

An Inferential Analysis of the Effect of Activity-Based Instruction on the Persistent Misconceptions of Physics Students

Emily M. Reiser¹

Mark E. Markes²

ABSTRACT

A "persistent misconception" can be defined as a wrong-to-same-wrong response on pre and posttest conceptual evaluations in which the distracters are chosen from common physical misconceptions held by students. Although standard analysis techniques can reveal the presence of persistent misconceptions, they do not directly show how a particular student's response in one type of class might have been modified had he or she taken a different type of class. In this paper pre and posttest Force and Motion Conceptual Evaluation results are characterized in terms of wrong to same wrong, wrong to different wrong, wrong to right, etc., and a method (based on sampling theory) of inferring how activity-based instruction might have changed the response of traditional lecture students is presented. The results indicate that a wrong-to-same-wrong response by a lecture student would be more likely to be converted to a wrong-to-right response by activity-based instruction than would a wrong-to-different-wrong response.

¹ deceased

² corresponding author: markesme@unk.edu

I. INTRODUCTION

About twenty five years ago researchers began to develop conceptual test instruments to evaluate student understanding of force and motion (Beichner, 1994; Hestenes & Wells, 1992; Hestenes, Wells, & Swackhamer, 1992; Thornton & Sokoloff, 1998). As a result of this development several evaluation instruments have emerged. Two of the more commonly used instruments are the Force Concept Inventory (FCI) (Hestenes et al., 1992) and the Force and Motion Conceptual Evaluation (FMCE) (Thornton & Sokoloff, 1998). Both the FCI and FMCE underwent similar lengthy and rigorous developments. In the first stage of this development short answer responses to carefully-written conceptual questions and the results of student interviews were used to identify many of the student misconceptions associated with a particular question. In the second stage of development these misconceptions were used as the basis for writing “distracters” for multiple-choice questions. Thus the presence of an identical wrong answer by a student on a pre and posttest is an indication of a pre-existing and persistent misconception.

The physics education research community began to systematically study student use of misconception-based distracters about ten years ago. In 1995, Thornton (Thornton, 1995) introduced the term “Conceptual Dynamics” to refer to the phenomenological study of student concepts as a function of time. In Thornton’s method the student responses to multiple-choice FMCE questions are mapped into categories called “student views” that are associated with particular beliefs about the relationship between force and motion. Thornton monitored the change in the frequencies describing the use of these views over the course of instruction. More recently, Bao and Redish have introduced a method called “Model Analysis” (Bao, 1999; Bao & Redish, 2006) that is also based on a predetermined set of student views or “models”. As in Conceptual Dynamics the models imply certain beliefs about force and motion on the part of the students. Model Analysis has been applied to analyze pre and posttest responses obtained from both the FCI and FMCE (Bao, 1999; Bao, Hogg, & Zollman, 2002; Bao & Redish, 2006).

Conceptual Dynamics and Model Analysis both involve the application of a predetermined set of models. However, there is one significant aspect of a student’s incorrect posttest response that can be addressed without constructing a set of student

models: This aspect is, whether or not the student's incorrect posttest response is the same or different from the pretest response. In this paper a study is made of the correct/incorrect responses from pre and posttests using the FMCE. For the purpose of this study wrong-to-same-wrong responses will be referred to as "persistent misconceptions". In addition, it has been found possible using this type of characterization to make statistical inferences concerning how activity-based instruction might have modified a traditional lecture student's wrong-to-same-wrong responses. Both the FMCE test results and an analysis of the data will be presented in this paper.

The paper is arranged as follows: In Section II the instruction methods and FMCE test results are presented. In Section III an analysis methodology that infers the effect that activity-based instruction might have had on lecture students who show persistent misconceptions after lecture instruction is presented. In Section IV the results of applying the method to the FMCE test responses are given, and the paper concludes in Section V with a summary and conclusions.

II. INSTRUCTION METHODS AND FMCE TEST RESULTS

Instruction Methods

This study involves two instructors and six sections of algebra-level, introductory physics (one section of each class type each fall semester from 1995 to 1997) at a Midwestern university in the USA. Three of these sections employed a Lecture-Based (LB) format and were taught by the first instructor. The other three sections employed an Activity-Based (AB) format and were taught by the second instructor. Both instructors had over 20 years of teaching experience at the time the data were collected, and both are popular teachers who receive excellent student evaluations. The LB sections met three 50-minute periods plus one 75-minute period per week with one three-hour laboratory each week. The AB sections met three 110-minute periods plus one 75-minute period per week with an integrated laboratory. Both types of class had a total of 6.75 contact hours per week. The mean number of LB students per section was 29, and the mean number of AB students per section was 21.

The LB sections required a standard textbook and problems were assigned from this textbook (Cutnell & Johnson, 1995). The AB sections were based on about 15 to 20 minutes of lecture instruction per 110-minute session with the remainder of the time used by the students, working in groups, to complete sections of a workbook (Markes & Markes, 1995) or do experiments. For data collection the AB sections primarily utilized the Calculator-Based Laboratory (CBL) developed by Texas Instruments. The CBL technology was essentially a TI-85 calculator, various “probes”, and the CBL interface box that connected the calculator to the probes. The activity workbook was written by the instructor and was designed to integrate activity-based methods with the CBL system. It was influenced to some extent by early versions of *Workshop Physics* (Laws, 2004). In addition, the activity workbook suggested readings from the same textbook used in the LB sections. However, it was primarily a stand-alone, guided-inquiry workbook organized into activity units focused on particular topics.

FMCE Test Results

Pre and posttesting was done using an early version of the FMCE. The primary difference between the newer version and the version used in this study is that the newer version (Thornton & Sokoloff, 1998) has more distracters for some of the questions. A total of 87 LB students and 64 AB students were tested. The version of the FMCE used has 43 multiple-choice questions of which 42 were analyzed. (One question regarding the forces on a book at rest on a table was omitted due to a spreadsheet capacity limitation.) Pretesting was done in both classes sometime during the first two weeks of instruction, and posttesting was done near the end of the semester after all topics on the FMCE had been covered in both classes. About equal amounts of time were devoted to force and motion topics in both types of class.

