

Running head: DEEP APPROACHES TO LEARNING

**Measuring Deep Approaches to Learning
Using the National Survey of Student Engagement**

Thomas F. Nelson Laird
Assistant Professor
Indiana University Center for Postsecondary Research
1900 East Tenth Street
Eigenmann Hall, Suite 419
Bloomington, IN 47406-7512
tflaird@indiana.edu
Phone: 812.856.5824

Rick Shoup
Research Analyst
Indiana University Center for Postsecondary Research

George D. Kuh
Chancellor's Professor and Director
Indiana University Center for Postsecondary Research

Paper presented at the Annual Meeting of the Association for Institutional Research,
May 14 – May 18, 2005
Chicago, IL

Abstract

**Measuring Deep Approaches to Learning
Using the National Survey of Student Engagement**

The concept of deep learning is not new to higher education. However, deep learning has drawn more attention in recent years as institutions attempt to tap their student's full learning potential. To more fully develop student talents, many campuses are shifting from a traditional passive, instructor-dominated pedagogy to active, learner-centered activities. Using exploratory and confirmatory factor analysis on multiple years of data from the National Survey of Student Engagement, this study examines the structure and characteristics of items about student uses of deep approaches to learning. Institutions and researchers can use the resulting scales to assess and investigate deep approaches to learning.

Measuring Deep Approaches to Learning Using the National Survey of Student Engagement

Colleges and universities are devoting substantial effort to design active, learner-centered teaching and learning environments. Findings from the National Survey of Student Engagement (NSSE) (2000, 2001, 2002, 2003, 2004, 2005) suggest that these efforts are paying off in that the vast majority of students at least “sometimes” engage in various forms of active and collaborative learning activities during a given academic year. A fundamental goal of the redesign is to utilize student’s full learning potential, which is not often captured through traditional pedagogical methods. In particular, the shift from passive, instructor-dominated pedagogy to active, learner-centered activities promises to take students to deeper levels of understanding and meaning as they apply what they are learning to real life examples in the company of others (Lave & Wegner, 1991, Tagg, 2003).

The concept of “deep” learning has drawn more attention in recent years as institutions attempt to tap their student’s full learning potential. Deep learning is not a new concept in higher education. Much of the research on deep learning stems from the seminal research of Marton and Säljö (1976). A key notion in deep learning is that students take different approaches to learning, with the outcomes of learning closely associated with the chosen approaches (Ramsden, 2003). Two common approaches to learning are “surface” and “deep” processing (Beattie, Collins, & McInnes, 1997).

Students using “surface-level processing” focus on the substance of information and emphasize rote learning and memorization techniques (Biggs, 1989; Tagg, 2003). The goal of studying for a test or exam is to avoid failure, instead of grasping key concepts and

understanding their relation to other information and how the information applies in other circumstances (Bowden & Marton, 1998).

In contrast, students using “deep-level processing” focus not only on substance but also the underlying meaning of the information. Scholars (Biggs, 1987, 2003; Entwistle, 1981; Ramsden, 2003; Tagg, 2003) generally agree that deep learning is represented by a personal commitment to understand the material which is reflected in using various strategies such as reading widely, combining a variety of resources, discussion ideas with others, reflecting on how individual pieces of information relate to larger constructs or patterns, and applying knowledge in real world situations (Biggs, 1989). Also characteristic of deep learning is integrating and synthesizing information with prior learning in ways that become part of one’s thinking and approaching new phenomena and efforts to see things from different perspectives (Ramsden, 2003; Tagg, 2003). As Tagg (2003) put it, “Deep learning is learning that takes root in our apparatus of understanding, in the embedded meanings that define us and that we use to define the world” (p. 70).

Surface and deep approaches to learning are not unalterable behaviors, though they may be influenced by personal characteristics such as ability (Biggs, 1987). But using one or the other approach is also affected in part by the learning task itself and the conditions under which the task is performed (Biggs, 1987; Ramsden, 2003). Thus, students may use both surface and deep approaches at different points in their studies. Although students may adopt different approaches in different situations, the general tendency is to adopt a particular approach and stick with it (Biggs, 1987; Entwistle, 1981; Ramsden, 2003).

The reason deep learning is important is because students who use such an approach tend to earn higher grades, and retain, integrate and transfer information at higher rates (Biggs 1988,

1989; Entwistle & Ramsden, 1983; Prosser & Millar, 1989; Ramsden, 2003; Van Rossum & Schenk, 1984; Whelan, 1988). Additionally, deep learning is associated with an enjoyable learning experience while the surface approach tends to be less satisfying (Tagg, 2003).

