Comparing Two Results: Hake Gain and Dellow Gain, to Analyze FCI Data in Active Learning Process[*]

Azita Seyed Fadaei

South Seattle College, Washington, USA

The purpose of this study is comparing Hake gain and Dellow gain to analyze the effects of active learning method (Interactive Conceptual Instruction [ICI]) on the level of students' understanding in Newton's laws. We analyzed the main study by Hake Method in 2014. In this paper, we are going to analyze the results by Dellow's Method. Testing the effectiveness of research's idea by Force Concept Inventory (FCI) and presenting the data, on which it was constructed as Hgain(g), Dgain(G), Dloss, and Retention is the goal of this paper. Results show differences in understanding of students in ICI.

Keywords: Hake gain, Dellow gain, FCI, active learning

Introduction

In this study, we focus on comparing two methods (traditional and active learning, which we refer to Interactive Conceptual Instruction [ICI]) (Savinainen & Scott, 2002) of teaching to achieve to a better understanding of the environment through Newton's laws. Forming a meaningful understanding of Newton's laws of motion is formidably difficult for beginners. ICI teaching approaches have been developed to try to assist students to develop a meaningful understanding.

The course had consisted of two 70-minutes lecture sections per week for formal teaching and applying ICI in the experimental group, 20-minutes per week was added as an extra class. Regular teacher for both groups was the same. The extra class for the experimental group had been starting with some demonstrations or videotaped experiments, and students were asked to identify objects interacting with the object of interest, and then to make front view free-body diagrams. We were guiding them through the process of carrying out, making sense of, and modeling their experiments in multiple representation activities. The learning processes were collaborative and cooperative. Collaborative learning, in which students work in groups to solve problems, perform exercises, has positive effects on achievement, and cooperative learning relies upon social interactions among students (Gabel, 2003).

Teaching Strategy

The ICI approach would offer the potential to promote enhanced learning gains in conceptual understanding of mechanics. In this study, the extra class for the ICI group, the students in groups

(peer-instruction) are asked to answer the demonstration questions. The students' answers for multiple representation activities (MRA) exercises were collected before the teacher gave the correct answers. Especially students will have certain difficulties, it is important that the tendency to make a particular error be deliberately exposed and then explicitly addressed. Once an error is elicited through an appropriate task, the student can be helped to recognize and confront the difficulty. At that point, it is crucial that the instructor insists that the difficulty is resolved. If this is not done, the difficulty is likely to remain latent and arise later in a different context (McDermott, 1998).

Research Instrument and Force Concept Inventory (FCI)

The independent variables are two different methods of performing teaching which is divided into traditional and active learning. The dependent variable in this study is conceptual learning of Newton's laws, the skill which we will explore the changes in these two domains separately for each group before and after the study's treatments and compare them with each other. To evaluate the effect of the treatment on student learning (dependent variable), some baseline is needed.

The FCI is an instrument used to assess students' beliefs about force and Newton's laws. Students have many common-sense views about motion both before and after formal instruction. Questions on the FCI test were designed to be meaningful to students without formal training in mechanics and target their preconceptions on the subject (Hestenes, Wells, & Swackhammer, 1992).

It is part of a sequence that led to the development of the instrument (Halloun & Hestenes, 1985). Hestenes et al. (1992) designed FCI to probe student beliefs about force and how their beliefs compare with the many dimensions of the Newtonian concept. One of their outstanding virtues is that the questions probe for conceptual understanding of basic concepts of Newtonian mechanics in a way that is understandable to the novice who has never taken a physics course, while at the same time rigorous enough for the initiate. Most physicists would probably agree that a low score on the FCI test indicates a lack of understanding of the basic concepts of mechanics (Hake, 1998).

So, we chose to use FCI to guide us to highlight the addressed misconceptions in questions. We did not focus on diagnosis of conceptions, but by directly referring to questions in FCI to provide an instruction to handle misconceptions (Bulbul & Eryurt, 2011). We gave the FCI post-test unannounced near the final week of classes, and preferably as part of the final exam with significant course credit given for post-test performance. Giving course credit probably motivates students to take the post-test more seriously, and thereby, demonstrate more adequately their understanding, especially if time devoted to the post-test subtracts from time spent on the rest of the final exam (Hake, 2002).

