written FCI question with a supplemental video. This format was the same in all four quiz forms. Thus, each test form contained a baseline of the traditional FCI question followed by four questions with a video. We asked the students to select one of the test forms based on their last name. Thus, students A-D were supposed to use Form A, and so forth, with the divisions chosen to give a roughly even number of students taking each form. After each of the five questions we asked how they came up with the answer to the previous question. We had an open-ended final question, which was: What impact (if any) did the videos have in determining your answers to the questions.

B. Data analysis

We first coded each answer as incorrect or correct. Next, we focused on the students’ repeated responses to the FCI written questions and the same questions with the video supplement to look for a gain or loss in their scores. We define a gain when they had the question incorrect initially but answered correctly in the second instance. Conversely a loss when they answered the questions correctly the first time but answered it incorrectly the second time. We also categorized it as neutral when both answers were either correct or incorrect. We conducted a dependent t-test for each question with repeated responses. We tested for significant differences in students’ performance on questions without and with videos, *i.e.* between those who only answered the written FCI questions and those who answered the same questions with videos in three different locations. We conducted an independent t-test for each question.

Next, we focused on their responses to the open-ended questions. We initially coded each question based on two criteria. The first was if they specifically mentioned the video in their answer. The second was if we inferred that they were referring to the video. One example for an inference was, in question 12 (A cannonball is shot out of a cannon on a hill), if they wrote something like “the ball left the cannon at a high rate of speed.” We then compared the number of correct to incorrect responses in each category.

We then focused on the open-ended question: “What impact (if any) did the videos have in determining your answers to the questions?” We coded answers into a positive (meaning the students said the videos helped them understand the problem or that they were able to visualize the situation because of it), negative (a student could say that

the videos made it more confusing for them), neutral (the videos did not help or hurt their understanding or they had issues with a video loading), or other (their comment did not reflect upon their use of the videos or it was left blank). For the number of students correctly answering different numbers of video modified questions, we analyzed whether they indicated the videos had a positive impact (thus excluding the written only response).

III. RESULTS

Table II focuses on the questions that were first asked without the video and then again with it. For overall percentage correct, paired t-tests indicated that Q12, Q14, and Q28 showed significant differences. Q14 and Q28 had slight increases when the video was present, however, there was the largest decrease with Q12. As for gain and loss, McNemar tests showed a significant difference on Q12, indicating a larger number of loss than that of gain when video was present. Q4, Q14, and Q28 did not show significant differences between the numbers of gain and loss even though the numbers of gain were slightly greater than those of loss. Overall, Q12 had a significantly negative impact, which means the video added in this question confused students. Videos added in Q14 and Q28 slightly improved students’ performance on these questions, but the numbers of gain and loss were equivalent. The video added in Q4 did not make a significant impact on students’ performance and the numbers of gain and loss.

TABLE II. Questions written out then with video

Question	Number of Students with Change	Overall Correct (%)	McNemar test (df = 1)
Q4 Written	Gain = 7	37 of 92 (40.2)	$\chi^2 = 0.82$
Q4 Video	Loss = 4 Neutral = 81	40 of 92 (43.5)	$p = 0.37$
Q12 Written	Gain = 2	73 of 122 (59.8)	$\chi^2 = 12.80$
Q12 Video	Loss = 18 Neutral = 102	57 of 122 (46.7)	$p < .001$
Q14 Written	Gain = 8	36 of 109 (33.0)	$\chi^2 = 3.60$
Q14 Video	Loss = 2 Neutral = 99	46 of 109 (42.2)	$p = .06$
Q28 Written	Gain = 20	41 of 129 (31.8)	$\chi^2 = 0.44$
Q28 Video	Loss = 16 Neutral = 93	45 of 129 (34.9)	$p = .50$

In addition to the paired analyses in two different formats, Table III shows the percentages of correct for each original FCI question without video and 3 same questions with videos. Independent t-tests indicated that there were no significant differences between students' performances on questions without and with videos for these questions.

TABLE III. Percent correct of questions with video

Qstn.	Without Video	With Video	t-test
Q 4	40.2 (37 of 92)	38.0 (135 of 355)	$t=.38, p=.70$
Q 12	59.8 (73 of 122)	55.0 (182 of 331)	$t=.92, p=.36$
Q 14	37.3 (41 of 110)	39.8 (134 of 337)	$t=-.46, p=.64$
Q 28	31.8 (41 of 129)	31.3 (99 of 316)	$t=.09, p=.93$

Note that in Table IV we did not separate out the questions in our analysis for the cases where they had the same question written out and then used as a video. In this situation, we analyzed the second question, which included the video.