Existing Measures of Deep Learning

There are several measures that assess aspects of deep learning. The broader goal of these instruments is to assess differences in how students study and learn, which they accomplish in different ways. The measures have similar features, with later instruments building on the research base of the earlier ones (Entwistle & McCune, 2004).

Two widely used assessments of deep learning are Bigg's Study Process Questionnaire (SPQ) and Entwistle and Ramsden's Approaches to Study Inventory (ASI) (Biggs, 1987, Ramsden & Entwistle, 1981, Entwistle & Ramsden, 1983). Both inventories were designed for use in higher education (Entwistle & McCune, 2004) and have been revised in recent years to update wording, reduce items, and incorporate new research on learning (Biggs, Kember & Leung, 2001; Gibbs, Habeshaw, & Habeshaw, 1989; Entwistle & Tait, 1994). The SPQ consists of 42 items, with three "main approach" scales (deep, surface, and achieving) and six sub-scales that divide the core scales into motives and strategies. SPQ scores are indicators of the preferred, ongoing, and contextual approaches to learning (Biggs, Kember & Leung, 2001). SPQ items address higher-order learning (e.g., "While I am studying, I often think of real life situations to which the material that I am learning would be useful"), integration (e.g., "I try to relate what I have learned in one subject to that in another") and reflection (e.g., "In reading new material I often find that I'm continually reminded of material I already know and see the latter in a new light"). Conceptually similar to the SPQ, the ASI contains 64 items and 16 subscales that contribute to three main factors: reproducing orientation, meaning orientation, and achieving

orientation (Entwistle & McCune, 2004). In the ASI combinations of subscale scores are described as orientations to studying.

The ASI and SPQ have many similarities in how they describe the different ways students approach their academic work. For example, the SPQ's "deep approach" and the ASI's "meaning orientation" address behaviors associated with deep learning. As expected, there is empirical evidence of conceptual overlap of these instruments (Entwistle & McCune, 2004). The common elements tying the instruments together are the two distinctive types of learning processes (deep vs. surface), each with distinctive intentions and motives (Entwistle & McCune, 2004).

While the ASI and SPQ emphasize the relative stability of study strategies, the Motivated Strategies for Learning Questionnaire (MSLQ) addresses the impact of student's perceptions of their teaching-learning environments and how they accordingly may adapt their learning processes. The MSLQ emphasizes the role of self-conscious reflection on studying, drawing on the ideas of metacognition and self-regulation (Entwistle & McCune, 2004).

The MSLQ is designed to assess college students' motivational orientations and their use of different learning strategies for a college course (Pintrich, Smith, Garcia & McKeachie, 1993). The MSLQ taps into 3 broad motivational constructs: expectancy, value and effort (Pintrich et al., 1993). Expectancy addresses the student's beliefs about whether they can perform a task (Pintrich et al., 1993). Value addresses why students engage in particular academic tasks (Pintrich et al., 1993). Anxiety taps into students' worry and concern over taking exams (Pintrich et al., 1993). The MSLQ also has nine scales that address learning strategies. These nine scales fall into 3 general areas: cognitive (use of basic and complex strategies for processing information), metacognitive (how a student controls and regulates their own

cognition), and resource management (how student control other resources besides their own cognition) (Pintrich et al., 1993).

Deep Learning and NSSE

NSSE is an annual survey of first-year and senior college students at four-year institutions that measures students' participation in educational experiences that prior research has connected to valued outcomes (Chickering & Gamson, 1987; Kuh, 2001, 2003; Pascarella & Terenzini, 2005). The survey focuses on student participation in effective educational practices. For example, students are asked to identify how often they make class presentations, participate in a community-based project as a part of a course, and work with faculty members on activities other than coursework. In addition, students identify the degree to which their courses emphasize different mental processes (e.g., memorizing, evaluating, synthesizing), how many hours per week they spend studying, working, or participating in co-curricular activities, as well as how they would characterize their relationships with people on campus.

