Student Perceptions of Physics by Inquiry at Ohio State

Gordon J. Aubrecht, II[*†], Yuhfen Lin[†], Dedra Demaree[†], David Brookes[**], and Xueli Zou[‡]

[*]Ohio State University, Marion, OH 43302
[†] Ohio State University, Columbus, OH 43210
[**] Rutgers, The State University of New Jersey, NJ 08854
[‡] California State University, Chico, Chico, CA 95929

Abstract. Physics by Inquiry (PbI) has been adopted and taught at the Ohio State University for more than a decade. A Q-type instrument, the LPVI, was used to assess students' perceptions of this teaching method and measure the consonance between these and instructors' goals. We present methods of analysis to make use of all the information collected with this survey. In applying the LPVI to different sections of the PbI course, we found many similarities in students' perceptions and also some interesting differences. We also found similarities and differences between students' perceptions and the goals of PbI.

Keywords: physics inquiry, preservice teachers.
PACS: 01.40.Fk, 01.40.Gm, 01.40.Jp

INTRODUCTION

At the Ohio State University, Physics by Inquiry (PbI) is currently implemented in three courses using the books developed by the Physics Education Group at the University of Washington [1, 2]. Data gathered from two of these courses is presented in this paper. We use a Q-type instrument called the Laboratory Program Variables Inventory (LPVI) originally developed by Abraham [3] for assessing chemistry laboratories. The LPVI was modified for use in physics as reported in [4, 5, 6]

Q methodology was first introduced by Stephenson in 1935 [7] to "reveal subjective structures, attitudes, and perspectives from the standpoint of the person or persons being observed.[8]" A detailed presentation of Q methodology can be found in [9]. A general feature of Q methodology is that statements are sorted into rankings following a quasi-normal distribution. The sorting process is called Q-sorting. The modified LPVI contains twenty-five descriptive statements that could be applicable to most laboratory-based classrooms. These statements are presented in the appendix of [4].

PbI is intended to achieve multiple goals. Students should learn some physics material, be introduced to how science works, use evidence to support claims, be able to make the connection between models and measurement, and use the models developed to help explain phenomena in context. The main goal of PbI is not only to teach concepts but to help students understand the nature and methods of science. Regular standardized tests merely focus on rote or on conceptual understanding and are thus possibly not the most appropriate measure of assessment of student accomplishment in PbI. The modified LPVI, on the other hand, can give us some insight into the attitudes and perceptions of students. The LPVI can help us see if students recognize the overarching goals of the course.

The goal of this paper is to show how the modified LPVI can give us insight into the perceptions of students as to whether the PbI course in achieving its instructor-intended aims, although, of course, it cannot show whether students have learned the content.

METHOD

Students were instructed to sort the 25 statements of the modified LPVI into five groups. Group I is considered most descriptive of the course; Group V least descriptive. Students are forced to rank the statements into groups of size 2, 6, 9, 6, 2, forming a quasi-normal distribution. For this study, 84 students in Physics 107 (the batteries and bulbs segment) from the Ohio State University main campus (3 sections), and Marion campus (1 section) completed the Q-sort. Ten Physics 107 instructors also completed the sorting, but their data are not analyzed here. Thirty students from Physics 108 (the geometric optics segment plus some additional materials) also completed the Q-sort (2 sections from Columbus, 1 from Marion).

RESULTS AND ANALYSIS

First we discuss the five most and least characteristic statements as chosen by students and construct a normalized matrix and perform a χ^2 analysis that allows us to see which statements are significantly different from randomly chosen values.

Ranking of Statements

To compare students sorting for each statement, we began by creating a "score" for each statement by assigning a score of +2 for Group I, +1 for Group II, and so on until a score of -2 is assigned to Group V. The scores were then summed and divided by the number of students to get an average score for each statement. A higher positive score means students rank that statement as more descriptive more often; a more negative score means students rank the statement as less descriptive. Our results are presented in Table 1.

TABLE 1 Most and Least Characteristic Statements Ranked by Average Scores.