The equivalence of the LB and AB students before instruction can be examined using the pretest responses. Figure 1 shows a histogram of the normalized pretest-wrong responses per student grouped into four bins: 0–24, 25–29, 30–34, and 35–39. The picket heights were calculated by first counting the total number of pretest wrongs for each student in the LB and AB class types. Each student was then placed into the appropriate

bin. The bin values were then normalized by dividing by the total number of students in the appropriate class type (LB or AB). From Figure 1 it can be seen that there is a fairly good match between the fractions of students in each bin. A chi-square test for independence yielded $p = 0.20$ which in most cases would not be considered large enough to invalidate the hypothesis of independence. In general, the overall pretest results are very consistent with the findings of other researchers who have found that an average student will miss about 75 percent of the questions on the pretest.

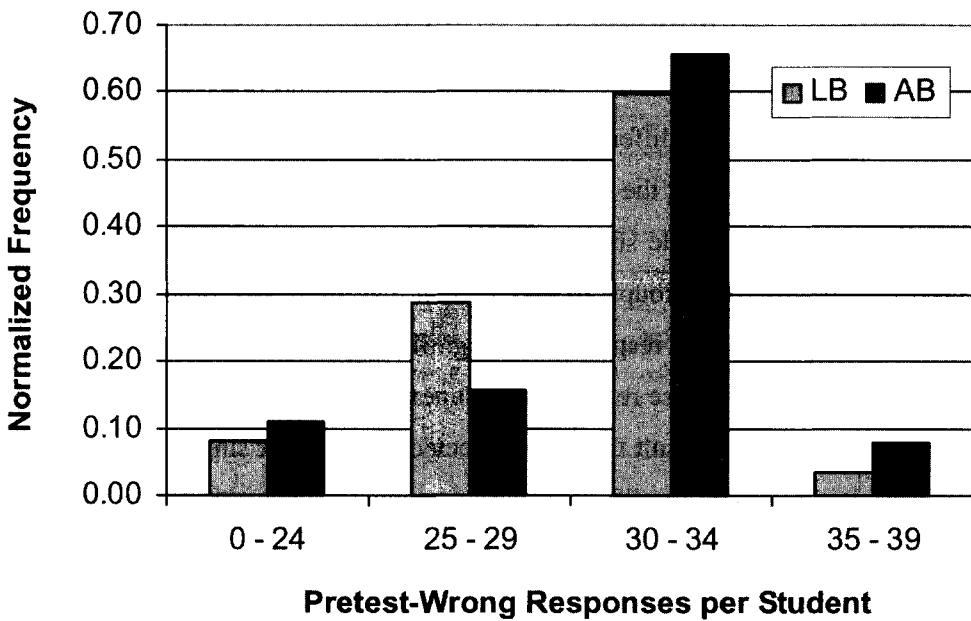


Figure 1

Normalized pretest-wrong responses per student.

Based on the pretest there are two response categories for a particular student and question combination: “pretest right” and “pretest wrong”. When both the pretest and the posttest are considered, each pretest category can be divided into response groups. There are two response groups for pretest-right questions: right to right (rr) and right to wrong (rw). For pretest-wrong questions there are three response groups: wrong to same wrong (ww), wrong to different wrong (ww'), and wrong to right (wr).

The questions on the FMCE can be divided into eight question groups: sled on ice (sled_F), toy car on ramp (car_F), coin tossed vertically into air (coin_F and coin_A), car on a horizontal surface (car_FG, car_VG, and car_AG), and Newton's Third Law (NTL). An "F", "V", or "A" denotes a question group associated with force, velocity, or acceleration respectively. In addition, a "G" denotes a question group in which some of the information is given in graphical form.

Figure 2 shows the total number of responses in each question group normalized with respect to the total number of pretest-wrong responses in that question group and class type (n_{ww} , $n_{w\acute{w}}$, n_{wr}). If the number of pretest-wrong responses is assumed to be the sample size, then (n_{ww} , $n_{w\acute{w}}$, and n_{wr} are the proportions associated with the responses. The theory of the difference between two proportions (Keller, 2001) was used to test the null hypothesis that the LB and AB populations were equivalent. In each case it was assumed that the sample size equaled the number of pretest-wrong responses for that class type and question group.³ It can be seen that the AB sections had significantly fewer wrong-to-same-wrong responses and significantly more wrong-to-right responses than did the LB sections. These results are not unexpected for AB classes when compared to LB classes. However, a result that is unexpected is the rather similar response for both types of class in the wrong-to-different-wrong category. One way to explain this result is to assume that AB instruction is mainly effective for questions for which students have misconceptions (revealed as n_{ww} in lecture). However, the possibility exists that ww students might be converted to $w\acute{w}$ and an approximately equal number of $w\acute{w}$ students converted to wr producing the same observed result. In the next section an inferential analysis method is proposed, based on right-to-wrong student responses, that suggests that students in the ww group are more likely to be converted to wr .

