SURVEY OF STUDENT UNDERSTANDING OF ELECTRIC FORCE AND FIELD

Santosh Kumar Umar*, Ashok Kumar Mittal and Vivek Kumar Tiwari
Physics Department, University of Allahabad, Prayagraj, U.P., India
santosh.91k@hotmail.com

The basic purpose of this study is to find whether students can identify and overcome their deficiencies in a narrow range of concepts through self-reflection, self-study and self-sought discussion, induced by a large number of practice question focused on the concepts. For this purpose, a narrow range of concepts pertaining to electric force and field were tested. The pre-test consisted of the ten conceptual multiple-choice questions pertaining to electric force and field in the well-known Conceptual Survey of Electricity and Magnetism (CSEM). After pre-test was administered, a large number of similar questions were sent to the students by e-mail and they were advised to ensure that they can answer such questions before a post-test to be administered within a few weeks. This survey covers more than 1000 Physics students at different levels of physics learning at different institutions.

INTRODUCTION

Physics is a concept-based subject. A central goal of Physics Education is to develop a proper conceptual understanding of students on different topics. Physics Education Research (PER) as a field of study emerged to provide reproducible and quantifiable conclusions related to conceptual understanding. It has created a large body of reliable knowledge to help improvement in physics instruction and techniques. Physics Education Researchers have developed several standardized concept-tests in different topics in Physics that have to satisfy stringent requirements of reliability and validity (Adams & Wieman, 2011; Engelhardt, 2009). These tests are administered before instruction and after instruction. The learning gains on such concept tests help compare the effectiveness of different instruction strategies. The standard definition of learning gain in PER is the ratio of 'increase in percentage score' and '(100 – percentage pre-score)' the latter being the maximum possible increase in the percentage score. PER studies have consistently demonstrated that interactive engagement techniques lead to significant gains in conceptual learning as compared to traditional lecture-based instruction (Hake, 1998) leading to the introduction of reformed Physics programs at several places (Crouch & Mazur, 2001). The Force Concept Inventory (FCI) (Hestenes, Wells & Swackhamer, 1992), Force and Motion Conceptual Evaluation (FMCE) (Thornton & Sokoloff, 1998), Brief Electricity and Magnetism Assessment (BEMA) (Chabay & Sherwood, 1997) and Conceptual Survey of Electricity and Magnetism (CSEM) (Maloney, O'Kuma, Hieggelke, & Van Heuvelen, 2001) are amongst the concept tests which have been widely used for investigating students' understanding. In most studies, the pre-tests are administered in the initial weeks of a semester course and post-tests towards the end of the semester. Concept tests such as BEMA and CSEM cover many sub-topics and concepts in Electricity and Magnetism. Such coarse-grained tools can help compare broadly the efficacy of different instruction strategies, but are not useful for identifying conceptual deficiencies in an individual student and for taking remedial action. For this purpose, it is necessary to administer fine-grained concept tests and reduce the time lag between pre- and post-tests. Here, the pre-test and the post-test mean tests administered before and after a remedial intervention. Accordingly, the ten questions in CSEM pertaining to electric force and field were taken as constituting a fine-grained pre-test. This subset (CSEFF = Conceptual Survey of Electric Force and Field) consists of ten multiple-choice questions (or items). One of the reasons, for choosing CSEM was that the performance on individual questions in pre- and post-tests in the CSEM are available for comparison with our survey, even though the meaning of pre- and post-test in CSEFF is different. The broader purpose of our studies is to quantify the impact of different types of remedial actions for overcoming conceptual deficiencies. In today's world it is easy to get information, therefore an important, perhaps the most important, learning goal is to develop the ability to answer questions through self-study, through self-sought discussion with peers and instructors, and through self-reflection. This ability is expected to improve with the advancement of learning stages. Unlike FCI, CSEM is a broad survey instrument. There cannot be a single concept inventory for such a broad range of topics. However, CSEFF, a sub-set of CSEM, can serve as a single concept inventory for Electric Force and Field, comparable to FCI. In the present study, after administering the pre-test, a large number of similar questions were sent to the students by e-mail and they were advised to ensure that they can answer such questions before a post-test to be administered within a few weeks. The pre-test, post-test and the set of intervention practice questions can be obtained by email. In this way, an attempt is made to evaluate the unsupervised self-learning skills that students have developed for answering a class of similar questions pertaining to a very narrow range of concepts. In addition, the survey was also carried out over participants in a refresher course for college and university teachers from all over India, held at the Physics Department, a leading Indian University. All the participants had Ph.D. degrees and varying lengths of teaching experience. They are well set on the path of Physics learning and teaching as a profession. Compared to students, they can be expected to have better self-learning skills and therefore better ability to overcome any deficiencies on the pre-test.

