Learning Styles and Physics Achievement in Secondary Education

PGO by Jelle Aalbers, Karel Kok and Koert Poelman

Abstract. As teachers, one of our most important tasks is guiding our students' study behavior. While there is ample (and sometimes contradictory) literature on study advice, very little is known about which study behaviors actually lead to success in specific subjects at specific levels. We explore the correlation between physics achievement in 3/4 HAVO/VWO schools and components of a standard learning strategies & motivation test, LEMO. We find that a memorization-oriented learning style is negatively correlated with grades, especially for girls. No learning style was found to correlate positively with grades, leading us to question the utility of the LEMO-test or even the significance of the conventional taxonomy of learning styles when it comes to physics.

Learning Styles and Physics Achievement in Secondary Education

July 1, 2012

Thesis by:
Jelle Aalbers, Karel Kok and Koert Poelman

Centrum voor onderwijs en leren

Supervisor:
prof. dr. Jan Vermunt

Utrecht University
Centrum voor onderwijs en leren
Postbus 80127,
508 TC Utrecht
the Netherlands
Jelle Aalbers: `j.aalbers@students.uu.nl`
Karel Kok: `k.kok@students.uu.nl`
Koert Poelman: `k.poelman@students.uu.nl`

Contents

Chapter 1

Motivation

As teachers, we know that students each study in their own way, yet some are more efficient at it. We all like to see our students succeed, and pity especially those students that put forward their best efforts yet fail to achieve nonetheless. Particularly from these students, we often get the question “how should I study for physics?”, yet find ourself short of an appropriate answer. To wit, there is no lack of study advice literature, but much of it is general, or contradictory, or not backed up by evidence. In teacher training courses, the focus is usually on what the teacher should do, rather than what the pupils should do.

Thus it is natural to inquire which study behavior constitutes effective learning in our specific subject (in our case, physics). This knowledge, we hope, will be useful in our classrooms and those of our colleagues as we advise and give guidance to our pupil’s study behavior.

Chapter 2

Theory

The field of pupil learning studies is a conceptual goldmine, but an etymological quagmire: different authors use different meanings for the same key terms. We will here follow Derry in distinguishing between learning strategies and tactics Derry (1989), and Vermunt (1994) in the definition of learning styles, see Table 2.1.

Term	Meaning	Example
Learning Style	Disposition towards preferring certain types of learning strategy	meaning-directed, reproduction-directed
Learning Strategy	A plan to accomplish a learning goal (Derry, 1989)	Using a structured reading strategy such as READS (Review, Examine, Ask, Do, Summarize)
Learning Tactics	“specific learning skills” required to accomplish a learning strategy (Derry, 1989)	Reviewing, summarizing, etc.

Table 2.1: lexicon of presently employed definitions of key terms.

The conceptual framework implied by this table is visualized in figure 2.1. In this study, we focus primarily on learning styles. According to Vermunt (1994), learning styles have four components: cognitive, metacognitive, dispositional and epistemological (“opvatting over het leren”) components. This is similar to the components discerned by Woolfolk & Perry (2006). In his research, Vermunt uncovered four dominant learning strategies in secondary education, see Table 2.2. The operational meaning of the terms in this table is further elucidated by their instantiation in the LEMO-test, which we discuss below.

Much has been written about the use of learning strategies in physics, yet the majority of these scriptures take the form of study advice rather than empirically validated research. As such, the literature on how to study physics appears rife with generalities and commonplaces interposed with genuine insights about the subject-specific challenges and how to confront them – the distinction is often in the eye of the beholder.
As two representatives of this large body of literature, we chose an essay by Hollabaugh

	Meaning-directed	Reproduction-directed	Application-directed	Undirected
Cognitive	Reflectivity, critical thought	Reductive analysis, mnemonic aids	Concretisation, example formulation	(barely any)
Metacognitive	Self-directed	Other-directed	Both	Neither
Dispositional	Personal interest	Achievement motive	Professional relevance	Ambivalent
Epistemological	Knowledge as a relational construct	Knowledge as a series of facts	Knowledge as skills to be mastered for use	Ambivalent

Table 2.2: dominant learning styles in secondary education, paraphrased from (Vermunt, 1994).

(2004) because it is often reprinted on the web and in books, including one of the most popular University Physics textbooks Young & Freedman (2004) and a pamphlet by Chapman (1949), because of its antiquity and breadth of scope.