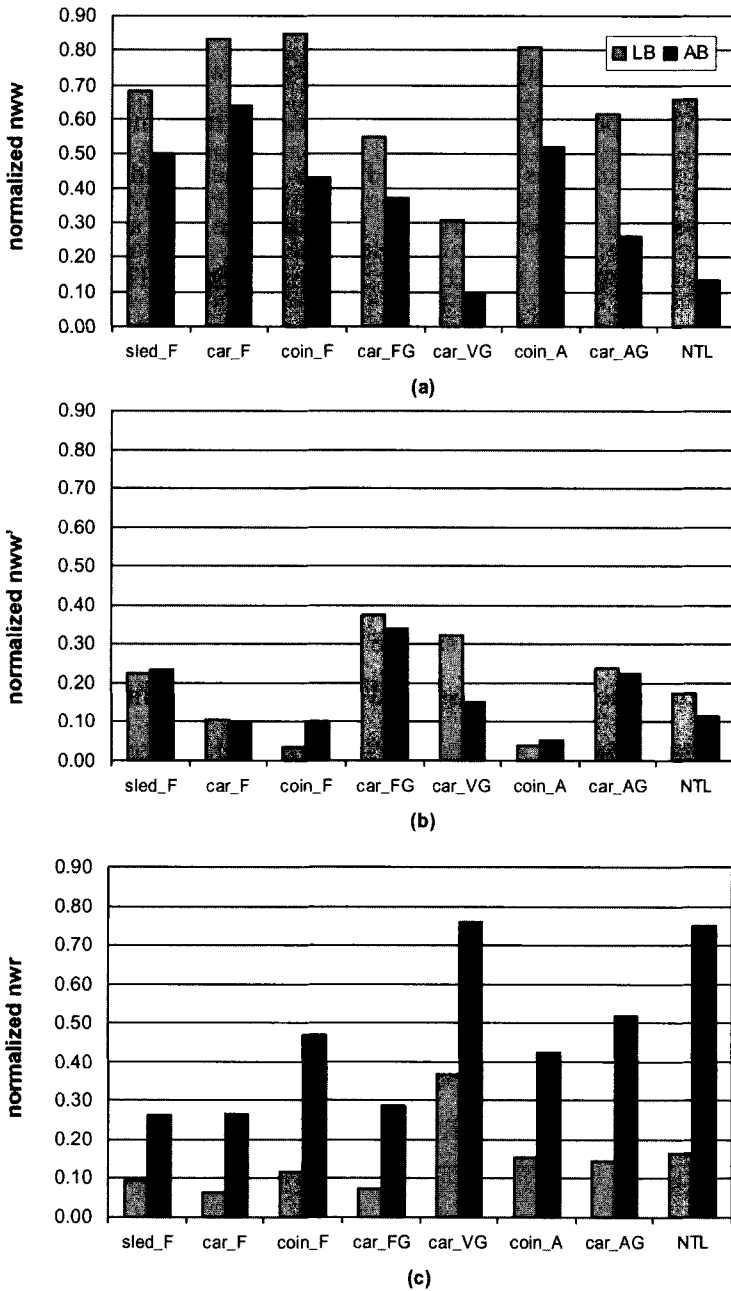


Figure 2 The LB and AB proportions for the ww , ww' and wr responses. A null hypothesis assuming equivalent LB and AB populations, using a pooled estimator, yields p -values less than 0.01 for all ww and wr responses. The same null hypothesis yields p -values greater than 0.02 for five of the eight ww' responses.

III. INFERENCE ANALYSIS

For the inferential analysis all question groups were combined, and each student was considered to be a primary sampling unit chosen at random from a large hypothetical population of N students. As shown in Figure 3 each student in the population is hypothetically assumed to be instructed by both the LB and AB methods with a re-initialization step between the LB and AB instruction. The process begins with a pretest of all students in the population followed by instruction using LB methods. After instruction the students are posttested and labeled population P_0 . The LB students taking part in this study are considered to be a random sample taken from the population P_0 . After posttesting the students are hypothetically returned to their pre-instruction state and re-instructed using AB methods and posttested a second time. Following AB instruction and posttesting the population is labeled P_1 , and the AB students participating in this study are considered to be a random sample taken from P_1 . Since the population is large, sampling with replacement can be assumed even though in practice AB students cannot be found in LB samples.

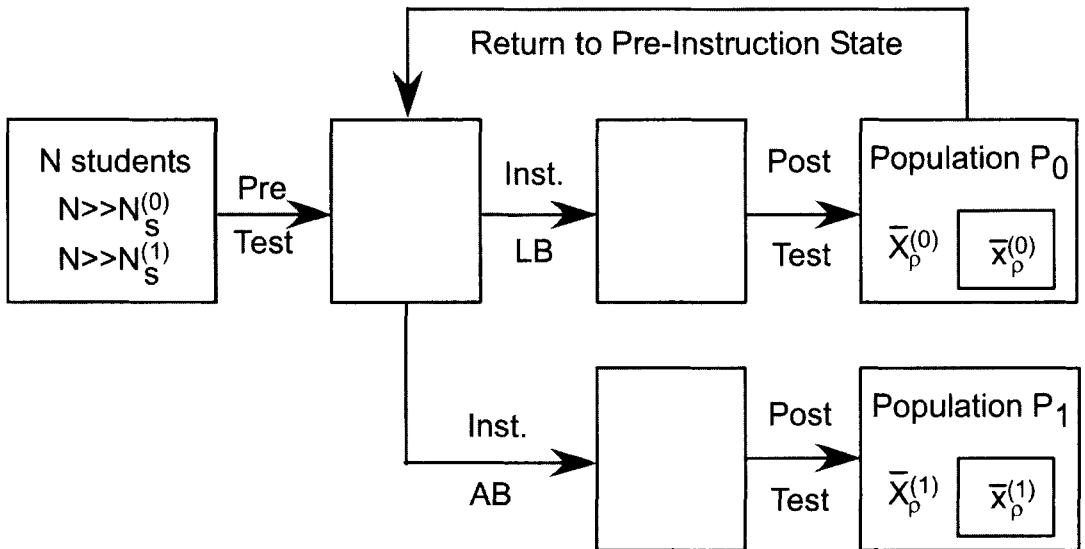


Figure 3 The hypothetical instruction procedure.

It might be claimed that this hypothetical process is unacceptable because it cannot be performed in practice. However, it has frequently been pointed out that students enter into instruction with previously-acquired conceptual elements and a sense of “mechanism” (Halloun & Hestenes, 1985; McDermott, 1984; Viennot, 1979) and (diSessa, 1993; Dykstra, Boyle, & Monarch, 1992; Hammer, 1996; Redish, 1994; Smith, diSessa, & Roschelle, 1993) Furthermore, since assimilation plays an essential role in the acquisition of knowledge, the student’s knowledge state prior to instruction and the method of instruction are likely to interact to produce the product of the instruction process. It therefore is reasonable to ask: “what would have been the result if a given student had been instructed via an alternative process?” If it is assumed the AB students participating in the study are taken from the same pre-instruction population as the LB students, then re-initialization of the hypothetical Population of LB students simply reproduces the population from which the AB students are chosen in practice. Here it is assumed that N is large enough that sampling with replacement can be assumed with a negligible probability that a former LB student would be chosen in the AB sample.