FCI Analyzing Method

The two analyses of pre-test and post-test provide information on different aspects of course effectiveness. A cumulative analysis is used to determine instruction increases the likelihood of students acquiring and retaining baseline knowledge. Marginal analysis is used to determine whether course design is flexible enough to continually rebalance acquisition and retention efforts as student performance changes from one instructional method to the other.

The method used to quantify changes in performance is a definitive feature of any pre-/post- testing design. In the next part, we will state the definitions of Hgain(g), Dgain(G), Dloss, and Retention in analyzing results.

The following index is frequently used to measure the change in group performance from a pre-instruction to a post-instruction test.

$$g = \frac{\begin{Bmatrix} average\ grade\ on\ the \\ post\text{-}instruction\ test \end{Bmatrix} - \begin{Bmatrix} average\ grade\ on\ the \\ pre\text{-}instruction\ test \end{Bmatrix}}{100 - \begin{Bmatrix} average\ grade\ on\ the \\ pre\text{-}instruction\ test \end{Bmatrix}} \tag{1}$$

The ratio in (1) often referred to as normalized change (which in this study, I refer Hgain or Hake gain), expresses the difference between average test scores as a fraction of the maximum possible difference between these scores.

Hake (1998) used (1) to quantify the effectiveness of instructional films. Hake (1998) used (1) to gauge the relative effectiveness of various instructional techniques employed in introductory physics courses. Meltzer (2002) used (1) to explore the relationship between mathematics preparation and concept learning in physics. This important study relied on the intuitive notion that when comparing two courses:

"The course with the larger value of normalized change (g) is the more effective course".

Unfortunately, as Dellwo (2010) demonstrated, this classic assessment rule can lead to counterintuitive conclusions.

David R. Dellwo (2010) employed an alternate assessment rule obtained by decomposing normalized change (1) into component measures:

$$g = G - \gamma L \tag{2}$$

Here, G is a normalized gain (which in this study, I refer Dgain or Dellwo gain) measuring the likelihood that a mistake on the group's pre-instruction test is corrected on the post-instruction test. Similarly, L is a normalized loss (which in this study, I refer Dloss or Dellwo loss) measuring likelihood that a correct response on the group's pre-instruction test is rendered incorrect on the post-instruction test. The non-negative parameter γ is a renormalization factor dependent on the population's pre-instruction performance. Consequently, (2) expresses normalized change (1) as the difference among two non-negative indices, normalized gain, and renormalized loss. The decomposition (2) gives rise to an alternative assessment rule that avoids the counterintuitive conclusions associated with the last result, and reads in part:

"The course with the larger value of normalized gain (G) and the smaller value of renormalized loss (γL) is the more effective course".

Normalized change (1) for a group of N students taking a diagnostic test with M questions can be expressed in the following form:

$$g = \frac{\theta_{post} - \theta_{pre}}{1 - \theta_{pre}} \tag{3a}$$

$$\theta_{pre} = \frac{\begin{Bmatrix} Number\ of\ questions\ students\ answer \\ correctly\ on\ the\ pre\text{-}instruction\ test \end{Bmatrix}}{NM} \tag{3b}$$

$$\theta_{post} = \frac{\begin{Bmatrix} Number\ of\ questions\ students\ answer \\ correctly\ on\ the\ post\text{-}instruction\ test \end{Bmatrix}}{NM} \tag{3c}$$

The derivation of (2) is based on the following observation.

$$\left\{\begin{array}{l} Number\ of\ questions \\ students\ answer\ correctly \\ on\ the\ post\text{-}instruction\ test \end{array}\right\} - \left\{\begin{array}{l} Number\ of\ questions \\ students\ answer\ correctly \\ on\ the\ pre\text{-}instruction\ test \end{array}\right\}$$

$$= \left\{\begin{array}{l} Number\ of\ questions\ students\ answer \\ correctly\ on\ the\ post\text{-}instruction\ test \\ and\ incorrectly\ on\ the\ pre\text{-}instruction\ test \end{array}\right\} - \left\{\begin{array}{l} Number\ of\ questions\ students\ answer \\ incorrectly\ on\ the\ post\text{-}instruction\ test \\ and\ correctly\ on\ the\ pre\text{-}instruction\ test \end{array}\right\}$$