Much of the advice in these documents concerns general study habits, advice on note-taking and problem-solving strategies. A more exhaustive comparison of these documents, which is available from the authors on request, reveals that both Chapman and Hollabaugh consider the essential characteristic of physics the inevitability of conceptual understanding and the importance. Hollabaugh quips:

> *"Physics continually builds on fundamental ideas and it is important to correct any misunderstandings immediately."*

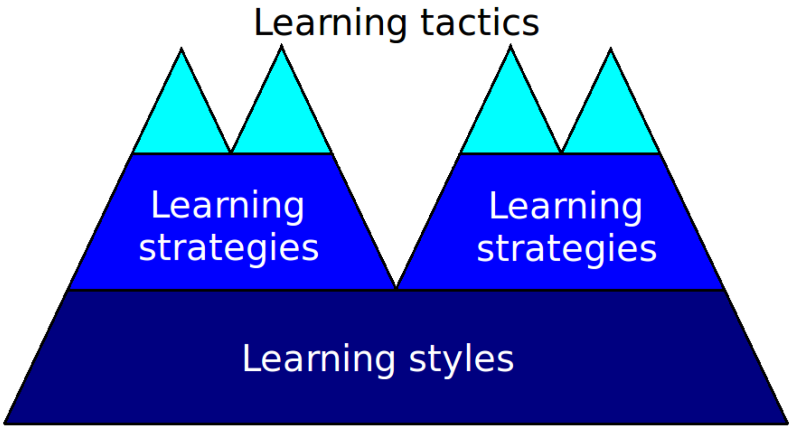

Figure 2.1: Learning styles, strategies and tactics

Chapman states it even more strongly:

> *"You may or may not be able to bluff your way through an essay question in economics but definitely you cannot do it in a physics problem. Either you can reason how to do the problem or you can't."* (Chapman, 1949)

The case for this proposition may be examined in light of a key dilemma in physics teaching (and learning): whether to study a law's derivation or simply memorize its algebraic expression ("the formula"). Chapman offers four reasons why the former is to be preferred, in paraphrase:

1. Derivations tell you the limits of validity of a formula (because in the derivation, simplifying assumptions are often made)
2. They will illustrate fundamental principles
3. They will allow you to recover the formula if you forget it on the exam
4. Deriving technical knowledge is a skill that will be useful even when the knowledge itself has become obsolete.

Another key influence in conventional understanding of how students (do not) learn physics has been the development of the Force Concept Inventory (Hestenes & Swackhammer, 1992). This FCI is a test designed to probe student understanding of mechanics concepts using straightforward multiple-choice questions. The results of the FCI often show that few students, even those that perform well on standard examination problems, understand the basic concepts, and even fewer retain them beyond the exam. Intuitive, Aristotelian notions – heavier objects fall faster, circular motion maintains itself, motion always requires a force – are akin to bothersome weeds, despite sincere eradication efforts, they resurface time and time again.

The test has taken the position of a benchmark of the degree to which a physics course succeeds in developing student understanding, as evidenced by its inclusion in teacher training courses such as Marjolein Vollebrecht's course at Utrecht University and Sanjoy Mahajan's course at MIT, Mahajan (2009) and its use in scientific literature, to show, among others, the effectiveness of interactive teaching in physics (Hake, 1988) or the limited transfer that occurs from practicing the same type of physics problem repeatedly (Kim & Pak, 2002).

Chapter 3

Hypothesis

The above considerations, especially Chapman and Hollabaugh's essays as well as the Hestenes e.a. and the Kim & Pak results, appear to point to a hypothesis: meaning-directed learning, rather than reproduction-directed learning, is the learning style which predominantly leads to success in physics. The recommended learning strategies that associate with this style are mainly structured problem-solving (with an emphasis on reflection after the problem is done, especially if mistakes were made), and to a lesser extent, conceptually structured note-taking and repeated revisiting of course readings.

Since this hypothesizes a causal relationship, the test would have to be a randomized trial, not an observational study. We will describe a suggestion towards such an experiment in the discussion, but due to time and budget constraints, only an self-report study is feasible for us. As correlation does not imply causation, and neither does lack of correlation imply lack of causation, the experiment would have inadequate falsificational power vis-à-vis the hypothesis of interest.