Understandably, several NSSE items tap behaviors indicative of deep approaches to learning. Although this was not the explicit intent of these items, a review of deep learning research and existing deep learning measures lends support to the use of NSSE items to create a deep learning scale. For example, four items (e.g., Applying theories or concepts to practical problems or in new situations) assess higher-order or cognitively intense learning that requires much more academic effort than simply memorizing facts for an exam. Other items (e.g., Examined the strengths and weaknesses of your own views on a topic or issue) probe the extent to which students reflect on the learning process. A final set of items (e.g., Worked on a paper or project that required integrating ideas or information from various sources) addresses the extent to which students are able to integrate and use information obtained from various sources.

Each of these three general areas is either captured by an existing measure of deep learning or otherwise represented in deep learning research.

Purpose of the Study

The main purpose of this study is to examine the factor structure underlying the items on NSSE identified as tapping deep approaches to learning. To accomplish this, we use an exploratory approach with data from 2004 and a confirmatory approach using data from 2005. The goal of the exploratory analysis is to identify factors, assess the reliability of those factors, and determine the relationships between the factors. Based on the results from 2004, we propose a model identifying the factor structure and, using data from 2005, we test this model as well as several others to determine if our proposed model fits the data better than the alternative models.

Methods

Data Source

The data for this study come from NSSE, an annual survey of college students at four-year institutions that measures students' participation in educational experiences that prior research has connected to valued outcomes (Chickering & Gamson, 1987; Kuh, 2001, 2003; Pascarella & Terenzini, 2005). The survey is available at the NSSE website, www.nsse.iub.edu. The standard NSSE sampling scheme draws equal numbers of first-year and senior students, with the size determined by the number of undergraduate students enrolled at the institution.

Each year, NSSE tests new survey items. In 2004, based on growing interest in deep learning, a set of items about reflective learning were included at the end of the online NSSE survey to augment core survey questions about higher order learning and integrative learning. Students who completed the paper version of the survey did not receive the reflective learning

items. The 2004 data were used for the exploratory factor analysis of the NSSE higher-order learning, integrative learning, and reflective learning items (i.e., the deep learning items).

Based on a favorable exploratory factor analysis and acceptable statistical properties of the reflective items, three of the experimental reflective items were added to the core NSSE survey for the 2005 administration. All students received the deep learning items in 2005, regardless of administration mode. The 2005 data were used for the confirmatory factor analysis.

Samples

Two different samples are used in this study. The first, from the 2004 administration of NSSE, consists of 110,886 randomly selected first-year and senior students from 450 U.S. four-year colleges and universities. The second, from the 2005 administration of NSSE, consists of 41,966 first-year students and seniors from 519 U.S. four-year colleges and universities. This sample represents one-fifth of the randomly selected respondents in 2005. It was necessary to reduce the sample to roughly this size in order for the confirmatory analyses to run. The 41,966 students were randomly selected from the total of 209,834 respondents.

Out of the 2004 sample, approximately 63% were female, 81% were white (5% African American, 5% Asian, 3% Hispanic, 1% Native American, < 1% other racial/ethnic background, and 5% multi-racial or ethnic), and 53% were first-year students. In addition, 30% were first generation college students, 52% lived on or near campus, about 12% were members of a social fraternity or sorority, and 93% are full-time students.

Similarly, of the respondents in the sample from 2005, approximately 65% were female and 51% were first-year students. In addition, 31% were first generation college students, about 11% were members of a social fraternity or sorority, and 91% are full-time students. A smaller percentage of the 2005 sample was white (72%) were white (7% African American, 5% Asian,

5% Hispanic, 1% Native American, 2% other racial/ethnic background, 2% multi-racial or ethnic, and 6% indicated that they preferred not to identify their race). The survey question about race was asked differently in 2005 compared to 2004, which likely accounts for most of the differences. Finally, a larger percentage of students (61%) lived on or near campus.

Some of the differences in respondent characteristics may also be attributable to the fact that all of the respondents in this study from 2004 completed the online version of the NSSE survey, since the reflective learning items were only administered online. Online completers differ in some ways from those students who fill out the paper survey. For example, a larger percentage of women and students of certain racial/ethnic groups (African American, Latino/a, and American Indian) fill out the paper version of the survey. Also, paper completers are more likely to be older, part-time, live off campus, have parents with less formal education, and have transferred from a different institution (Carini, Hayek, Kuh, Kennedy, & Ouimet, 2003). However, the slight differences in these characteristics should not present difficulties for the current study, as there is no evidence to suggest the relationships between the variables under study vary by these characteristics.