The following research questions are addressed in this study:

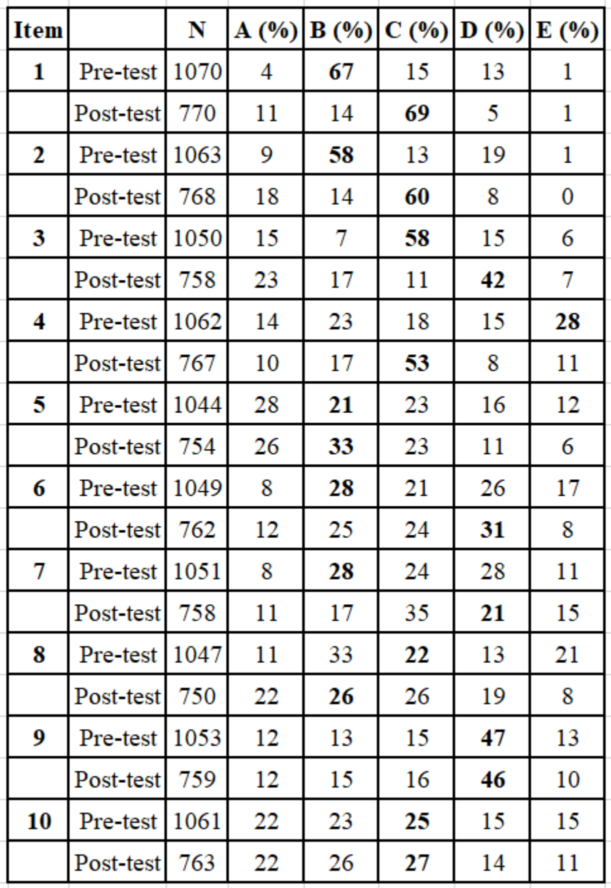
- How do statistical test characteristics of CSEFF compare with recommended values?
- How do the option choices for different items vary with the ability of students?
- To what extent does the unsupervised intervention of self-directed study, induced by several similar problems, help improve conceptual understanding of narrowly focussed concepts?
- What conceptual difficulties are indicated by CSEFF and how do they compare with those reported for CSEM?
- To what extent does the conceptual understanding of electrostatic force and field improve with learning stages from B.Sc. I Year to M.Sc. Final.

METHODOLOGY

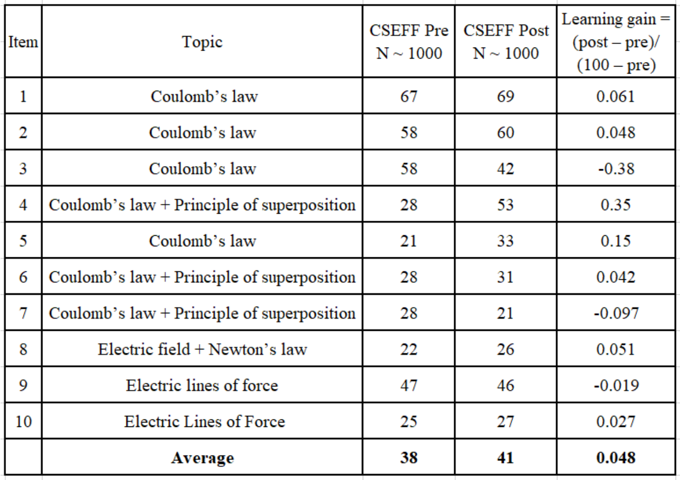
The CSEFF was administered as a pre-test to students at various stages of physics learning from a B.Sc. First Year to M.Sc. Final Year and to participants of the refresher course, without any pre-announcement. The reason for not giving any advance notice was to evaluate the internalization of concepts and not memorization for a test. After the completion of the pre-test, a set of 57 problems were sent to the students by email. Twenty of the self-study questions (two for each item on the pre-test) were minor variants of the ten items on the pre-test, therefore presumably had the same difficulty level. The remaining were a mix of simpler and more complicated versions of the items on the pre-test. The answers were not given, as the purpose was to evaluate problem-driven self-learning skills. Students were advised to ensure that they can answer such questions before a post-test to be administered any time between one to four weeks. A post-test consisting of questions very similar to the pre-test was administered, again without announcing the exact date of such test. Each item in the post-test was a minor variant of the corresponding item in the pre-test, allowing the computation of item-wise learning gains.