We therefore grudgingly formulate a more humble hypothesis: that a meaning-directed learning style is correlated positively with physics achievement in secondary education. This is by no means an empty statement, nor of negligible value to education. If this more humble hypothesis is falsified, it implies that either the grand hypothesis (meaning-directed learning leads to better results) is false, or that it is true, yet other interfering factors are present. We will further elaborate on this in the discussion.

Chapter 4

Methodology

Variables

The variables of interest in our study are achievement in physics in secondary education and learning style. We will instantiate physics achievement as the mean z-score on physics classroom tests conducted by students in pursuit of the higher levels of Dutch secondary education (HAVO/VWO). Since our subjects study at different schools, they take different classroom tests, hence the z-scores should be calculated relative to each group which does take the same test. The score thus obtained is denoted by #.

A more rigorous approach to measuring physics achievement would be to administer a standard test, such as the FCI discussed above, in all classrooms. We have chosen differently for two reasons. First, while taking a standardized tests might make it easier to measure physics insight, it would frustrate a measurement of what we are truly interested in: the correlation between learning styles and achievement; especially since we are measuring in different classrooms, we cannot escape the presence of a multitude of interfering factors such as teacher quality and school culture which in usual grading systems are usually corrected for by renormalization. Second, we had to deal with time constraints.

Learning style, being inherently multidimensional, will be instantiated as the vector whose components are z-scores on the A-B scales of the LEMO test described in the subsequent section.

Instrument

To measure the learning style of students, we use the LEMO-test developed by Donche, Van Petegem, Van de Mosselaer and Vermunt (Donche & Vermunt, 2010). This test was originally derived from the Inventaris LeerStijlen (ILS) by (Vermunt, 1994), an extensively validated questionnaire, whose length was however experienced as burdensome by some users in education. Table 4.1 indicates which components of LEMO measure which component of Vermunt' s learning style model. The LEMO-test has reduced the number of questions to 53.

LEMO has been deployed, among others, to first year students of 2006-2007 at the Plantijn Hogeschool in Antwerpen and later to several other schools. With a database of

977 respondents, the LEMO test was validated (Crohnbach's $\alpha \geq 0.68$) on construct validity, one-dimensional validity and predictability. For example, it was found that students with an undirected learning style usually had poor grades, while internal motivation was associated with deep learning and external motivation with surface learning.

However, no conclusions could yet be drawn regarding the success of meaning-oriented versus reproductive-oriented learning, let alone the success of learning styles in specific subjects.

We have used a version of LEMO adapted (mainly linguistically) to its use in secondary education.

	Meaning orientation	Reproductive orientation	Applied orientation	Disorientation
Cognitive component	A1, A2	A3, A4	A5	
Metacognitive component	B1	B2	B1, B2	B3
Learning orientation	C2	C1		C3

Table 4.1: implementation of Vermunt's learning style model in LEMO, adapted from (Vermunt, Donche, & Petegem).

For a more detailed description and analysis of these components, we refer to (Donche & Vermunt, 2010).

Subjects and conditions

The LEMO test has been taken by 6 different HAVO/VWO classes which have been taught by the authors during the 2011-2012 academic year. Two HAVO 4 classes taught by Karel Kok, two VWO 4 classes taught by Jelle Aalbers and two VWO 3 classes taught by Koert Poelman. We have produced a digitally available version of LEMO as an excel file for our students to fill in. We have e-mailed our students this file, and have collected the data. Of a total of 170 pupils 66 have responded, a response rate of $0,4$. We could have reached a higher response rate by having our pupils take the LEMO test in the classroom. We chose this digital method of taking the test for convenience, since data analysis if far easier this way. Also, since the LEMO test was taken at the end of the academic year, time was valuable.

Chapter 5

Results

Table 5.1 shows our key results: Pearson's r for # and any of the LEMO-components and questions along with the associated two-tailed p-value associated with the null-hypothesis $r = 0$. Table 5.2 shows the mean scores of various subgroup divisions on each of the LEMO-components and questions, along with the difference in their means expressed in pooled standard deviations

$$\sigma_{\text{pooled}} = \sqrt{\frac{(n_A - 1){\sigma_A}^2 + (n_B - 1){\sigma_B}^2}{n_A + n_B - 2}}$$

where A and B stand for the various subgroups, along with the two-tailed p-value associated with the null-hypothesis $\mu_A - \mu_B = 0$. We chose only to report data which is supported by a p-value bigger than 0.05.