TABLE IV. Students Comments about video

Question	Video Code	Responses per Code
Q 4	Mentioned – 28	Correct – 9 (32%) Incorrect – 19
	Inferred – 31	Correct – 13 (42%) Incorrect – 18
Q 12	Mentioned – 22	Correct – 6 (27%) Incorrect – 16
	Inferred – 21	Correct – 1 (5%) Incorrect – 20
Q 14	Mentioned – 88	Correct – 32 (36%) Incorrect – 56
	Inferred – 16	Correct – 5 (31%) Incorrect – 11
Q 28	Mentioned - 51	Correct – 4 (8%) Incorrect – 47
	Inferred - 85	Correct – 38 (45%) Incorrect – 47

With regards to the last open-ended question, we assigned a +1 to those who said that the videos helped them in some way. We then assigned a -1 to those who said the videos had no impact or that the videos hurt their understanding of the situation. If the videos did not load, they did not watch them or they did not have a comment relating to the videos, we scored that as a zero. Then, for each score on the number of correct responses on the questions with videos ($N = 100$ for 0, $N = 152$ for 1, $N = 82$ for 2, $N = 68$ for 3, and $N = 52$ for 4) we found the average 'helpfulness' score, which is the net positive impact (Table V), for each group in that category. We calculated the helpfulness score by subtracting the negative proportion from the positive proportion. Our results are shown in Table V.

The majority of the students' responses ($N=243$) in the open-ended questions stated that these videos were beneficial due to the visualization factor. Thirty-five

students stated they felt the videos helped them better understand the concept. Fifteen students felt it facilitated their decision-making process while six felt it made the question easier. Twenty-four students simply stated the videos helped them but did not clarify how it helped them. The comments that were more negative in nature included 17 students who said the videos were confusing.

TABLE V. Questions only with video format

Number Correct	Proportion Found Videos Positive / Helpful	Proportion Found Videos Negative / Confused	Net Positive Impact
0	0.80	0.08	0.72
1	0.81	0.11	0.70
2	0.77	0.16	0.61
3	0.75	0.13	0.62
4	0.60	0.33	0.27

We then graphed these results. Figure 1 shows the plots of this analysis. In the top figure, we plotted the proportion of students with each score who found the videos helpful (positive impact) versus their score, shown in blue. We repeated this with the proportion of students with each score who found the videos harmful (negative impact) versus their score, shown in red. In the bottom plot of Fig. 1, we show the net positive impact (positive impact - negative impact) for the students with each score versus their score.

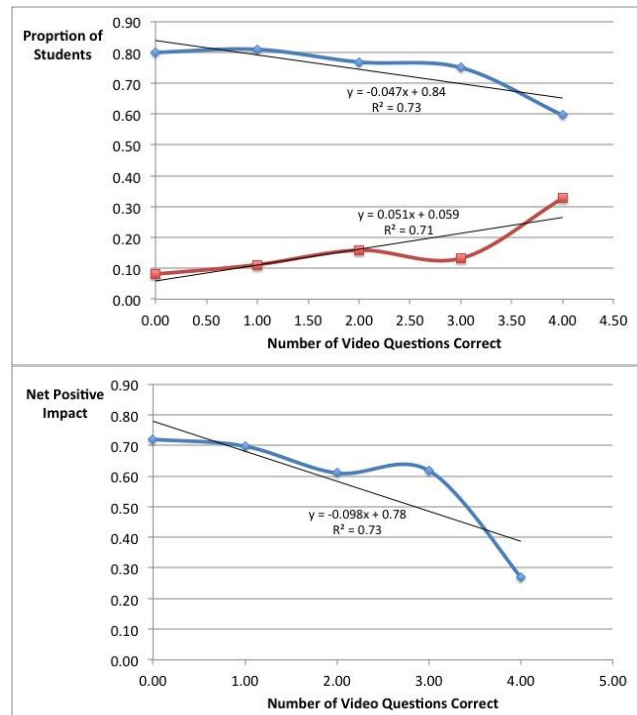


Figure 1: Comparison of helpfulness of videos to scores.

IV. DISCUSSION

We first address our question as to whether the supplemental videos improve student performance on the FCI. When Dancy and Beichner [11] replaced the solutions to the FCI with animations, they found that the students gave answers that were more aligned with their thoughts and beliefs. Though we made the videos as a supplement to the question stem as opposed to a replacement for the answer, we found very similar results. Out of the four problems, three of them showed a gain for students correctly answering the video enhanced question, but only two of three were significant changes (Table II). The greatest change was a negative one in that more students changed from a correct to an incorrect solution. Though these were overall changes, every question had some students go from incorrect to correct and vice versa. This caused us to go deeper in the data.

Globally, students expressed that they found the videos to be helpful: roughly 77% indicated that the videos had a positive impact as compared to 14% who thought the videos had a negative or no impact. Students thought that the videos helped them understand and, especially, to visualize the situation that the problem presented. However, this sense of there being a positive impact from the videos was not reflected in improved scores. Indeed, students who performed better on the questions found the videos to be less helpful.

One interpretation of this observation is that what the videos did was to reinforce their erroneous beliefs. For example, consider question 4 in which a large truck collides with a small car. Students see in the video what they would expect from everyday experience: the small car receives greater damage than the large truck. From this viewing, they conclude that the truck exerted a larger force during the collision, counter to the correct Newton's 3rd Law interpretation of the problem. This mistake was apparent from some of the students' comments regarding the video in this question.