This observation together with definitions (3b) and (3c) imply

$$\theta_{post} - \theta_{pre} = G\,(1 - \theta_{pre}) - L\,\theta_{pre} \tag{4}$$

$$G = \frac{\left\{\begin{array}{l} Number\ of\ questions\ students\ answer\ correctly\ on\ the\ post\text{-} \\ instruction\ test\ and\ incorrectly\ on\ the\ pre\text{-}instruction\ test \end{array}\right\}}{\left\{\begin{array}{l} Number\ of\ questions\ students\ answer \\ incorrectly\ on\ the\ pre\text{-}instruction\ test \end{array}\right\}} \tag{5a}$$

$$L = \frac{\left\{\begin{array}{l} Number\ of\ questions\ students\ answer\ incorrectly\ on\ the\ post\text{-} \\ instruction\ test\ and\ correctly\ on\ the\ pre\text{-}instruction\ test \end{array}\right\}}{\left\{\begin{array}{l} Number\ of\ questions\ students\ answer \\ correctly\ on\ the\ pre\text{-}instruction\ test \end{array}\right\}} \tag{5b}$$

The numerator in (5a) is the number of questions on which students demonstrate a gain in knowledge and the denominator is the maximum possible gain. Consequently, the ratio G is a normalized gain measuring the conditional probability that a mistake on the group's pre-instruction test is corrected on the post-instruction test. Similarly, the numerator in (5b) is the number of questions, on which students demonstrate a loss in knowledge and the denominator is the maximum possible loss. Consequently, the ratio L is a normalized loss measuring the conditional probability that a correct response on the group's pre-instruction test is rendered incorrect on the post-instruction test.

In summary, equation (4) expresses a change in test score as a difference between the fraction of questions on which students demonstrate a gain in knowledge and the fraction on which they demonstrate a loss of knowledge. Finally, to obtain (2) define

$$\gamma = \frac{\theta_{pre}}{1 - \theta_{pre}} \tag{5c}$$

and divide (4) by $(1-\theta_{pre})$. The scaling factor (5c) is a non-negative parameter whose value is larger than 1 if $\theta pre > \frac{1}{2}$, equal to 1 if $\theta pre = \frac{1}{2}$, and smaller than 1 if $\theta < \frac{1}{2}$. The scale γ is referred to as the group's aspect ratio and specifies the odds that the group gives a correct answer on the pre-instruction test.

i. *A is more effective than B if:* $\left\{\begin{array}{c} G_A > G_B \ and \ \gamma_A L_A \leq \gamma_B L_B \\ or \\ G_A \geq G_B \ and \ \gamma_A L_A < \gamma_B L_B \end{array}\right.$ \hfill (6a)

ii. *A and B are equally effective if:* $G_A = G_B \ and \ \gamma_A L_A = \gamma_B L_B$ \hfill (6b)

$$\text{iii. } \textit{A and B are not comparable if: } \begin{cases} G_A > G_B \text{ and } \gamma_A L_A > \gamma_B L_B \\ \quad\quad\quad \text{or} \\ G_A < G_B \text{ and } \gamma_A L_A < \gamma_B L_B \end{cases} \tag{6c}$$

Notice, (6a) restates "The course with the larger value of normalized gain (G) and smaller value of renormalized loss (γL) is the more effective course" in algebraic form and defines a consistent ordering of courses in the sense that if A is more effective than B and B is more effective than C, then A is more effective than C. Also, (6c) offers an assessment option not offered by "The course with the larger value of normalized change (g) is the more effective course", some courses are not comparable.

If A is a more effective course than B in the sense of (6a), then GA-GB is a value-added measure of improved effectiveness due to larger gains (Suskie, 2004). Also, γ is a value-added measure of improved effectiveness due to smaller renormalized losses experienced by students in a more effective course. Consequently,

$$g_A - g_B = \left(G_A - G_B \right) + \left(\gamma_B L_B - \gamma_A L_A \right) \tag{7}$$

is a value-added measure of the total improvement in effectiveness when (6a) or equivalently "The course with the larger value of normalized gain (G) and smaller value of renormalized loss (γL) is the more effective course" applies and one course can claim the larger gains as well as the smaller renormalized losses.