Pearson's r is the cosine of the angle between the z-score vectors of both variables (in the space where each subject is a dimension), or equivalently, the standard deviation of the product of the variates divided by the product of the standard deviation of the variates, or equivalently, the cosine of the angle between. Intuitively, it is a measure of the goodness-of-fit of a linear regression. More formally, it is the cosine of the angle between the lines resulting from converse pairwise linear regressions of the data.

A p-value is the lowest type-I error probability threshold which still implores a rejection of the null hypothesis. For a two-sided test, if the test statistic t follows a t-distribution with d degrees of freedom and is centered on 0 if the null hypothesis is true, the p value is thus

$$p = 2(1 - \tau_d(|t|))$$

where τ_d is the cumulative t-distribution with d degrees of freedom.

For the correlation tests, we do not know in advance either the variance of the data or whether the correlation will turn out positive or negative, hence the appropriate, and at any rate the conventional, significance test to use is a two-sided test of the null hypothesis $r = 0$. For our tests, we used the statistic

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

which follows a t-distribution with $(n - 2)$ degrees of freedom centered on $r = 0$ if the populations are normally distributed and the actual correlation is 0 (Montgomery&Runger 1991 p 463).

For the mean comparison tests, a two-tailed test of is again appropriate for the same reason. We use the statistic

$$t = \frac{\mu_A - \mu_B}{\sqrt{\sigma_A/n_A + \sigma_B/n_B}}$$

which, if the population is normally distributed and the means are equal approximately follows a t-distribution centered around 0 with degrees of freedom (Montgomery & Runger p.342)

$$d = \frac{\left(\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}\right)^2}{\frac{\left(\frac{\sigma_A^2}{n_A}\right)^2}{n_A - 1} + \frac{\left(\frac{\sigma_B^2}{n_B}\right)^2}{n_B - 1}}$$

We did not opt for the simpler and t-exact pooled-σ t-test because it is quite sensitive to inequality of the variances of the two subgroups, and because the invention of practical Turing machines for the consumer market reduces he migrainogenic potency of substantial repetitive computational numerals.

Table 5.3 shows the calculated value's of Cronbach alpha for the various LEMO-test aggregate items

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} {\sigma_i}^2}{\sigma^2}\right)$$

where k is the number of constitutive items, ${\sigma_i}^2$ the variance of the scores on the constitutive items and σ^2 the variance of the aggregated score.

After distributing the questionnaire, some students contacted us having noticed a few confusing questions. Their concerns pointed us to spelling errors introduced in the questionnaire due to manual copying the questions from an analog source using inadequately immaculate monks, which as an unintended side-effect of over reading a certain science-fiction book are only capable of transcoding 42 questions without introducing errors which materially affect interpretation. As a result, in question 43, "Ik zie *niet* in waarom ik eigenlijk studeer en ik maak mij daar geen zorgen om" (emphasis added), the emphasised word was accidentally left out, negating the intended meaning of the question. We have resolved this matter by reverse-coding the score of this question, i.e. to note the score as six minus the question score. This does not resolve the issue completely: the question as it was given to the students is more confusing, leading potentially to less extreme values in the responses. Unfortunately, question 43 was also the one question given double weight by LEMO in the calculation of its corresponding aggregate item (C3). Fortunately, as can be seen from table 5.3, Cronbach's alpha does not seem to be affected, giving some measure of confidence that this issue, though perilous and pernicious did not become perditious.