Alternatively, one might expect that more naïve students would find comfort in the presence of the videos, reducing their need to mentally picture the action in the question. They would be more likely to say that the videos had a positive impact on their ability to answer the questions. At the same time, these naïve students would be expected to perform more poorly. With this interpretation, the videos didn't mislead the students, but rather led to answers that better reflected their beliefs. This interpretation is supported by the comparable results from the other Newton's 3rd Law question, question 28 (Table III).

The results from questions pertaining to projectile motion, questions 12 and 14 in Table III, show that students

correctly chose the path of the fired cannon ball more often than they did the pumpkin dropped from a moving airplane. The perspective shown in the cannon video matched the perspective being used in the question. For question 14, as observed from the airplane and as shown in portions of the video, the pumpkin would seem to fall straight downward, though the question asks about motion from the perspective of an observer on the ground. Some students' comments seemed to indicate that they took the perspective from the airplane, potentially because this perspective was reinforced by the video.

Our second question was how the students perceived the helpfulness of videos while solving questions from the FCI. Over two-thirds of the students stated that they liked the videos (Table IV). However, the more students thought the video helped them visualize the question scenario, the lower they scored in correctly answering the questions. Thus, it appears that the videos reinforced the beliefs of those students who had incorrect notions about the physics concepts. However, our goal was to better uncover students' beliefs about physics concepts, which may not result in a higher percentage of correct answers to the FCI.

In terms of the videos themselves, naturally they need to be directly related to the nature of the question. Students keyed in to certain details that the more physics savvy would ignore, such as the car taking more damage or that the cannon was large and "exerted a great deal of force." The videos also varied in that they may have shown the entire question (the cars colliding) or everything up to the answer (the pumpkin being dropped out of the plane). However, the video that students found least helpful was one which had subtle differences between it and the question. The question that had male and female students with a sizeable mass difference sitting in chairs and pushing off from one another, we could not replicate exactly. Thus, students commented on the differences and how it was not as helpful.

V. CONCLUSIONS

Clearly, there is more work that needs to be done. Future studies will involve more questions with the video formats and follow up interviews so we can determine whether the videos better reflect the students' beliefs. Even with future work needed, this study showed that videos have the potential to provide the more naïve students with a better mental picture of the questions, thereby better assessing their comprehension of the question scenario. We want to assess students understanding of the physics in a real-world context, not in a hypothetical environment created for an exam. Thus, if the students better comprehend the question scenario; we can then get a more accurate assessment of their understanding of the physics concepts.

-
- [1] W.W. Cobern, D. Schuster, B. Adams, B.A. Skjold, E.Z. Muğaloğlu, A. Bentz, and K. Sparks, *International Journal of Science Education*, 36(13), 2265-2288 (2014).
- [2] M. Yorke, *Higher Education*, 45(4), 477-501 (2003).
- [3] E. Etkina, A. Van Heuvelen, S. White-Brahmia, D.T. Brookes, M. Gentile, S. Murthy, D. Rosengrant, and A. Warren. *Physical Review special topics-physics education research* 2, 020103 (2006).
- [4] D. Rosengrant, *The Physics Teacher*, 49(1), 36-39 (2011).
- [5] D. Rosengrant, E. Etkina, and A. Van Heuvelen, in *AIP Conference Proceedings, 2007*, Vol. 883, No. pp. 149-152.
- [6] K. Semsar, J.K. Knight, G. Birol, and M.K. Smith, *CBE—life sciences education*, 10(3), 268-278 (2011).
- [7] C. Singh and D. Rosengrant, in *Proceedings of the Physics Education Research Conference* (Rochester, NY, 2001), edited by S. Franklin, J. Marx, and K. Cummings.
- [8] D. Hestenes, M. Wells, and G. Swackhamer, *The physics teacher*, 30(3), 141-158 (1992).
- [9] C. Henderson, *The Physics Teacher*, 40(9), 542-547 (2002).
- [10] D. Huffman, and P. Heller, *The physics teacher*, 33(3), 138-143 (1995).
- [11] S.N. Rebello, and D.A. Zollman, *American Journal of Physics*, 72(1), 116-125 (2004).
- [12] M.H. Dancy, and R. Beichner, *Physical review special topics-physics education research*, 2(1), 010104 (2006).
- [13] J. Stewart, H. Griffin, and G. Stewart, *Physical Review Special Topics-Physics Education Research*, 3(1), 010102 (2007).
- [14] B. Stengel, *Journal of Curriculum Studies* 29, 1 (1997).
- [15] H. Munby, M. Cunningham, and C. Lock, *Science Education*, 84(2), 193-211 (2000).
- [16] I. Halloun, R. Hake, E. Mosca, and D. Hestenes, (1995). Retrieved 4/10/2021. <http://modeling.asu.edu/R&E>.
- [17] <https://www.youtube.com/watch?v=hx1mBKeUeLo> Retrieved 6/2/2021
- [18] <https://www.youtube.com/watch?v=tKN37in85II> Retrieved 6/2/2021
- [19] <https://www.youtube.com/watch?v=orRjAZ8XasM> Retrieved 6/2/2021