Measures

Three groups of items present on NSSE are identified in Table 1. The first two, higher order learning and integrative learning, were identified and used for investigations prior to 2004 (e.g., NSSE, 2003). The reflective learning items were developed for the 2004 administration of NSSE as an additional way to potentially tap students' use of deep approaches to learning.

The higher order learning items focus on the amount students believe that their courses emphasize advanced thinking skills such as analyzing the basic elements of an idea, experience, or theory and synthesizing ideas, information, or experiences into new, more complex

interpretations. The integrative learning items center around the amount students participate in activities that require integrating ideas from various sources, including diverse perspectives in their academic work, and discussing ideas with others outside of class. Central to the reflective learning behaviors is the notion that students can learn and expand their understanding by investigating their own thinking and then applying their new knowledge to their lives. The items ask, for example, how often students examined the strengths and weaknesses of their own views, learned something that changed their understanding, and applied what they learned in a course to their personal life of work.

Data Analyses

For the sample from 2004, exploratory factor analysis was used (Principle Components). Although, the items were in previously identified groups, we were interested in exploring the factor structure that would result from combining the three groups of items into an analysis. In particular, we were interested in examining whether the reflective learning items would be explained by a separate factor or whether they would load on factors associated with the other two groups. As with other factor analyses run on NSSE data (e.g., Kuh, 2001), an oblique rotation (Oblimin with Kaiser normalization) was used since resulting factors were assumed to correlate.

Building on the exploratory results, a proposed model of the factor structure was tested on the sample from 2005 using confirmatory factor analysis employing the EQS 6.1 statistical software program (Bentler, 1995; Bentler & Wu, 2002). Model results were compared to results for two other plausible models. The models and model specifications are addressed in the results.

To judge model fit in the confirmatory factor analyses, we relied on the guidelines given by Raykov, Tomer, and Nesslerode (1991) and reported the goodness of fit measures known as the Normed Fit Index (NFI), Non-normed fit index (NNFI, which is also known as the Tucker-Lewis Index or TLI), and the Comparative Fit Index (CFI). In addition, we followed Boomsma's (2000) recommendation to use the misfit index known as the Root Mean Square Error of Approximation (RMSEA). Under current standards, acceptable fitting models have indices that exceed .90 and RMSEA scores equal to or below .10. Models that fit well have indices greater than .95 and RMSEA scores equal to or below .05.

Results

Exploratory Factor Analysis

Table 2 presents the results of the exploratory factor analysis. Using the common criteria of retaining factors with eigenvalues greater than 1, three factors were derived (Eigenvalues for Factors 1, 2, and 3 were 6.15, 1.64, and 1.04, respectively). The three factors explain nearly 60% of the variance in the items. Not surprisingly, the items associated with each factor correspond to the *a priori* item groupings. The factor loadings are relatively strong for all items and there were no cross-loadings of a size comparable to primary factor loadings.

The factor correlations suggested that the underlying constructs of higher order learning, integrative learning, and reflective learning were moderately related (correlations ranging from .44 to .55). Scales created by taking the mean of all of the associated items for a particular grouping correlated somewhat higher ($r = .53$ between HL and IL, $r = .60$ between HL and RL, and $r = .48$ between IL and RL; $p < .001$ for all correlations; see Appendix A for all item and scale means, standard deviations, and correlations for the 2004 sample). The internal consistency of the item groupings was good for higher order learning and reflective learning ($\alpha =$

.82 and .89, respectively) and was acceptable for integrative learning ($\alpha = .71$). Item and scale means, standard deviations, and correlations are provided in Appendix A for the 2004 sample.

The clear factor loadings and relative strength of the correlations suggested the possibility of a second-order factor. A principle components analysis based on the three scales derived from the first analysis produced a single factor, further indicating this possibility. To test the resulting second-order factor model, data from the 2005 administration of NSSE was used along with confirmatory factor analysis.

Confirmatory Factor Analyses

Figure 3 identifies three potential factor structures for the items in Table 1 that appeared on the 2005 NSSE survey. These models are representations of models a through c described by Rindskopf and Rose (1988), simply adapted for the current investigation (i.e., the number of items loading on a factor matches the number of items in this study, not the actual models presented in their article). Note that in all of the models there are only three reflective learning items instead of six. In 2005, the three highest loading Reflective Learning items (RL1-3) were added to the core survey (only three could be added due to space constraints).