Vraag	Iedereen		Jongens		Meisjes		HAVO		VWO		3^e klas		4^e klas	
	r	p	r	p	r	p	r	p	r	p	r	p	r	p
A4 Memoriseren	-0.38	0.00^1			-0.45	0.01			-0.43	0.01	-0.49	0.00^1		
A2 Kritisch verwerken			0.36	0.05										
B3 Stuurloos	-0.41	0.00^1			-0.44	0.01	-0.51	001	-0.34	0.03			-0.47	0.00^1
D1 Zelfeffectiviteit	0.40	0.00^1			0.49	0.00^1	0.54	0.00^1	0.35	0.03			0.55	0.00^1
9. Ik zoek uit wat de overeenkomsten en verschilen zijn tussen de onderwerpen die in een les worden behandeld	0.26	0.03					0.50	0.01					0.51	0.00^1
12. Ik zet de belangrijkste feiten op een rijtje en die leer ik dan uit mijn hoofd.	-0.31	0.01			-0.36	0.03								
16. Ik constateer dat ik moeite heb met het leren van de leerstof	-0.31	0.01			-0.45	0.01	-0.44	0.02					-0.38	0.02
20. Definities leer ik zo letterlijk mogelijk van buiten	-0.30	0.02			-0.37	0.03			-0.33	0.04				
25. Ik constateer dat ik het moeilijk vind om te bepalen of ik de leerstof voldoende beheers	-0.40	0.00^1			-0.40	0.02	-0.47	0.02	-0.34	0.03			-0.49	0.00^1
28. Ik constateer dat ik niet goed begrijp wat de leerkracht bedoelt als hij aangeeft wat ik moet leren	-0.31	0.01			-0.38	0.02							-0.34	0.05
29. Ik prent de betekenis van alle begrippen die ik niet ken stuk voor stuk in mijn hoofd	-0.37	0.00^1			-0.43	0.01			-0.39	0.01	-0.45	0.01		
51. Ik denk dat ik vrij goed ben in studeren vergeleken met sommige andere leerlingen	0.30	0.02			0.36	0.03			0.38	0.01				
52. Ik ben tevreden op het gebied van bekwaamheid van studeren	0.36	0.00^1			0.46	0.01	0.63	0.00^1					0.57	0.00^1
53. Ik heb vertrouwen in de manier waarop ik studeer	0.42	0.00^1	0.42	0.02	0.43	0.01	0.63	0.00^1					0.60	0.00^1

Table 5.1: Significant pearsons-r for correlations with # (grades). ^{1} indicating $p< 0.005$.

Vraag	J-M		HAVO-VWO		3^e-4^e	
	σ	p	σ	p	σ	p
A4 memoriseren	-0.54	0.03				
B1 zelfsturing	0.60	0.02				
B3 stuurloos	-0.57	0.02				
C3 demotivatie					-0.53	0.03
D1 zelfeffectiviteit					0.75	0.00[1]
4. Ik gebruik wat ik in de les leer bij mijn activiteiten buiten school	0.74	0.00[1]				
5. Ik lees naast de verplichte leerstof ook andere boeken of teksten die met de leerstof te maken hebben	0.76	0.00[1]				
10. Ik analyseer de afzonderlijke feiten van de theorie stap voor stap			0.66	0.01	-0.59	0.02
11. Ik zet de belangrijkste feiten op een rijtje en die leer ik dan uit mijn hoofd			0.54	0.03		
24. Ik constateer dat ik het moeilijk vind om te bepalen of ik de leerstof voldoende beheers	-0.71	0.00[1]				
30. Met behulp van wat ik op school leer, bedenk ik oplossingen voor praktische problemen	0.82	0.00[1]				
33. De extra uitleg en aanwijzingen die de leerkracht tijdens de les geeft, zijn een onmisbare hulp bij het leren			0.60	0.01	-0.56	0.03
36. Omdat ik nieuwe dingen wil bijleren	0.53	0.03				
37. Eerlijkgezegd, weet ik het niet; ik heb het gevoel dat ik mijn tijd verdoe op school					-0.68	0.01
47. Omdat ik dit persoonlijk zeer waardevol vind					0.64	0.01
52. Ik ben tevreden op het gebied van bekwaamheid van studeren					0.51	0.04
52. Ik heb vertrouwen in de manier waarop ik studeer					0.74	0.00[1]

Table 5.2: Differences between means of various subgroups expressed in pooled standard deviations.
[1] indicating p< 0.005.

A1	Relateren en structureren	0.77
A2	Kritisch verwerken	0.81
A3	Analyseren	0.74
A4	Memoriseren	0.79
A5	Concrete verwerking	0.65
B1	Zelfsturing	0.60
B2	Externe sturing	0.74
B3	Stuurloos	0.68
C1	Moeten studeren	0.71
C2	Willen studeren	0.85
C3	Demotivatie	0.92
D1	Zelfeffectiviteit	0.80

Table 5.3: Cronbach's alpha's for the different learning scales

Chapter 6

Discussion

The negative correlation between A4 ('memoriseren') and # is consistent with our hypothesis that it is meaning-directed learning rather than reproduction-directed learning that leads to success in physics. The negative correlation between A4 and # is particularly pronounced for girls, even being significant of its own. For boys this is not the case. It should be noted that girls score significantly higher on A4 than boys, leading us to believe that memorization is particularly harmful when used in excess.