In the analysis each student in the Population N can be associated with the 42 questions of the FMCE pretest. The number of pretest-wrong responses $N_w(i)$ and pretest-right responses $N_r(i)$ for student i are related as

$$N_w(i) + N_r(i) = 42. \tag{1}$$

When pre and posttest results are considered, a k label is required to indicate the type of instruction, and there are two response groups for pretest-right questions (rr and rw) and three response groups for pretest-wrong questions (ww , $w\acute{w}$, and wr). Therefore, each student can be associated with five random variables $n_{rr}^{(k)}(i)$, $n_{rw}^{(k)}(i)$, $n_{ww}^{(k)}(i)$, $n_{w\acute{w}}^{(k)}(i)$, and $n_{wr}^{(k)}(i)$. These are the total numbers of outcomes (frequencies) in the five response categories for student i in class type k .

The frequencies $n_{rr}^{(k)}(i)$, $n_{rw}^{(k)}(i)$, $n_{ww}^{(k)}(i)$, $n_{w\acute{w}}^{(k)}(i)$, and $n_{wr}^{(k)}(i)$, are normalized with respect to either the total number of pretest-right responses or the total number of pretest-wrong responses as:

$$X_{\rho}^{(k)}(i) = n_{\rho}^{(k)}(i)/N_{\varepsilon(\rho)}(i). \tag{2}$$

Here $\rho = rr, rw, ww, ww'$ or wr , $\in(\rho) = r$ when $\rho = rr$ or rw , and $\in(\rho) = w$ when $\rho = ww, ww'$, or wr . The quantities $X_{rr}^{(k)}(i)$ and $X_{rw}^{(k)}(i)$ are the conditional probabilities that, if student answers a question correctly on the pretest, he or she will answer it either correctly (rr) or incorrectly (rw) on the posttest. Analogously, if student answers a pretest-question incorrectly, the conditional probabilities for a ww, ww' or wr or response are $X_{ww}^{(k)}(i), X_{ww'}^{(k)}(i)$ and $X_{wr}^{(k)}(i)$ respectively, and average conditional probabilities can be defined over the population N as

$$\bar{X}_{\rho}^{(k)} = \frac{1}{N} \sum_{i=1}^N X_{\rho}^{(k)}(i). \quad (3)$$

Sample-based estimates of conditional probabilities will be denoted by a lower-case “x”. However, the definitions are very similar except the number of pretest-right, and pretest-wrong responses are functions of the class type k (i.e. the particular sample LB or AB):

$$x_{\rho}^{(k)}(i) = n_{\rho}^{(k)}(i)/N_{\in(\rho)}^{(k)}(i), \quad (4)$$

where ρ and $\in(\rho)$ are defined as for the Population N above. The averages are defined over the LB and AB samples as

$$\bar{x}_{\rho}^{(k)} = \frac{1}{N_s^{(k)}} \sum_{i=1}^{N_s^{(k)}} x_{\rho}^{(k)}(i), \quad (5)$$

where $N_s^{(k)}$ is the number of students in the LB ($k = 0$) or AB ($k = 1$) samples.

The changes in the population due to re-initialization and re-instruction using activity-based methods can be written as

$$\Delta \bar{X}_{\rho} = \bar{X}_{\rho}^{(1)} - \bar{X}_{\rho}^{(0)}, \quad (6)$$

where the $\bar{X}_{\rho}^{(k)}$ are the mean probabilities per student in population N after instruction of type k . Because $\bar{X}_{rr}^{(k)} + \bar{X}_{rw}^{(k)} = 1$ and $\bar{X}_{ww}^{(k)} + \bar{X}_{ww'}^{(k)} + \bar{X}_{wr}^{(k)} = 1$ (for both $k = 0$ and $k = 1$), the changes $\Delta \bar{X}_{\rho}$ must satisfy

$$\Delta\bar{X}_{rr} + \Delta\bar{X}_{rw} = 0 \tag{7}$$

and

$$\Delta\bar{X}_{ww} + \Delta\bar{X}_{wr} + \Delta\bar{X}_{wr} = 0. \tag{8}$$

Equations (7) and (8) imply the effect of re-initialization and re-instruction of the lecture students by activity-based methods can be represented by transfers of probability *A*, *B*, *C*, and *D* as shown in Figure 4. The transfers resulting from a hypothetical re-instruction using activity-based methods can be expressed as

$$D = -\Delta\bar{X}_{rw}, \tag{9}$$

and

$$A + C = \Delta\bar{X}_{wr}, \tag{10}$$

$$A + B = -\Delta\bar{X}_{ww}. \tag{11}$$

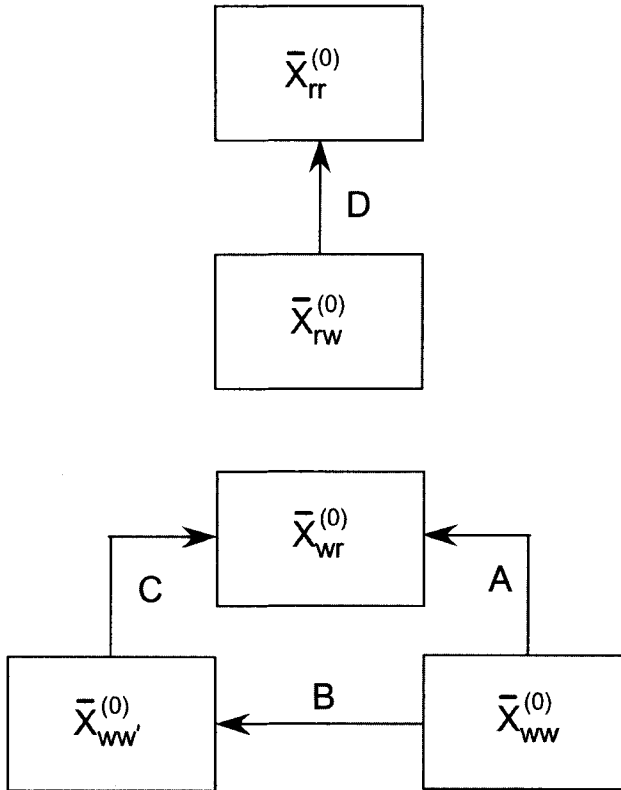


Figure 4 The LB probabilities and the transfers of probability that transform the LB probabilities into the AB probabilities.