	107 students		**108 students**	
	Statement	**Avg.**	**Statement**	**Avg.**
Most	17	0.88	17	1.10
desc.	1	0.60	13	0.87
	5	0.56	23	0.77
	23	0.56	16	0.63
	22	0.54	22	0.53
to	…	…	…	…
	21	-0.49	21	-0.43
	24	-0.56	14	-0.47
	14	-0.60	24	-0.53
Least	7	-0.88	7	-0.60
desc.	6	-1.74	6	-1.70

Among the Physics 107 and 108 students, statement 17 ("students discuss their data and conclusions with each other") was chosen as most descriptive. Statement 6 ("the instructor lectures to the whole class") was chosen as the least descriptive statement. It is gratifying to see the students say that students work together within their groups and that the instructor does not lecture to the class at all in PbI.

The average scores analysis provides a quick overview of trends in perception among a group of students. There is, however, much more information that can be extracted from the data. For example, are the trends statistically significant, and are there interesting distributions or trends that our naïve ranking missed?

Matrix and χ^2 Analysis

In order to examine the details of the rankings more thoroughly, we used the five categories and twenty-five statements to construct a matrix of values (not shown due to lack of space). The frequency with which each statement appeared in each Group (I – V) make up the values in the matrix. We then normalized each matrix element by dividing by the number of students who completed the modified LPVI.

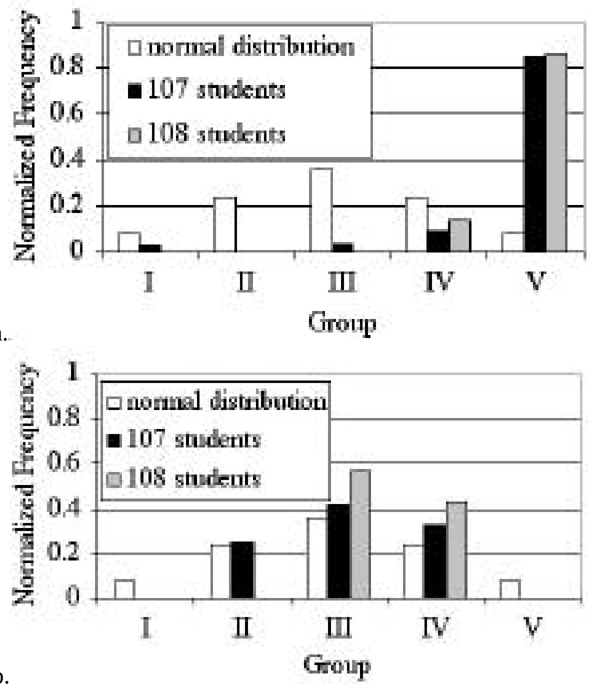
We examined the results on a statement by statement basis, looking for significant trends in student opinion. Our statistical null hypothesis is as follows: If a particular statement was sorted completely randomly by a group of n students, it would occur in Groups I and V 8% of the time, in Groups II and IV 24% of the time, and in Group III 36% of the time. In order to determine whether the matrix obtained from the sortings differed from these random values, we subtracted the random values from our original matrix. This gave us a new matrix M with matrix elements M_{ij}, where i refers to the rows of the matrix (modified LPVI statements) and j refers to the columns (Groups I – V). The sum of any row i of the matrix ($\Sigma^5_{,j=1} M_{ij}$) is equal to zero. A positive (negative) value for M_{ij} would mean that statement i is ranked into Group j more (less) often than one would expect from a random distribution. The matrix elements range from +0.92 to -0.36. These matrix elements are difference values, and we denote the range of these values difference ranges. Difference values for statement 2 from Physics 107 students are shown in Table 2.

TABLE 2 Difference Values for Statement 2.

I	**II**	**III**	**IV**	**V**
-0.04	-0.03	+0.18	-0.04	-0.07

We performed a χ^2 analysis on each question, for 107 students and 108 students, to find if any of the differences between the normal distribution (null hypothesis) and the students' data were statistically significant (had confidence level > 95%).

We noticed two types of pattern in student responses that produced significant χ^2 values. The first is when students agree with each other that a statement is descriptive or not descriptive of the course. A typical example for statement 6 ("the instructor lectures to the whole class") is shown in Fig 1a below.



a.

b.