DATA ANALYSIS

Table 1 shows the percentage of students who chose different options for each item on CSEFF. Table 2 shows item-wise percentage correct response on pre-test and post-test. It also shows the learning gains for each item. Table 3 shows the item-wise correct response at different stages of learning. It also includes a column of the reported response in the pre-test given to introductory students in US universities.

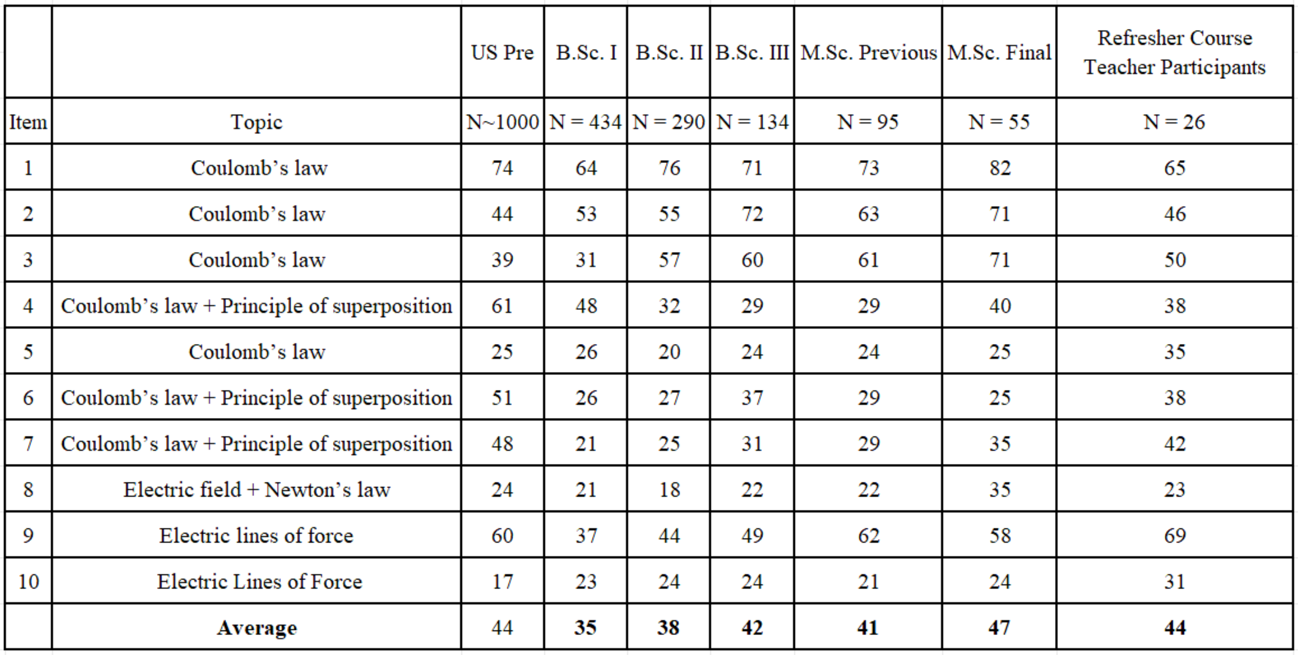


Item		**N**	**A (%)**	**B (%)**	**C (%)**	**D (%)**	**E (%)**
1	Pre-test	1070	4	**67**	15	13	1
	Post-test	770	11	14	**69**	5	1
2	Pre-test	1063	9	**58**	13	19	1
	Post-test	768	18	14	**60**	8	0
3	Pre-test	1050	15	7	**58**	15	6
	Post-test	758	23	17	11	**42**	7
4	Pre-test	1062	14	23	18	15	**28**
	Post-test	767	10	17	**53**	8	11
5	Pre-test	1044	28	**21**	23	16	12
	Post-test	754	26	**33**	23	11	6
6	Pre-test	1049	8	**28**	21	26	17
	Post-test	762	12	25	24	**31**	8
7	Pre-test	1051	8	**28**	24	28	11
	Post-test	758	11	17	35	**21**	15
8	Pre-test	1047	11	33	**22**	13	21
	Post-test	750	22	**26**	26	19	8
9	Pre-test	1053	12	13	15	**47**	13
	Post-test	759	12	15	16	**46**	10
10	Pre-test	1061	22	23	**25**	15	15
	Post-test	763	22	26	**27**	14	11