Given A , B , C , and D , the effect of a hypothetical re-instruction of the population of N lecture-based students is determined. However, without additional information, Equations (10) and (11) cannot be solved for A , B , and C because there are more unknowns than independent equations.

A solution is made possible by an additional assumption that relates the values C and D . Recent interviews of students who answered right to wrong on certain FMCE questions have indicated that they generally could not state specifically why they were able to answer a question correctly on the pretest and yet failed to answer it correctly on the posttest. A typical response was that the pretest answer must have been a “lucky guess”. Another possibility, although none of the interviewed students mentioned it, is that the pretest-right response could have been obtained in spite of faulty physical understanding. An examination of the FMCE reveals this is possible for at least some of the questions (McDermott & Redish, 1999). Given the pretest-right response was fortuitous, or made in spite of a faulty physical model, right-to-wrong responses are essentially equivalent to wrong-to-different-wrong responses, and there must be some relationship between C and D .

This relationship can be obtained as follows. First express C and D as

$$C = \frac{1}{N} \sum_{i=1}^N \frac{n_{ww}^{(0)}(i)}{N_w(i)} \frac{n_{ww,wr}(i)}{n_{ww}^{(0)}(i)} = \frac{1}{N} \sum_{i=1}^N X_{ww}^{(0)}(i) f_{ww,wr}(i), \quad (12)$$

and

$$D = \frac{1}{N} \sum_{i=1}^N \frac{n_{rw}^{(0)}(i)}{N_r(i)} \frac{n_{rw,rr}(i)}{n_{rw}^{(0)}(i)} = \frac{1}{N} \sum_{i=1}^N X_{rw}^{(0)}(i) f_{rw,rr}(i), \quad (13)$$

where $n_{ww,wr}(i)$ and $n_{rw,rr}(i)$ are the numbers of responses transferred from ww to wr and from rw to rr by hypothetical re-instruction of lecture student i . The factors $f_{ww,wr}(i) = n_{ww,wr}(i)/n_{ww}^{(0)}(i)$ and $f_{rw,rr}(i) = n_{rw,rr}(i)/n_{rw}^{(0)}(i)$ are the “transition probabilities” for a response to be transferred from ww to wr and from rw to rr respectively for student i . Analogous expressions can be written for A and B

$$A = \frac{1}{N} \sum_{i=1}^N X_{ww}^{(0)}(i) f_{ww,wr}(i) \quad (14)$$

and

$$B = \frac{1}{N} \sum_{i=1}^N X_{rw}^{(0)}(i) f_{ww,ww}(i). \quad (15)$$

Assuming the transition probabilities f and the probabilities X are independent, Equations (12) through (15) can be written as:

$$C = \bar{X}_{ww}^{(0)} \bar{f}_{ww,wr}, \quad (16)$$

$$D = \bar{X}_{rw}^{(0)} \bar{f}_{rw,rr}, \quad (17)$$

and

$$A = \bar{X}_{ww}^{(0)} \bar{f}_{ww,wr}, \quad (18)$$

$$B = \bar{X}_{ww}^{(0)} \bar{f}_{ww,ww}, \quad (19)$$

where $\bar{X}_{\rho}^{(0)}$ and $\bar{f}_{\rho,\rho}$ denote mean values obtained over the population.

Based on the assumed equivalence of the ww and rw responses as discussed above, the mean transition probabilities $\bar{f}_{ww,wr}$ and $\bar{f}_{rw,rr}$ can be set equal, and Equations (9), (16), and (17) yield:

$$C = -\frac{\bar{X}_{ww}^{(0)}}{\bar{X}_{rw}^{(0)}} \Delta \bar{X}_{rw}. \quad (20)$$

In practice only sample-based quantities are available. However, sample-based statistics can be used to estimate population-based statistics to within a certain estimated confidence interval determined from the standard deviation of the sample. The probability transfer C can be estimated as

$$C \approx -\frac{\bar{x}_{ww}^{(0)}}{\bar{x}_{rw}^{(0)}} \Delta \bar{x}_{rw}. \quad (21)$$

With C given, Equation (10) can be used to obtain A using $\Delta\bar{x}_{wr}$ to estimate $\Delta\bar{X}_{wr}$

$$A \approx \Delta\bar{x}_{wr} - C, \tag{22}$$

and similarly with A known, Equation (11) yields:

$$B \approx -\Delta\bar{x}_{ww} - A. \tag{23}$$

Given A , B , and C the average transition probabilities can be estimated from Equations (16), (18), and (19)

$$\bar{f}_{ww,wr} \approx C/\bar{x}_{ww}^{(0)}, \tag{24}$$

$$\bar{f}_{ww,wr} \approx A/\bar{x}_{ww}^{(0)}, \tag{25}$$

$$\bar{f}_{ww,ww} \approx B/\bar{x}_{ww}^{(0)}. \tag{26}$$

IV. APPLICATIONS TO FMCE RESULTS

Tables 1 and 2 show the sample-based, pretest-right and pretest-wrong, mean conditional probabilities obtained from the FMCE test results and Equations (4) and (5). The sample size is denoted by N_s , and the sampling standard deviations of the means (which are shown in parenthesis) have been estimated using the approximation displayed in Equations (A3) and (A4) of the Appendix. The sampling standard deviations for the differences between AB and LB instruction can be calculated from the standard deviations given in Table 2 and Equation (A10) from the Appendix. For $\Delta\bar{x}_{ww}$ and $\Delta\bar{x}_{wr}$ the measured difference is about ten times the standard deviation, and for $\Delta\bar{x}_{ww}$ the measured difference is about one standard deviation. A null hypothesis that LB and AB instruction are equivalent yields p-values less than 0.01 for the ww and wr responses. Table 2 shows the same general trend apparent in Figure 2. Overall, AB instruction significantly increases the probability of a wrong-to-right response, significantly decreases the probability of a wrong-to-same-wrong response, and changes the probability of a wrong-to-different-wrong response by a much smaller amount. It can be calculated from the values given in

Table 1 that AB instruction decreases the mean conditional probability \bar{x}_{rw} for a right-to-wrong response by about $\Delta\bar{x}_{rw} = -0.048$.