On the other hand, (7) is not a measure of total improvement in effectiveness when (6c) applies and neither course can claim both larger gains and smaller renormalized losses. In this case, one of GA-GB and γBLB-γALA is positive while the other is negative, so (7) is the difference between two value-added measures:

$$g_A - g_B = \left(G_A - G_B \right) + \left(\gamma_B L_B - \gamma_A L_A \right) = \begin{cases} -\left(G_B - G_A \right) + \left(\gamma_B L_B - \gamma_A L_A \right) & \text{if} & \left(G_A - G_B \right) < 0 \\ \left(G_A - G_B \right) - \left(\gamma_A L_A - \gamma_B L_B \right) & \text{if} & \left(\gamma_B L_B - \gamma_A L_A \right) < 0 \end{cases} \tag{8}$$

That is, gA-gB is a difference between added effectiveness due to larger gains in one course and added effectiveness due to smaller renormalized losses in the other course.

Finally, in view of (8), the classic assessment rule: "The course with the larger value of normalized change (g) is the more effective course" declares A more effective than B when either of the following applies:

1. The added effectiveness due to smaller renormalized losses in A offsets the added effectiveness due to larger gains in B;

2. The added effectiveness due to larger gains in A offsets the added effectiveness due to smaller renormalized losses in B.

Data Collection Procedures

Testing the effectiveness of idea by FCI and presenting the data on which it was constructed as Hgain(g), Dgain(G), Dloss, and Retention is the goal of this study. Will use the analysis to show how focusing on the FCI, a multiple-choice test is used to analyze students understanding in Newton's laws. So, we investigate results for two groups of students that can be explained with the analyzing data.

For survey classification and analysis purposes we defined:

1. Average Hgain for both courses as the ratio of the actual average gain to the maximum possible average gain;

2. Average Dgain measuring the likelihood that a mistake on the group's pre-instruction test is corrected on the post-instruction test;

3. Average Dloss, measuring the likelihood that a correct response on the group's pre-instruction test is rendered incorrect on the post-instruction test;

4. Average γ is a renormalization factor dependent on the population's pre-instruction performance and the average value of renormalized loss (γDloss);

5. Comparing averages for both groups to investigate the effectiveness of method;

6. FCI analyzing items for both groups.

Hgain, Dgain, Dloss, and Retention for groups

David Dellwo's paper on "Reassessing Hake's Gain" was presented at the Washington DC *Winter 2010 AAPT Meeting*. In that paper, Dellwo describes some sensible arguments on how to improve upon the Hake gain definition. Dellwo reasons a true "gain" would be one in which the student had pretested a problem incorrectly and posted the problem correctly. Likewise, a true "loss" would be a correctly pretested problem but incorrectly post tested problem. This allows for a true gain (Dellwo gain or "Dgain"—our definition, not his) and true loss (Dellwo loss or "Dloss"). Lastly, the student correct pre- and post- test problems are defined as retention "R%".

In Table 1 for the average of the class are included.

Table 1

Class Average Hgain, Dgain, Dloss, Retention, Pre-test, Post-test for the Experimental Group

FCI results for class		Hgain (%)	R (%)	Dgain (%)	Dloss (%)	Pre-test (%)	Post-test (%)
The experimental group	Class Av.	24	46	32	54	13	34
	St. Dev.	15	29	13	29	7	12
	Count	21	21	21	21	21	21

Notes. Pre-test = 13% ± 7% s.d.;

    Post-test = 34% ± 12% s.d.;

    Hgain = 24% ± 15% s.d. (pre-tested incorrectly and posted correctly);

    Dloss = 54% ± 29% s.d. (correctly pre-tested but incorrectly post-tested it), (+, -);

    Dgain = 32% ± 13% s.d. (incorrectly pre-tested and posted it correctly), (-, +);

    R = 46% ± 29% s. (both pre- and post- test problems), (+, +).

In analysis columns in Table 2. These new calculations for the average of the class are included.

Table 2

Class Average Hgain, Dgain, Dloss, Retention, Pre-test, Post-test for the Control Group

FCI results for class		Hgain (%)	R (%)	Dgain (%)	Dloss (%)	Pre-test (%)	Post-test (%)
The control group	Class Av.	7	25	20	75	15	21
	St. Dev.	6	26	7	26	5	7
	Count	20	20	20	20	20	20

Notes. Pre-test = 15% ± 5% s.d.;

    Post-test = 21% ± 7% s.d.;

    Hgain = 7% ± 6% s.d. (pre-tested incorrectly and posted correctly);

    Dloss = 75% ± 26% s.d. (correctly pre-tested but incorrectly post-tested it), (+, -);

    Dgain = 20% ± 7% s.d. (incorrectly pre-tested and posted it correctly), (-, +);

    R = 25% ± 26% (both pre- and post- test problems), (+, +).