Item	Topic	CSEFF Pre N ~ 1000	CSEFF Post N ~ 1000	Learning gain = (post – pre)/ (100 – pre)
1	Coulomb's law	67	69	0.061
2	Coulomb's law	58	60	0.048
3	Coulomb's law	58	42	-0.38
4	Coulomb's law + Principle of superposition	28	53	0.35
5	Coulomb's law	21	33	0.15
6	Coulomb's law + Principle of superposition	28	31	0.042
7	Coulomb's law + Principle of superposition	28	21	-0.097
8	Electric field + Newton's law	22	26	0.051
9	Electric lines of force	47	46	-0.019
10	Electric Lines of Force	25	27	0.027
	Average	**38**	**41**	**0.048**

Table 1: Percentage of student chosing different option, **Table 2:** Learning gain



		US Pre	B.Sc. I	B.Sc. II	B.Sc. III	M.Sc. Previous	M.Sc. Final	Refresher Course Teacher Participants
Item	Topic	N~1000	N = 434	N = 290	N = 134	N = 95	N = 55	N = 26
1	Coulomb's law	74	64	76	71	73	82	65
2	Coulomb's law	44	53	55	72	63	71	46
3	Coulomb's law	39	31	57	60	61	71	50
4	Coulomb's law + Principle of superposition	61	48	32	29	29	40	38
5	Coulomb's law	25	26	20	24	24	25	35
6	Coulomb's law + Principle of superposition	51	26	27	37	29	25	38
7	Coulomb's law + Principle of superposition	48	21	25	31	29	35	42
8	Electric field + Newton's law	24	21	18	22	22	35	23
9	Electric lines of force	60	37	44	49	62	58	69
10	Electric Lines of Force	17	23	24	24	21	24	31
	Average	44	**35**	**38**	**42**	**41**	**47**	**44**

Table 3: Item wise correct response at different stages of learning

Statistical characteristics

Some of the statistical characteristics (Ding, Chabay, Sherwood & Beichner, 2006), that are often used in PER for ensuring their reliability and validity are item difficulty index, item discrimination index, point-biserial index, KR-21 reliability index, Ferguson delta. It is necessary to evaluate these characteristics for the CSEFF as they depend also on the student population.

Difficulty index of an item: The difficulty index of an item in the test is defined by;

$$P = N_1/N \tag{1}$$

Here N is total number of students and N_1 is the number of students choosing the correct option, thereby scoring 1 for the item. The difficulty index for different items of CSEFF is shown in Fig 1(a). None of the items has difficulty index less than 0.2 or more than 0.7. The average difficulty index for the whole test is 0.38.

Discrimination index of an item: The discrimination index of an item gives information about the contribution of the item for discriminating between high scoring (> 75 percentile) and low scoring (< 25 percentile) students on the entire test. It is defined by

$$D = 4(N_{H-} - N_L)/N \tag{2}$$

where N_H and N_L are the number of correct responses to the item by the high scoring and low scoring groups respectively. According to classical test theory, items (e.g. 1, 2 and 3) with difficulty index around 0.5 are likely to have high discrimination index and items with difficulty index away from 0.5 (whether on the high side or the low side) are likely to have low discrimination index. Fig.1(b) shows the discrimination index for different items of CSEFF. The average value of the discrimination index is 0.48. An item is considered to have satisfactory discrimination if the discrimination index is greater than 0.4. Only items 8 and 10 failed this condition.

Point Biserial reliability Index: The point biserial index, a measure of the consistency of one item with the whole test, is defined by the correlation between the item score and the total score. It is given by

$$r_{pbs} - \frac{\bar{X}_1 + \bar{X}}{\sigma x}\sqrt{\frac{P}{1-P}} \qquad (3)$$

where $\bar{X}_1$ is the total score average for the students who answered the item correctly, $\bar{x}$ is the total score average for the whole test, σx is the standard deviation for the whole test and P is the difficulty index. Fig.1(c) shows the point biserial coefficient for each item of CSEFF. An item is considered good if its point biserial is > 0.25 and the recommended average value is ≥ 0.2. For CSEFF, all items had point biserial > 0.3 with an average of 0.4.