An estimate of the probability transfer C can be obtained using the value of $\Delta\bar{x}_{wr}$ calculated above, other conditional probabilities found in Tables 1 and 2, and Equation (21). The result is shown in Table 3. Table 3 also shows the mean transition probability $\bar{f}_{ww,wr}$ estimated from C using Equation (24) and $\bar{x}_{ww}^{(0)}$ from Table 2. Table 3 also shows the transfers A and B , calculated from Equations (22) and (23), as well as the associated mean transition probabilities calculated from Equations (25) and (26). Included in Table 3 are standard deviations obtained using results found in the Appendix. The standard deviations are obtained by systematically applying Equations (A10), (A15), and (A20) with sample means used to estimate population means and standard deviations of the means as given in Tables 1 and 2.

It is interesting to note that the transition probability $\bar{f}_{ww,wr}$ is much smaller than the transition probability $\bar{f}_{ww,wr}$. This is consistent with the proposal, stated in Section II, that AB instruction is more effective in converting ww to wr responses than in converting $w\bar{w}$ to wr responses. It was proposed in Section II that this would explain why the normalized wrong-to-different-wrong responses are about the same for the LB and AB sections (Figure 2b). The inferential analysis presented in this section supports this assumption.

Table 1

Mean conditional probabilities for the pretest-right responses obtained from the student samples. Sample sizes are denoted by N_s and the estimated standard deviations of the sampling means were obtained using the approximations displayed in Equations (A3) and (A4) of the Appendix.

Method	N_s	$\bar{x}_{rr}^{(k)}$	$(\sigma_{\bar{x}_{rr}^{(k)}})$	$\bar{x}_{rw}^{(k)}$	$(\sigma_{\bar{x}_{rw}^{(k)}})$
LB ($k = 0$)	87	0.779	(0.019)	0.221	(0.019)
AB ($k = 1$)	64	0.827	(0.025)	0.173	(0.025)

Table 2

Mean conditional probabilities for the pretest-wrong responses obtained from the student samples. Sample sizes are denoted by N_s , and the estimated standard deviations of the sampling means were obtained using the approximations displayed in Equations (A3) and (A4) of the Appendix.

Method	N_s	$\bar{X}_{ww}^{(k)}$	$(\sigma_{\bar{X}_{ww}}^{(k)})$	$\bar{X}_{ww'}^{(k)}$	$(\sigma_{\bar{X}_{ww'}}^{(k)})$	$\bar{X}_{wr}^{(k)}$	$(\sigma_{\bar{X}_{wr}}^{(k)})$
LB ($k = 0$)	87	0.658	(0.015)	0.215	(0.012)	0.127	(0.011)
AB ($k = 1$)	64	0.345	(0.023)	0.191	(0.018)	0.464	(0.031)

Table 3

Probability transfers and transition probabilities predicted using the inferential analysis of Section III. Estimated standard deviations were obtained by applying Equations (A10), (A15), and (A20) from the Appendix with the standard deviations of the means given in Tables 1 and 2.

	A (s.d.) $\bar{f}_{ww,wr}$ (s.d.)	B (s.d.) $\bar{f}_{ww,ww'}$ (s.d.)	C (s.d.) $\bar{f}_{ww,wr}$ (s.d.)
Probability Transfers	0.290(0.045)	0.023(0.053)	0.047(0.031)
Transition Probabilities	0.441(0.069)	0.035(0.080)	0.219(0.145)

V. SUMMARY AND CONCLUSIONS

This paper has presented a method, based on sampling theory, that infers the effect that activity-based instruction would have on traditional lecture students were it possible to return them to their initial knowledge state and re-instruct them using activity-based methods. This is done by categorizing responses on the FMCE pre and posttests as wrong to same wrong (ww) wrong to different wrong (ww'), wrong to right (wr), right to

wrong (wr), or right to right (rr). In addition, it is assumed (based on student interviews) that a $w\acute{w}$ response is equivalent to an rw response. The principal theoretical object of the analysis is a single large hypothetical population of students that is initially instructed using traditional lecture methods then re-initialized and re-instructed using AB instruction. The lecture students participating in this study are assumed to be a random sample taken from this population after LB instruction. The activity-based students participating in this study are assumed to be a random sample taken after re-initialization and re-instruction using AB instruction. The population is assumed to be very large so that sampling with replacement can be assumed.

The primary quantities of interest are the mean transition probabilities $\bar{f}_{\rho,\beta}$ with ρ and $\beta = ww, w\acute{w},$ or wr . The equivalence of the $w\acute{w}$ and the rw responses is used to justify the assumption $\bar{f}_{w\acute{w},wr} = \bar{f}_{rw,rr}$. This assumption allows all probability currents (the flow of conditional probability from one type of response to another) and mean transition probabilities (normalized number of responses transferred between response groups) to be calculated. It is found that $\bar{f}_{ww,wr}$ and $\bar{f}_{w\acute{w},wr}$ are resolved at one standard deviation with the estimated value of $\bar{f}_{w\acute{w},wr}$ about twice the estimated value of $\bar{f}_{ww,wr}$. The transition probability $\bar{f}_{ww,w\acute{w}}$ is not completely resolved from $\bar{f}_{w\acute{w},wr}$. However, $\bar{f}_{ww,w\acute{w}}$ is small. The estimated value of $\bar{f}_{ww,wr}$ is about twelve times as large as $\bar{f}_{ww,w\acute{w}}$. Generally, the results suggest $\bar{f}_{ww,wr} > \bar{f}_{w\acute{w},wr} > \bar{f}_{ww,w\acute{w}}$ with $\bar{f}_{ww,w\acute{w}}$ substantially smaller than either $\bar{f}_{ww,wr}$ or $\bar{f}_{w\acute{w},wr}$.