FIGURE 1. Normalized frequency distribution for both Physics 107 and 108 students, compared against the null hypothesis for a. statement 6 b. statement 10.

Statements that students considered descriptive, have positive scores for matrix elements M_{i1} and M_{i2}, and negative scores for M_{i4} and M_{i5}. Conversely, statements that students considered non-descriptive, have negative scores for matrix elements M_{i1} and M_{i2}, and positive scores for M_{i4} and M_{i5}. Therefore we can construct an analog of the average score for statement i by adding matrix elements: $M_{i1} + M_{i2} - M_{i4} - M_{i5}$. This presents us with another way of finding the most and least descriptive statements. The results, along with the χ^2 analysis, are shown in Table 3 for Physics 107 and Physics 108 students. (Note that the 95% confidence level is at a χ^2 of approximately 9.4.)

A second pattern that we observed, which could not be identified by the average score or matrix ranking methods, is shown in Table 4 below for all four statements having such a pattern. This represents an agreement among students that a particular statement is neither descriptive nor non-descriptive of the class. Such a pattern in the matrix represents a sharp spike in the Group III category. This is shown in Fig. 1b.

We binned the difference values from the 107 students' matrix and the 108 students' matrix to graph the distribution of matrix elements for each group (Fig. 2). The maximum number of difference values occurs in the -0.1 to 0.0 bin (slightly below zero) because of the very few highly positive uncompensated matrix elements seen to the right in Fig. 2. This figure indicates that there is reasonable student agreement about which statements are not descriptive than there is about which statements are descriptive. It can also be seen that the distributions are quite similar to each other.

TABLE 3 Most and Least Characteristic Statements from χ^2 compared to Average Score.

	Physics 107 students			
	State-ment	**Av. Score**	χ^2	**p(χ^2)**
Most descriptive	17	0.61	83	<0.001
	1	0.36	38	<0.001
	22	0.43	29	<0.001
	23	0.38	24	<0.001
	5	0.41	24	<0.001
	16	0.28	15.3	0.004
	13	0.28	13.9	0.008
to . . .	. . .	. . .	. . .	. . .
	21	-0.38	21	<0.001
	14	-0.43	29	<0.001
	24	-0.49	38	<0.001
Least descriptive	7	-0.62	69	<0.001
	6	-0.92	670	<0.001
	Physics 108 students			
	State-ment	**Av. Score**	χ^2	**p(χ^2)**
Most descriptive	17	0.73	46	<0.001
	13	0.60	23	<0.001
	23	0.60	21	<0.001
	16	0.47	11.1	0.025
to . . .	. . .	. . .	. . .	. . .
	24	-0.47	13.4	0.009
	7	-0.40	14.4	0.006
Least descriptive	21	-0.43	18.0	0.001
	6	-0.90	212	<0.001

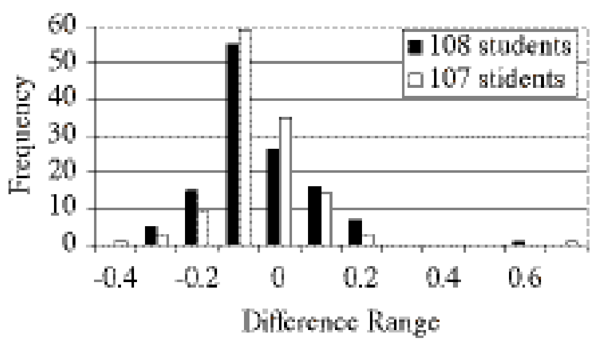


FIGURE 2. The number of matrix elements in each defined range of difference values (from -0.5 to -0.4 to 0.7 to 0.8).

TABLE 4. Neutral Statements with Significant χ^2.