However, contrary to our hypothesis, none of the correlations between the other A-components and # were strong enough to reach an acceptable level of significance. In fact, the only question on the learning style (A) components of LEMO which correlated at all positive at a $p > 0.05$ level is question 9, and it correlates quite weakly still. The only qualifying observation we can make is that for boys alone, component A2 does seem to correlate positively with grades, though with only borderline significance.

The correlations between # and the components B3 and D1 are not surprising, as they measure self-reported confidence in students' ability to study. Common sense tells us negative grades cause low confidence in study ability, and this would likely result in the correlation. To what degree the more interesting effect, low confidence in study ability causing low grades, is also present, this study cannot discern.

Another interesting result is that boys report greater intrinsic motivation and greater confidence in their study ability. As we also see them using less of the apparently ineffective memorization learning style, it is somewhat surprising that boys did not have significantly higher grades than girls.

Comparing classes from year 3 and 4, we see that in year 4 students have greater motivation yet less confidence in their ability. They key factor here would seem to be two qualitative changes in how the subject is offered: it is no longer mandatory, and its content is delivered at a higher pace.

Further research

To further analyze this data, one could perform a factor analysis on the raw question scores and the grades and try to construct new aggregate components to replace the LEMO scores, with special focus to explaining the variation in grades. Even better, one could perform this procedure on a more extensive learning styles test, such as Vermunt's ILS. If this analysis reveals an internally consistent component which correlates strongly with grades, it could give a powerful clue as to what study behaviors are effective for physics.

Another way to extend this study is to take a more in-depth approach: to observe or interview students and code the transcripts for relevant learning style, strategy, and tactic clues. Correlating these to grades, we would obtain a powerful insight into what successful students do in physics. However, such a study naturally faces scaling constraints: interviews and observations are time-consuming when applied to a large sample. Combining these recommendations, one could first do observations and interviews, arrive at a preliminary conclusion, then validate this using a questionnaire.

As mentioned above, our experiment is insufficient to try and falsify the grand hypothesis that meaning-directed learning leads to better physics achievement. To do this, one could imagine a randomized trial in which the results of an experimental group of classes which receive instruction in meaning-directed learning are compared to a control group which receives instruction in reproduction-directed learning, or even no instruction at all. Ideally, this would be conducted with a large group of similarly constituted classes, and by teachers which are uninvolved with the experiments, unaware of our hypothesis, and which have measurably similar instructional aptitude. To further improve methodological soundness, a standardized test to measure physics prowess could be used in both groups, or the learning style instruction could be provided to students by texts or video lectures so even the teachers themselves are unaware of the experimental conditions.

Despite the substantial rigor of such a test, it would still faces several challenges, chiefly that instruction in learning styles may not be effective, especially for high-achieving students that have already achieved success with a learning style of their own, and/or low-motivation students which are not willing to experiment with different learning styles regardless of the potential benefit.

Chapter 7

Conclusion

While LEMO did allow us to identify a learning style which correlates with low grades (memorization), it did not allow us identify to identify a learning style which actually does correlate which high grades. If our experiment was adequate, then either the learning style which is associated with high grades in physics cannot be measured (adequately) by LEMO, or learning styles as such are not an important determinant for physics achievement.

Bibliography

Chapman, S. (1949). How to study physics. *Cambridge, Mass: Addison-Wesley Publishing Company*.

Derry (1989). Putting Learning strategies to work. *Educational Leadership*, *46*.

Donche, P. M. H., V. Petegem, & Vermunt, J. (2010). LEMO: een instrument voor feedback over leren en motivatie. *Plantyn: Mechelen*.

Hake (1988). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.*, *66*.

Hestenes, M., D. Wells, & Swackhammer, G. (1992). Force Concept Inventory. *The physics teacher*, (p. 141).

Hollabaugh, M. (2004). How to succeed in phyics by really trying. *R. A. Hugh D Young, University Physics, 11th edition*.

Kim, E., & Pak, S. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *Am. J. Phys.*, *70*.

Mahajan, S. (2009). Teaching College-Level Science and Engineering. *MIT Open CourseWare*.

Vermunt, J. (1994). Inventory of learning styles in higher education. *Leiden University, The Netherlands: ICLON-Graduate School of Education*.

Woolfolk, P., A.E. Winne, & Perry, N. (2006). *Educational Psychology*. Pearson.

Young, H. D., & Freedman, R. A. (2004). University Physics. *San Francisco: Addison-Wesley*.