While the exploratory results yielded three factors (Table 2), the eigenvalue for Factor 1 was large compared to the eigenvalues for the other two factors. This is an indication that a single first order factor, in which all items load freely on the factor, may be sufficient to explain the relationships between the items (Single Factor Model in Figure 1). Since this is a possibility, such a model was tested.

Our proposed model, based on the exploratory results, is the Second-Order Factor Model in Figure 1. In this model, there are first-order factors representing higher-order learning, integrative learning, and reflective learning, and a second-order factor representing deep learning

upon which the first-order factors load. With only three first-order factors, such a model is just identified (Rindskopf & Rose, 1988). So, to run this model, a constraint was added to help insure identification. Based on data from 2004 and 2005, the variances of higher order learning and reflective learning were close in size. Consequently, as was done by Byrne (1994) in an example, the variances of the disturbances associated with higher-order learning and reflective learning were forced to be equal.

A single second-order factor is a special case of the Correlated Three Factor Model in that the second-order factor simply specifies a structure to the correlations between the first-order factors (Rindskopf & Rose, 1988). The Correlated Three Factor Model is tested to determine whether it fits the data better than the model with the additional structure (Second-Order Factor). Such a result would suggest that a single second-order factor is not the ideal structure indicated in the data.

Rindskopf and Rose (1988) provide a fourth model, d, where a general factor, one upon which all items load, is added to the Correlated Three Factor Model. As they point out, these “bi-factor” models can run into identification problems. When they tested a bi-factor model, they needed additional constraints in order for it to be identified. We experienced similar problems but were unable to identify the model even with the addition of constraints.

Table 3 contains the model fit statistics for the Single Factor, Second-Order Factor, and Correlated Three Factor models. The single factor solution does not fit the data well (fit indices $< .90$ and $RMSEA > .10$), where as the other two models fit very well with nearly identical fit statistics (fit indices $> .95$ and $RMSEA = .05$). The Second-Order Factor Model is the preferred solution since it is more parsimonious than the Correlated Three Factor Model.

Table 4 presents the standardized factor loadings and item reliabilities (R^2) for the second-order factor solution (note the standardized factor loading, reliability, and correlation estimates for the Correlated Three Factor Model are given in Appendix C for comparison purposes; Appendix B contains item and scale means, standard deviations, and correlations for the 2005 sample). The items load on the first-order factors in a similar fashion to the exploratory results (see Table 2). The first-order factors load highly (factor loadings of .70, .99, and .90) on the second-order factor. The factor loading for integrative learning on the second-order factor is very close to 1, indicating that integrative learning is nearly perfectly predicted by the second-order factor. This may indicate a particularly strong connection between integrative learning and deep learning or may be an artifact of the constraints placed on the disturbances of the other two factors.

Discussion and Implications

The purpose of this study was to improve our understanding of the structure underlying the relationships between the items we investigated. Our results suggest that these 12 NSSE items are appropriately grouped into three distinct categories, which we have labeled higher-order learning, integrative learning, and reflective learning. Further, a second-order construct, which we call, deep approaches to learning, is a good representation of the relationships between the three correlated first-order constructs. The fact that we utilized two years of data and both exploratory and confirmatory factor analyses adds to our confidence in the results. Consequently, those using NSSE data can create a deep approaches to learning scale, which is a combination of the three “sub”-scales. Our results suggest that the deep approaches to learning scale and its sub-scales have adequate to good internal consistency (see Table 1) and our

comparison between NSSE deep learning items and other measures of deep learning suggests that there is sufficient substantive overlap.

There is room for improvement in the scales' internal consistencies and our understanding of the validity of the scale scores. In particular, the integrative learning scale could be improved through the testing and use of additional items and the reflective learning scale would be improved by making available the three items from 2004 that were cut for the 2005 administration. In this vein, there is probably room for improvement in the higher-order learning items as well. Future work in this area may include the development of an "enhanced" deep learning scale, which would require the administration of extra items, given that it is unlikely that room for additional items will appear on the core NSSE questionnaire in the next couple of years.

In addition, several studies could dramatically improve our understanding of the validity of the deep learning scale and sub-scale scores. It seems imperative that researchers seek an opportunity to administer the NSSE items along with other measures of deep learning (e.g., SPQ, ASI, and MSLQ). Examining the correlations between these measures and, perhaps, using multi-trait, multi-method analyses to further examine scale reliability and validity would be valuable.