Physics 107 students			Physics 108 students		
Statement	**χ^2**	**$p(\chi^2)$**	**Statement**	**χ^2**	**$p(\chi^2)$**
2	14.9	0.005	10	10.6	0.030
12	11.6	0.021	8	10.2	0.037

DISCUSSION AND CONCLUSION

Table 3 shows some similarities and differences between Physics 107 students and Physics 108 students. The rankings in Table 3 are ordered in terms of decreasing (or increasing) χ^2. We found several interesting features that χ^2reveals in addition to the information gleaned from the "average score" analysis. For Physics 107 students, we find additional significant statements 13 and 16 in the category of descriptive. Statements 13 ("the instructor or laboratory manual requires that students explain why certain things happen") and 16 ("questions in the laboratory manual require that students use evidence to back up their conclusions") refer to the evidence needed to explain observations. Table 3 shows that Physics 107 and 108 students agree that statements 13, 16, 17 ("students discuss their data and conclusions with each other"), and 23 ("in discussion with the instructor, assumptions are challenged and conclusions must be justified") are descriptive of their PbI experience. All of these statements reflect important goals of instruction in PbI.

While Physics 107 students chose statement 1 ("students follow the step-by-step instructions in the laboratory manual") as descriptive, with a $\chi^2 = 37.8$, Physics 108 students placed it tenth overall with a χ^2 of just 3.0. Apparently, Physics 108 students, who were allowed free rein to move light sources, masks, and screens felt less constrained by the book than Physics 107 students, struggling to understand how circuits work. For Physics 108 students, statement 5 ("laboratory activities are used to develop concepts"), while ranked in the top 5 in Table 1, is not statistically significant (χ^2 is just below the 95% confidence level). Physics 107 students, however, ranked statement 5 as descriptive ($\chi^2 = 23.9$).

Interesting are the statistical significance of statement 2 ("questions in the laboratory manual require the interpretation of data"), and 12 ("the laboratory manual requires that specific questions be answered") for Physics 107 students. Likewise statement 8 ("during laboratory students record information requested by the instructor or the laboratory manual"), and statement 10 ("the instructor or laboratory manual identifies the problem to be investigated") were given a statistically significant ranking by Physics 108 students. A look at the matrices for the two classes reveals that these statements are clustered in Group III. Such clustering is representative of a consensus among students that the statements in question are neither descriptive, nor non-descriptive of the course.

In the least descriptive category Physics 107 and 108 students mostly agree. The selection of statement 6 ("the instructor lectures to the whole class") and statement 24 ("students usually know the general outcome of an experiment before doing the experiment") should be considered to indicate success for our implementation of PbI because it shows consonance between instructor goals and student perception.

Other student choices reveal that we as instructors could do more to encourage independent thought and experimentation by students in the courses. The selection of statement 1 ("students follow the step-by-step instructions in the laboratory manual") among "most descriptive", and statements 7 ("students are asked to design their own experiments") and 21 ("students identify problems to be investigated") in the "least descriptive" category, provides formative assessment information for future changes in the course. We might "know" this anecdotally, but the Q-sort chi-square analysis gives us a clearer foundation for action.

The results and analysis presented in this paper show the value of this diagnostic Q-sort instrument for learning about student perception of laboratory courses. Further analysis of these data is under way. In particular, we are examining differences among class sections. Information from this analysis may help instructors learn how to teach better.

REFERENCES

1. McDermott, L. C., *Physics by Inquiry*, vol. 1, Wiley, New York, 1995.
2. McDermott, L. C., *Physics by Inquiry*, vol. 2, Wiley, New York, 1995.
3. Abraham, M. R., *J. Res. Sci. Teach.*, **19**, 155 – 165 (1982).
4. Lin, Y., Demaree, D., Zou, X., and Aubrecht, G., PERC Proceedings (2005).
5. Lin, Y., Demaree, D., Zou, X., and Aubrecht, G., Poster presented at PERC Conference (2004).
6. Zou, X., Lin, Y., Demaree, D., and Aubrecht, G., Poster presented at PERC Conference (2004).
7. Stephenson, W., *Nature*, **136**, 297– (1935).
8. Brown, S. R., *Qualitative Health Research*, **6**, 561–567 (1996).
9. Stephenson, W., *The study of behavior: Q-technique and its methodology*, University of Chicago Press, Chicago, 1953.