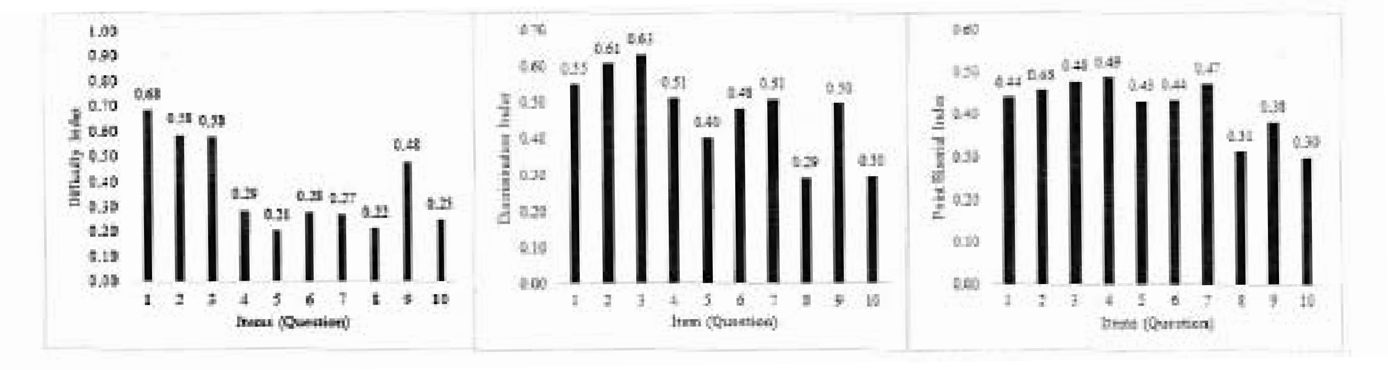


Figure 1: (a) Difficulty index, **Figure 1(b):** Discrimination index, & **Figure 1(c):** Point Biserial index

KR-21 test reliability index for the whole test: The Kuder Richardson reliability index measures the self-consistency of the whole test. KR-21 reliability index is the average correlation between all possible divisions into two parts, each having half the number of items. For a dichotomous multiple-choice test, it is given by

$$r_{test} = \frac{K}{K-1}\left(1 - \frac{\sum P(1-P)}{\sigma_X^2}\right) \qquad (4)$$

Here K is the total number of test items in the whole test, P is difficulty index and σx is the standard deviation of the total score. The recommended value for a good test is $r_{test} \geq 0.8$. For CSEFF the value of is 0.98.

Ferguson delta for the whole test: Ferguson Delta is a measure of the discriminatory power of the whole test. Let, be the number of students whose total score is i. Ferguson's delta is defined by;

$$\delta = \frac{K+1}{K}\left\{1-\sum_{i=0}^{R}\left(\frac{f_i}{N}\right)^2-\right\} = \frac{N^2-\sum_i^k =_0 \; f_1^{\,2}}{N^2-\frac{N^2}{K+1}} \tag{5}$$

For a good test, it is recommended that Ferguson delta should be > 0.9. The value of Ferguson's delta for CSEFF is 0.93.

Item response curves

Table I shows item-wise percentage responses to different options both for the pre-test and post-test. The correct option is shown in bold. This table gives the overall performance of all students without distinguishing between the varying abilities of students. The total score on the test may be taken as a measure of the ability of students. The response data is separated for students of varying ability as indicated by the total score. A table, like Table I, is prepared separately for each value of the total score. The data for all such tables is represented in the form of Item Response Curves (IRC) (Morris et al., 2006; Morris et al., 2012) as shown in Fig 2(a) for pre-tests and 2(b) for post-tests. As expected, with increasing ability the IRC's show an increasing trend for the correct option and a decreasing trend for the wrong options. For low scoring items like 5, 7, 8 and 10, the IRC's of the correct options start from 0 and rise steeply only at mid or high ability levels. For high scoring items like 1, 2 and 3, they start at relatively high values and rise slowly to flatten out as mid and high ability students have a high probability of answering them correctly.