It seems unlikely that the NSSE items will be able to function as a replacement for other, more in-depth measures of deep learning processes. However, the small number of items makes this scale a quick way to assess students' use of deep approaches to learning. In addition, the fact that the scales are contained on a questionnaire used by over 500 colleges and universities each year, make them an appealing way to assess students' use of deep approaches to learning at

the campus level, rather than a more focused setting (e.g., a classroom) where the other instruments are more likely to be used.

In our work to date, we created the deep learning scale and sub-scales by taking the mean of the component sub-scales or items. This simplistic approach is preferred over other approaches (e.g., using factor loadings), because it is a common approach and one that can be replicated easily by the institutional users of NSSE. In Appendices A and B, the correlations among the scales mirror those found in the factor analyses.

Conclusion

NSSE annually assesses a number of educational activities that have been linked to positive outcomes, but until recently did not explicitly target deep learning. To address prevailing research that deep processing is an important component of a learning environment, NSSE first piloted a set of reflective learning items in 2004. The results of this study suggest that these items, when combined with existing core survey items, assess three distinct aspects of a second-order factor that, in content, appears to be related to deep learning.

Although the NSSE deep learning scale was not an intentional measure, it nonetheless has promising properties. The next step is to conduct research to examine the link between the NSSE scale and established measures of deep learning. Even in the absence of this important work the results of this study suggest that the NSSE deep learning scale reliably addresses several aspects important to an active, learner-centered educational environment.

References

- Beatie, V., Collins, B., & McInnes, B. (1997). deep and surface learning: A simple or simplistic dichotomy? *Accounting Education*, 6(1), 1-12.
- Bentler, P.M. (1995). *EQS 6: Structural Equation Program Manual*. Encino, CA: Multivariate Software.
- Bentler, P. M. & Wu, E. J. C. (2002). *EQS 6 for Windows User's Guide*. Encino, CA: Multivariate Software, Inc.
- Biggs, J.B. (1987). *Student approaches to learning and studying*. Hawthorn, Victoria: Australian Council for Educational Research.
- Biggs, J.B. (1988). *Approaches to learning and to essay writing*. Buckingham: Open University Press. In R.R. Schmeck (ed.) *Learning Strategies and Learning Styles*. New York, NY: Plenum.
- Biggs, J.B. (1989). Approaches to the enhancement of tertiary teaching. *Higher Education Research and Development*, 8, 7-25.
- Biggs, J.B. (2003). *Teaching for quality learning at university*. Buckingham: Open University Press.
- Biggs, J.B., Kember, D., & Leung, D.Y.P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71, 133-149.
- Boomsma, A. (2000). Reporting analysis of covariance structures. *Structural Equation Modeling*, 7(3), 461-482.
- Bowden, J., & Marton, F. (1998). *The university of learning*. London, England: Kogan Page.
- Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Thousand Oaks, CA: Sage.
- Carini, R.M., Hayek, J.H., Kuh, G.D., Kennedy, J.M., & Ouimet, J.A. (2003). College student responses to web and paper surveys: Does mode matter? *Research in Higher Education*, 44, 1-19.
- Chickering, A. W. & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 39(7), 3-7.
- Entwistle, N.J (1981). *Styles of learning and teaching: An integrated outline of educational psychology for students, teachers and lecturers*. Chichester: Wiley.

- Entwistle, N.J. & McCune, V. (2004). The conceptual bases of study strategy inventories. *Educational Psychology Review*, 16(4), 325-345.
- Entwistle, N.J., & Ramsden, P. (1983). *Understanding student learning*. London: Croom Helm.
- Entwistle, N.J., & Tait, H. (1994). *The revised Approaches to Study Inventory*. Edinburgh: Centre for Research into Learning and Instruction, University of Edinburgh..
- Gibbs, G., Habeshaw, S., & Habeshaw, T. (1989). *53 interesting ways to appraise your teaching*. Bristol: Technical and Educational Services.
- Kuh, G. D. (2001). Assessing What Really Matters to Student Learning: Inside the National Survey of Student Engagement. *Change*, 33(3), 10-17, 66.
- Kuh, G. D. (2003). What we're learning about student engagement from NSSE. *Change*, 35(2), 24-32.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning I: Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.
- National Survey of Student Engagement (2000). *The NSSE 2000 report: National benchmarks of effective educational practice*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Survey of Student Engagement (2001). *Improving the college experience: National benchmarks of effective educational practice*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Survey of Student Engagement (2002). *From promise to progress: How colleges and universities are using student engagement results to improve collegiate quality*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Survey of Student Engagement (2003). *Converting data into action: Expanding the boundaries of institutional improvement*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Survey of Student Engagement (2004). *Student Engagement: Pathways to collegiate success*. Bloomington, IN: Indiana University Center for Postsecondary Research.
- National Survey of Student Engagement (2005). *Exploring Different Dimensions of Student Engagement*. Bloomington, IN: Indiana University Center for Postsecondary Research.