Richard Hake found that for Interactive Engagement (IE) and Traditional (T) introductory mechanics courses: The normalized gain for IE courses was <Hgain>IE = 0.48 ± 0.14sd.

That is over twice the gain for traditionally taught courses: <Hgain>T = 0.23 ± 0.04sd.

Jeff Saul and Edward "Joe" Redish found that for traditionally taught college algebra-based and calculus-based physics courses, <Hgain> is typically 0.2. The definition of Hgain levels are as follows (Hake, 1998):

(a) "High-g" courses as those with (<Hgain>) > 0.7;

(b) "Medium-g" courses as those with 0.7 > (<Hgain>) > 0.3;

(c) "Low-g" courses as those with (<Hgain>) < 0.3.

In this study, we found that for traditionally taught 10th-grade high school physics course, Hgain is typically 0.07 and for ICI method for the experimental group, Hgain is typically 0.24 (much lower than in Hake's article). Remember: These students are only 20% or so of the class that is "good at math", but the improvement in Hgain from traditional method to ICI is high. This is over triple the gain for traditionally taught the course.

Average Hgain reported here suggests that the ICI approach was very successful in promoting learning. Furthermore, we are aware of the difference between the learning Hgains achieved through ICI and those achieved through earlier more traditional approaches to teaching, but is it a sign for a course to be effective more than other? Depending on Dellwo's (2010) explanation, it is not a complete sign for being successful in a course by itself. So, the average values of Dgain and γ. Dloss for traditional and ICI courses are determined. Although the research Hgain results adopted does not permit direct comparisons in learning gains between matched control and experimental classes subjected to "new" and "traditional" teaching, we believe that comparisons with FCI data gathered in research studies internationally allow us to claim relative success for the ICI teaching approach in Dellwo's idea. In the next part, we will calculate the reported Dellwo's components to give more satisfaction in the result.

Table 3

Results for Both Groups

	Pre-test	Post-test	Hgain	Dgain	Dloss	Ret	γ	γ. Dloss
The experimental group	0.13	0.34	0.24	0.32	0.54	0.46	0.149	0.084
The control group	0.15	0.21	0.07	0.20	0.75	0.25	0.176	0.132

It is important to consider differences in student populations when comparing the normalized gains of different classes. For example, it might be incorrect to conclude that teaching methods used in an active teaching class with a normalized gain of 0.6 are necessarily more effective than those that produce a gain of 0.3 in a different class. The backgrounds of the students in the two classes could be a more important factor than the specific methods used in the classes (Coletta & Phillips, 2005). The probability of guessing correct answers for the FCI questions for both groups (13% for the experimental group and 15% for the control group) in pre-test are about 14%. It shows that the students' initial knowledge and understanding of mechanics were very poor and they are the same. This was by no means a surprise, because the students had received only limited instruction in mechanics in lower secondary school. Successive differences in nominal values of Hgain listed in Table 3. Quantify the added effectiveness. For example, the value 0.54 for the Dloss in ICI group means that 54% of diagnostic questions answered correctly on pretest were answered incorrectly on post-test and the value 0.32

for the Dgain means that 32% of diagnostic questions answered incorrectly on pre-test were answered correctly on the post-test. In addition data in Table 3 indicates that Dloss decreased from traditional instructional method to the ICI, and also renormalized gain decreased with γ. Dloss = 0.084. Although nominal values of γL from traditional method to ICI decreased slightly, the increase in Dgain is large enough to be statistically significant.

In the previous discussion was explained that "The course with the larger value of normalized gain (Dgain) and the smaller value of renormalized loss (γ. Dloss) is the more effective course" in algebraic form. It does not mean by "The course with the larger value of normalized change (Hgain) is the more effective course".