Impact of interventional questions

The extent to which unsupervised intervention of self-directed study, induced by several similar problems, helped improve conceptual understanding of electrostatic force and field is shown in Table II. The average learning gain is a modest 0.048. Most of the items showed very little change. Item 4 was the only item that showed substantial learning gain. There was a substantial decrease in the correct option for item 3 and an increase in the wrong option (a). In the post-test, the distance was decreased, instead of increasing. It seems that several students still have problems in dividing by a fraction, causing a substantial increase in option (a). In figure 2(c) we show for each item, the correct option IRC's for pre- and post-tests juxtaposed together. Item 7 in figure 2(c) reveals a surprising and inexplicable feature: low ability students show improvement in the post-test, whereas high ability students show a decline.

Identification of misconception

As in CSEM, item1 was the easiest item in CSEFF for students of different abilities and at different learning stages in pre-tests as well as post-tests. Analysis of distractor responses to this item throw some surprising insight about possible misconceptions. About 15% answered that the force on a charge remains unchanged if the magnitude of the other charge is increased by a factor of 4. About an equal number answered that the force will decrease by a factor of 4. These answers are far more frequent than the other wrong options for no apparent reason. This trend is seen for items 1 and 2 in the pre-tests as well as the post-tests. This suggests that the wrong options are the consequence of misconceptions rather than random choices.

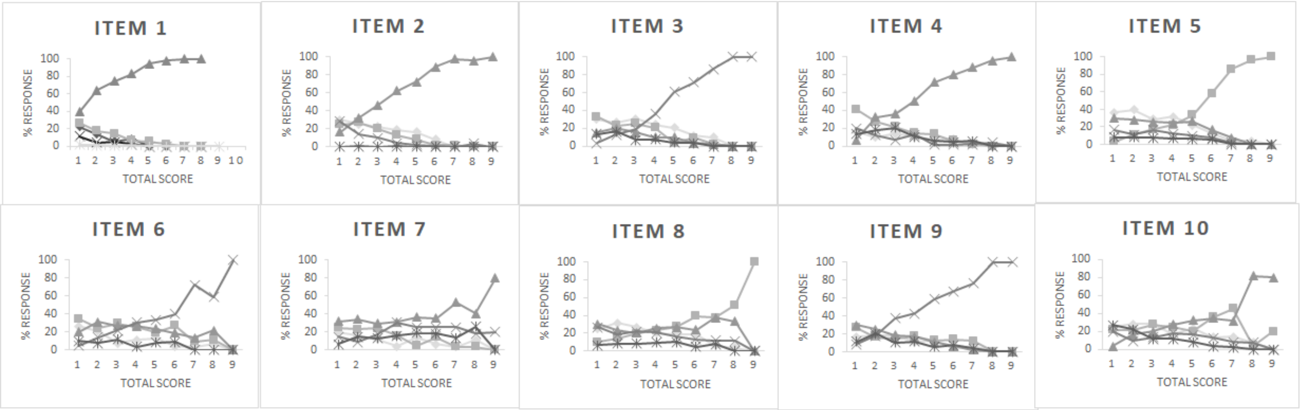


Figure 2(b) Post-test IRC's [✧ is option (a), ☐is option (b), Δ is option (c),x is option (d) and * is option (e)]

Figure 2(c) [◇for pre-test and □ for post-test]

In CSEM, the number of correct responses to item 2 were about 33% less than to item 1, although from an expert perspective, the two items are almost identical. Commenting on this result, it was remarked in (Maloney, *et al.*, 2001) that many students do not apply Newton's Third Law or symmetry of Coulomb's Law, that students seemed to believe that larger objects (in this case larger charges) exert larger forces than smaller objects. In CSEFF, the number of correct responses to item 2 were only about 14% less than to item 1. Responses to items 1, 2 and 5 do not support the view that students believe that greater charges exert greater forces. The distribution of wrong responses in CSEM is very different from those for CSEFF. Item 3 tests the understanding of distance dependence of electrostatic force between two charges. The problem with this item is that it is not independent of item 2. Those who get answer to item 2 wrong due to some misconception, cannot answer item 3 correctly. In CSEM, the number of persons who answered item 3 correctly was slightly less than those who answered item 2 correctly, suggesting that some students have problems with the inverse square dependence of the Coulomb's Law. In the CSEFF pre-test, however, both items 2 and 3 were answered correctly by the same percentage of students, suggesting that they do not have problems with the inverse square dependence.