- Pascarella, E. T. & Terenzini, P. T. (2005). *How college affects students: A third decade of research*. San Francisco: Jossey-Bass.
- Pintrich, P.R., Smith, D.A., Garcia, T., & McKeachie, W.J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychology Measurement*, 53, 801-813.
- Prosser, M., & Millar, R. (1989). The “how” and “why” of learning physics. *European Journal of Psychology of Education*, 4, 513-528.
- Ramsden, P. (2003). *Learning to teach in higher education*. London: RoutledgeFalmer.
- Raykov, T., Tomer, A. & Nesselroade, J.R. (1991). Reporting structural equation modeling results in Psychology and Aging: Some proposed guidelines. *Psychology and Aging*, 6(4), 499-503.
- Rindskopf, D. & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, 23, 51-67.
- Tagg, J. (2003). *The learning paradigm college*. Boston, MA: Anker.
- Van Rossum, E.J., & Schenk, S.M. (1984). The relationship between learning conception, study strategy and learning outcome. *British Journal of Educational Psychology*, 54, 73-83.
- Whelan, G. (1988). *Improving medical students' clinical problem-solving*. In P. Ramsden (ed.) *Improving learning: New perspectives*. London, England: Korgan Page.

Appendix A
 NSSE 2004 Deep Learning Means, Standard Deviations, and Correlations

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19			
1. DL	2.86	0.53	1.00																					
2. HL	3.06	0.66	0.81	1.00																				
3. IL	2.67	0.58	0.84	0.53	1.00																			
4. RL	2.86	0.68	0.84	0.48	0.60	1.00																		
5. HL1	3.22	0.75	0.64	0.79	0.41	0.38	1.00																	
6. HL2	2.97	0.84	0.69	0.83	0.47	0.41	0.59	1.00																
7. HL3	2.91	0.87	0.65	0.80	0.43	0.38	0.47	0.56	1.00															
8. HL4	3.14	0.84	0.62	0.79	0.39	0.36	0.50	0.51	0.51	1.00														
9. IL1	3.17	0.78	0.56	0.39	0.68	0.36	0.31	0.35	0.32	0.29	1.00													
10. IL2	2.77	0.90	0.56	0.34	0.69	0.40	0.26	0.30	0.30	0.23	0.42	1.00												
11. IL3	2.66	0.84	0.59	0.39	0.69	0.41	0.30	0.34	0.30	0.31	0.37	0.31	1.00											
12. IL4	1.94	0.84	0.55	0.34	0.67	0.38	0.25	0.31	0.28	0.25	0.29	0.28	0.34	1.00										
13. IL5	2.83	0.86	0.60	0.35	0.69	0.47	0.29	0.30	0.27	0.26	0.29	0.30	0.36	0.38	1.00									
14. RL1	2.84	0.83	0.69	0.38	0.49	0.84	0.30	0.32	0.32	0.27	0.29	0.35	0.33	0.32	0.39	1.00								
15. RL2	2.91	0.82	0.66	0.35	0.46	0.81	0.28	0.30	0.29	0.25	0.28	0.35	0.30	0.28	0.36	0.69	1.00							
16. RL3	2.84	0.82	0.68	0.38	0.47	0.83	0.30	0.32	0.31	0.29	0.29	0.32	0.32	0.30	0.37	0.63	0.65	1.00						
17. RL4	2.72	0.86	0.66	0.37	0.48	0.80	0.30	0.32	0.30	0.26	0.28	0.34	0.31	0.31	0.37	0.68	0.57	0.59	1.00					
18. RL5	2.92	0.85	0.69	0.41	0.50	0.79	0.31	0.34	0.33	0.35	0.31	0.31	0.37	0.32	0.40	0.54	0.53	0.62	0.50	1.00				
19. RL6	2.92	0.86	0.66	0.41	0.48	0.76	0.33	0.35	0.30	0.33	0.30	0.27	0.35	0.33	0.38	0.52	0.50	0.54	0.49	0.60	1.00			