The practical impossibility of returning lecture students to their pre-instruction state brings into question the hypothetical model in which the N students are returned to their pre-instruction state. The justification lies in the generally accepted assertion that learning is at least a partly constructive process that makes use of a student's pre-existing knowledge or conceptual elements to construct new knowledge. It is certainly possible that different instruction methods could interact with these initial elements differently. It therefore is reasonable to ask what would have been the effect of AB instruction on an LB student with a particular set of conceptual elements. Furthermore, it makes sense to ask how his or her responses would have been modified. It is this later question that justifies the model.

From the point of view that considers a student's initial knowledge state, a lecture student who answers wrong to same wrong on a particular question (persistent misconception) has at least a working set of conceptual elements and a sense of mechanism that can be applied to that problem. The results presented in this paper go on to suggest that activity-based instruction can work effectively with these students, taking advantage of these persistent misconceptions to make the construction of "expert knowledge" easier to accomplish. Thus it appears that persistent misconceptions can in fact be used to advantage in activity-based classes.

Appendix

Let there be M_1 ways of taking random samples of size n_1 from Population P_1 and M_2 ways of taking random samples of size n_2 from Population P_2 . For each element i of sample α from P_1 a measurement $x_{ww,wr}(i)$ can be made, and the mean value \bar{x}_α is

$$\bar{x}_\alpha = \frac{1}{n_1} \sum_{i=1}^{n_1} x_\alpha(i). \quad (A1)$$

Similarly, for each element j of sample β from Population P_2 a measurement $y_\beta(j)$ can be made, and the mean value \bar{y}_β is

$$\bar{y}_\beta = \frac{1}{n_2} \sum_{j=1}^{n_2} y_\beta(j). \quad (A2)$$

A theorem from statistics can be used to relate the variance of the sampling distribution of the mean to the variance of the sampled quantity over the population. In addition, the variance of a sampled quantity over a population can be estimated by the variance of that quantity over the sample. For Populations P_1 and P_2 :

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n_1} \approx \frac{s_x^2}{n_1} \quad (A3)$$

and

$$\sigma_{\bar{y}}^2 = \frac{\sigma_y^2}{n_2} \approx \frac{s_y^2}{n_2}, \quad (A4)$$

where s_x^2 is the variance of x over a sample from P_1 and s_y^2 is the variance of y over a sample from P_2 .

If the samples α and β are taken in pairs, one from P_1 and one from P_2 a distribution will exist for the difference between the means \bar{x}_α and \bar{y}_β . The variance of this difference is given by

$$\sigma_{\bar{x}-\bar{y}}^2 = \frac{1}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} [(\bar{x}_\alpha - \bar{y}_\beta) - (\bar{X} - \bar{Y})]^2, \tag{A5}$$

where \bar{X} and \bar{Y} are the means of the measurements x and y over the Populations P_1 and P_2 . Now define $\Delta\bar{x}_\alpha$ and $\Delta\bar{y}_\beta$ as :

$$\bar{x}_\alpha = \bar{X} + \Delta\bar{x}_\alpha, \tag{A6}$$

and

$$\bar{y}_\beta = \bar{Y} + \Delta\bar{y}_\beta. \tag{A7}$$

Substitution of Equations (A6) and (A7) into (A5) yields:

$$\sigma_{\bar{x}-\bar{y}}^2 = \frac{1}{M_1} \sum_{\alpha=1}^{M_1} \Delta\bar{x}_\alpha^2 + \frac{1}{M_2} \sum_{\beta=1}^{M_2} \Delta\bar{y}_\beta^2 - \frac{2}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} \Delta\bar{x}_\alpha \Delta\bar{y}_\beta. \tag{A8}$$

If the samples α and β are taken randomly from Populations P_1 and P_2 the variations $\Delta\bar{x}_\alpha$ and $\Delta\bar{y}_\beta$ will be independent and the last term will vanish resulting in

$$\sigma_{\bar{x}-\bar{y}}^2 = \frac{1}{M_1} \sum_{\alpha=1}^{M_1} \Delta\bar{x}_\alpha^2 + \frac{1}{M_2} \sum_{\beta=1}^{M_2} \Delta\bar{y}_\beta^2, \tag{A9}$$

which can be written

$$\sigma_{\bar{x}-\bar{y}}^2 = \sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2. \tag{A10}$$

The variance of the product $\bar{x}_\alpha \bar{y}_\beta$ is

$$\sigma_{\bar{x}\bar{y}}^2 = \frac{1}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} [\bar{x}_\alpha \bar{y}_\beta - \bar{X}\bar{Y}]^2. \quad (\text{A11})$$

To first order in $\Delta\bar{x}_\alpha/\bar{X}$ and $\Delta\bar{y}_\beta/\bar{Y}$:

$$\bar{x}_\alpha \bar{y}_\beta - \bar{X}\bar{Y} = \bar{Y}\Delta x_\alpha + \bar{X}\Delta y_\beta, \quad (\text{A12})$$

and

$$\sigma_{\bar{x}\bar{y}}^2 = \frac{\bar{Y}^2}{M_1} \sum_{\alpha=1}^{M_1} \Delta\bar{x}_\alpha^2 + \frac{\bar{X}^2}{M_2} \sum_{\beta=1}^{M_2} \Delta\bar{y}_\beta^2 + \frac{2\bar{X}\bar{Y}}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} \Delta\bar{x}_\alpha \Delta\bar{y}_\beta. \quad (\text{A13})$$

Again, because the samples are taken randomly from two different populations, they are independent and the last term vanishes. The variance is

$$\sigma_{\bar{x}\bar{y}}^2 = \frac{\bar{Y}^2}{M_1} \sum_{\alpha=1}^{M_1} \Delta\bar{x}_\alpha^2 + \frac{\bar{X}^2}{M_2} \sum_{\beta=1}^{M_2} \Delta\bar{y}_\beta^2, \quad (\text{A14})$$

which can be written

$$\sigma_{\bar{x}\bar{y}}^2 = \bar{Y}^2 \sigma_{\bar{x}}^2 + \bar{X}^2 \sigma_{\bar{y}}^2. \quad (\text{A15})$$