If A is a more effective course than B in the sense of (6a), then GA-GB is a value-added measure of improved effectiveness due to larger gains. Also, γ is a value-added measure of improved effectiveness due to smaller renormalized losses experienced by students in a more effective course. Consequently,

$$g_A - g_B = (G_A - G_B) + (\gamma_B L_B - \gamma_A L_A) \tag{7}$$

is a value-added measure of the total improvement in effectiveness when equivalently "The course with the larger value of normalized gain (Dgain) and smaller value of renormalized loss (γ. Dloss) is the more effective course" applies and one course can claim the larger gains as well as the smaller renormalized losses. So,

$$g_{(exp)} - g_{(cont)} = (Dgain_{(exp)} - Dgain_{(cont)}) + (\gamma.Dloss_{(cont)} - \gamma.Dloss_{(exp)})$$
$$= (0.32\text{-}0.20) + (0.132 - 0.084) = 0.168 \tag{8}$$

Inspection of Table 3 reveals that Dgain was larger from the traditional method to ICI. In addition, the renormalized loss was smaller from traditional to ICI. Consequently, the experimental group was more effective in promoting learning ICI during the instructional periods than the control group by an amount equal to the difference in Δg = 0.168. It shows the ICI course is more effective than the traditional course.

Although Dellwo's gain points to the overall success of the teaching, these data do not provide information about specific learning difficulties that students may have encountered in coming to an understanding of the force concept.

Conclusions

In this article, we suggest that using Dellow Gain for developing FCI data analyze. We see the results for Dellow Gain were similar to Hake Gain. Although this article has focused on the comparing of two methods, the results recognize the successes of ICI methodology in learning physics. Successive differences in nominal values of Hgain quantify the added effectiveness. Average Hgain reported here suggests that the ICI approach was very successful in promoting learning.

On the other hand, analyzing indicates that Dloss has decreased from traditional instructional method to the ICI approach, and renormalized Dellow Gain has decreased. Although nominal values of γL from traditional method to the ICI approach decreased slightly, the increase in Dgain is large enough to be statistically significant.

References

Azita, S. F. (2014). An investigation about traditional teaching in high school on conceptual understanding in force and motion in Iran. *Pensee Journal, 76*(5).

Azita, S. F., & Cesar, M. (2015, January). An investigation about misconceptions in force and motion in high school. *US-China Education Review A, 5*(1), 38.

Bulbul, M. S., & Eryurt, K. (2011). Adapting force concept inventory test into drama applications. In *Western Anatolia Journal of Educational Sciences (WAJES)*, Special Issue: Selected papers presented at WCNTSE, 37.

Coletta, V. P., & Phillips, A. J. (2005, December). Interpreting FCI scores: Normalized gain, pre-instruction scores, and scientific reasoning ability. *American Journal of Physics, 73*(12), 1172.

Dellwo, R. D. (2010). Course assessment using multi-stage pre/post testing and the components of normalized change. *Journal of the Scholarship of Teaching and Learning, 10*(1), 55-67.

Gabel, D. (2003). Enhancing the conceptual understanding of science. *Winter (Educational Horizons), 70*,

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

Hake, R. R. (2002, December 2-4). Assessment of physics teaching methods, assessment of physics teaching methods. In *Proceedings of the UNESCO-ASPEN Workshop on Active Learning in Physics*, University of Peradeniya, Sri Lanka. Retrieved from http://www.physics.indiana.edu/~hake/ ASPEN = Asian Physics Education Network http://www.swin.edu.au/physics/aspen/welcome.htm

Halloun, I., & Hestenes, D. (1995). Common sense concepts about motion. *American Journal of Physics, 53*, 1056.

Hestenes, D., Wells, M., & Swackhammer, G. (1992). Force concept inventory. *Physical Technology, 30*, 141-158.

McDermott, C. L. (1998). Students' conceptions and problem-solving in mechanics: Connecting research in physics education with teacher education. In T. Andrée, E. L. Jossem, and B. Jorge (Eds.), *International commission on physics education* (p. 42). All rights reserved under International and Pan-American Copyright Conventions.

Meltzer, D. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics, 70*(12), 1259-1268.

Savinainen A., & Scott, P. (2002). Using the Force Concept Inventory to monitor student learning and to plan teaching. *Physics Education, 37*(1), 0031-9120.

Suskie, L. (2004, June 4). *Assessing student learning: A common sense guide*. Publisher: Wiley.