In CSEFF, item 4 was answered correctly by only 28% students in the pre-test and 53% students in the post-test. This was the only item with substantial learning gain. From an expert perspective, items 6 and 7 are the most difficult in CSEFF, as they involve relatively lengthy problem description, intermediate stage inference and reasoning skills. These items were correctly answered by 28% students only, substantially less than CSEM students.

Items 8 and 10 were the most difficult in our survey, with difficulty index less than 0.25. Table III shows that in the reported CSEM studies also, pre-test scores on items 8 and 10 had difficulty index less than 0.25. As in CSEM, the number of students in CSEFF who answered item 8 correctly was much less than those who answered item 9 correctly, suggesting that many students still thought that application of constant force on a body would give rise to constant velocity.

CONCLUSION

CSEFF is a reasonable fine-grained instrument for assessment of conceptual understanding of electric force and field, which satisfies the statistical parameters recommended for concept tests. Item Response curves (Figure 2) provide graphs of how student performance on different options vary with ability. These graphs are easy to interpret and provide useful insight. The average learning gain due to self-learning questions was a modest 0.048. Only item 4 showed substantial learning gain. This could be because the correct option in the pre-test was 'none of these' which students tend to avoid. As argued in subsection III, substantial negative learning gain for item 3 could be due to students having problem in dividing by a fraction. Inexplicably, the performance of low ability students on item 7 improved in the post-test, whereas that of high ability students deteriorated. A serious limitation of our study is that it relies wholly on intrinsic motivation for students to try and do well in the post-test. In future studies we plan to introduce some extrinsic motivations so that students work on the intervention problems more seriously towards better performance on the post-test. It seems that a large number of students have problems with the principle of superposition and/or vector addition. A finer grained survey will be conducted for understanding the detailed nature of the misconceptions related to the principle of superposition as a follow-up of the present survey. Compared to CSEM, performance of CSEFF students was substantially poorer on problems involving lengthy description, intermediate stage inference and reasoning skill. Overall performance of students in CSEFF was found to increase slightly with advancement of learning stages from BSc I year to M. Sc. Final. The table shows a significant improvement only at the M.Sc. Final stage. This shows that students do not improve their understanding of electrostatic force and field with advancement of learning stages. Improvement at the final year stage could be because students start preparing for entrance tests for research programs, which test understanding of basic physics concepts better than the more memory-based tests in their regular courses. However, the performance of teacher participants in the Refresher Course was very disappointing and should be a cause of concern for Physics Education in India. Inferences drawn from CSEFF are suggestive and require carefully planned studies for confirmation.

ACKNOWLEDGEMENTS

Authors are thankful to the Physics Department, University of Allahabad and its affiliated Colleges for giving us the opportunity to carry out the survey. One of the author SKU is thankful to UGC, New Delhi for providing him financial assistance.

REFERENCES

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert like thinking. *International Journal of Science Education*, *33*(9), 1289-1312.

Chabay, R., & Sherwood, B. (1997). Qualitative understanding and retention. *AAPT Announcer*, *27*(2), 96.

Ding, L., Chabay, R., Sherwood, B., & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical review special Topics-Physics education research*, *2*(1), 010105.

Engelhardt, P. V. (2009). An Introduction to Classical Test Theory as Applied to Conceptual Multiple-choice Tests in Getting Started in PER, edited by C. Henderson and KA Harper, American Association of Physics Teachers, College Park. *MD Reviews in PER*, *2*.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics*, *66*(1), 64-74.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The physics teacher*, *30*(3), 141-158.

Maloney, D. P., O’Kuma, T. L., Hieggelke, C. J., & Van Heuvelen, A. (2001). Surveying students’ conceptual knowledge of electricity and magnetism. *American Journal of Physics*, *69*(S1), S12-S23.

Morris, G. A., Branum-Martin, L., Harshman, N., Baker, S. D., Mazur, E., Dutta, S., ... & McCauley, V. (2006). Testing the test: Item response curves and test quality. *American Journal of Physics*, *74*(5), 449-453.

Morris, G. A., Harshman, N., Branum-Martin, L., Mazur, E., Mzoughi, T., & Baker, S. D. (2012). An item response curves analysis of the Force Concept Inventory. *American Journal of Physics*, *80*(9), 825-831.

Thornton, R. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton’s laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. *american Journal of Physics*, *66*(4), 338-352.