For \bar{x}/\bar{y} , the variance is

$$\sigma_{\bar{x}/\bar{y}}^2 = \frac{1}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} \left[\frac{\bar{x}_\alpha}{\bar{y}_\beta} - \frac{\bar{X}}{\bar{Y}} \right]^2. \quad (\text{A16})$$

Substitution from Equations (A6) and (A7) yields:

$$\frac{\bar{x}_\alpha}{\bar{y}_\beta} = \frac{\bar{X}}{\bar{Y}} \left[1 + \frac{\Delta\bar{x}_\alpha}{\bar{X}} \right] \left[1 - \frac{\Delta\bar{y}_\beta}{\bar{Y}} + \left(\frac{\Delta\bar{y}_\beta}{\bar{Y}} \right)^2 - \dots \right]. \quad (\text{A17})$$

To first order in $\Delta\bar{x}_\alpha/\bar{X}$ and $\Delta\bar{y}_\beta/\bar{Y}$

$$\frac{\bar{x}_\alpha}{\bar{y}_\beta} - \frac{\bar{X}}{\bar{Y}} = \frac{\bar{X}}{\bar{Y}} \left[\frac{\Delta\bar{x}_\alpha}{\bar{X}} - \frac{\Delta\bar{y}_\beta}{\bar{Y}} \right] \quad (\text{A18})$$

and

$$\begin{aligned} \sigma_{\bar{x}\bar{y}}^2 &= \frac{1}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} \left[\frac{\bar{x}_\alpha}{\bar{y}_\beta} - \frac{\bar{X}}{\bar{Y}} \right]^2 = \left(\frac{\bar{X}}{\bar{Y}} \right)^2 \\ &\times \left[\frac{1}{M_1} \sum_{\alpha=1}^{M_1} \left(\frac{\Delta\bar{x}_\alpha}{\bar{X}} \right)^2 + \frac{1}{M_2} \sum_{\beta=1}^{M_2} \left(\frac{\Delta\bar{y}_\beta}{\bar{Y}} \right)^2 - \frac{2}{M_1 M_2} \sum_{\alpha=1}^{M_1} \sum_{\beta=1}^{M_2} \frac{\Delta\bar{x}_\alpha}{\bar{X}} \frac{\Delta\bar{y}_\beta}{\bar{Y}} \right]. \end{aligned} \quad (\text{A19})$$

Again $\Delta\bar{x}_\alpha$ and $\Delta\bar{y}_\beta$ are independent, and the last term in the square brackets vanishes. The variance is

$$\sigma_{\bar{x}\bar{y}}^2 = \left(\frac{\bar{X}}{\bar{Y}} \right)^2 \left[\frac{\sigma_{\bar{x}}^2}{\bar{X}^2} + \frac{\sigma_{\bar{y}}^2}{\bar{Y}^2} \right]. \quad (\text{A20})$$

References

- Bao, L. (1999). *Dynamics of Student Modeling: A Theory, Algorithms, and Application to Quantum Mechanics*. doctoral dissertation, University of Maryland, College Park, Maryland.
- Bao, L., Hogg, K., & Zollman, D. (2002). Model analysis of fine structures of student models: An example with Newton's third law. *American Journal of Physics*, 70, 766–778.
- Bao, L., & Redish E. F. (2006). Model analysis: Representing and assessing the dynamics of student learning. [Online Journal]. *Physical Review Special Topics-Physics Education Research*, (online journal), 2, 010103-1-010103-16.
- Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, 62, 750–672.

- Cutnell, J. D., & Johnson, K. W. (1995). *Physics*. (3rd ed.) New York: John Wiley and Sons.
- diSessa, A. (1993). Toward an Epistemology of Physics. *Cognition and Instruction*, 10, 105-225.
- Dykstra, D. I., Boyle, C. F., & Monarch, I. A. (1992). Studying Conceptual Change in Learning Physics. *Science Education*, 76, 615-652.
- Halloun, I. A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics*, 53, 1056-1065.
- Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning and an appropriate role for education research. *American Journal of Physics*, 64, 1316-1325.
- Hestenes, D., & Wells, M. (1992). A Mechanics Baseline Test. *The Physics Teacher*, 30, 159-166.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158.
- Keller, G. (2007). *Applied Statistics with Microsoft Excel*. Pacific Grove, California: Thomas Learning Inc.
- Laws, P. W. (2004). *Workshop Physics Activity Guide*. New York: John Wiley and Sons.
- Markes, C. T., & Markes, M. E. (1995). *Activities for PHYS 205, General Physics*. (unpublished).
- McDermott, L. C. (1984). Research on conceptual understanding in mechanics. *Physics Today*, 37, 24-32.
- McDermott, L. C., & Redish, E. F. (1999). Resource Letter: PER-1 Physics Education Research. *American Journal of Physics*, 67, 755-767.
- Redish, E. F. (1994). Implications of cognitive studies for teaching physics. *American Journal of Physics*, 62, 796-803.
- Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *The Journal of the Learning Sciences*, 3, 115-163.

- Thornton, R. K. (1995). Conceptual Dynamics: Changing Student Views of Force and Motion. In C. Bernardini, C. Tarsitani, & M. Vicentini (Eds.), *Thinking Physics for Teaching* (pp. 157–183). New York: Plenum Press.
- 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, 338–352.
- Viennot, L. (1979). Spontaneous reasoning in elementary dynamics. *European Journal of Science Education*, 1, 205–210.

Author Note

The authors wish to acknowledge many helpful discussions with Dr. Teara Archwamety, Dept. of Counseling and School Psychology, University of Nebraska-Kearney; Dr. C. Trecia Markes, Dept. of Physics and Physical Science, University of Nebraska-Kearney; Dr. Dewey Dykstra, Dept. of Physics, Boise State University; Dr. Ronald Thornton, Center for Science and Mathematics Teaching, Tufts University; Dr. Robert J. Whitaker, Dept. of Physics, Astronomy and Material Science, Missouri State University. They would also like to thank Ms. Shea Holman for spreadsheet development and data entry and Mrs. Karen Malmkar for conducting the student interviews. They also express their gratitude to the University of Nebraska-Kearney for the University Research and Creative Activity grant which funded this research.