N=1110,886; DL = Deep Learning; HL = Higher Order Learning; IL = Integrative Learning; RL = Reflective Learning; p < .001 for all correlations

Appendix B
 NSSE 2005 Deep Learning Means, Standard Deviations, and Correlations

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. DL	2.83	0.53	1.00															
2. HL	3.02	0.67	0.80	1.00														
3. IL	2.70	0.58	0.83	0.54	1.00													
4. RL	2.77	0.71	0.80	0.40	0.51	1.00												
5. HL1	3.15	0.77	0.64	0.80	0.43	0.32	1.00											
6. HL2	2.94	0.84	0.68	0.84	0.47	0.34	0.62	1.00										
7. HL3	2.90	0.86	0.65	0.81	0.44	0.33	0.50	0.57	1.00									
8. HL4	3.08	0.84	0.62	0.80	0.42	0.29	0.52	0.53	0.53	1.00								
9. IL1	3.20	0.77	0.53	0.38	0.67	0.28	0.30	0.33	0.31	0.29	1.00							
10. IL2	2.79	0.88	0.56	0.35	0.68	0.36	0.27	0.31	0.31	0.25	0.42	1.00						
11. IL3	2.74	0.82	0.60	0.43	0.71	0.36	0.34	0.37	0.34	0.35	0.37	0.34	1.00					
12. IL4	1.99	0.90	0.54	0.35	0.68	0.33	0.26	0.31	0.29	0.27	0.28	0.28	0.37	1.00				
13. IL5	2.79	0.85	0.60	0.37	0.68	0.43	0.31	0.31	0.28	0.28	0.28	0.30	0.38	0.37	1.00			
14. RL1	2.66	0.87	0.68	0.32	0.43	0.85	0.27	0.28	0.27	0.23	0.22	0.29	0.30	0.29	0.37	1.00		
15. RL2	2.81	0.83	0.68	0.32	0.43	0.87	0.26	0.28	0.27	0.23	0.23	0.32	0.30	0.26	0.36	0.64	1.00	
16. RL3	2.84	0.82	0.69	0.37	0.45	0.82	0.29	0.31	0.30	0.28	0.26	0.30	0.32	0.29	0.37	0.52	0.58	1.00

N=44,558; DL = Deep Learning; HL = Higher Order Learning; IL = Integrative Learning; RL = Reflective Learning; p < .001 for all correlations

Appendix C
Standardized Factor Loading, Reliability, and Correlation Estimates for
the Correlated Three Factor Model

Factors and Items	HL	IL	RL	R ²
<i>Factor Loadings</i>				
Higher Order Learning				
HL1	.75			.56
HL2	.80			.65
HL3	.72			.51
HL4	.69			.48
Integrative Learning				
IL1		.56		.31
IL2		.56		.32
IL3		.64		.41
IL4		.55		.30
IL5		.60		.36
Reflective Learning				
RL1			.76	.58
RL2			.82	.66
RL3			.71	.51
<i>Factor Correlations</i>				
	HL	1.00		
	IL	.71	1.00	
	RL	.47	.66	1.00

Note: R² refers to the amount of variance accounted for in a factor by a particular indicator and is a reliability estimate. HL = Higher Order Learning; IL = Integrative Learning; RL = Reflective Learning

Table 1.
Deep Learning Scale, Subscales, and Component Items

Deep Learning ($\alpha_{2004} = .77, \alpha_{2005} = .73$)

Combination of the 3 subscales listed below

Higher-Order Learning^a ($\alpha_{2004} = .82, \alpha_{2005} = .82$)

- HL1 Analyzed the basic elements of an idea, experience, or theory, such as examining a particular case or situation in depth and considering its components
- HL2 Synthesized and organized ideas, information, or experiences into new, more complex interpretations and relationships
- HL3 Made judgments about the value of information, arguments, or methods, such as examining how others gathered and interpreted data and assessing the soundness of their conclusions
- HL4 Applied theories or concepts to practical problems or in new situations

Integrative Learning^b ($\alpha_{2004} = .71, \alpha_{2005} = .71$)

- IL1 Worked on a paper or project that required integrating ideas or information from various sources
- IL2 Included diverse perspectives (different races, religions, genders, political beliefs, etc.) in class discussions or writing assignments
- IL3 Put together ideas or concepts from different courses when completing assignments or during class discussions
- IL4 Discussed ideas from your readings or classes with faculty members outside of class
- IL5 Discussed ideas from your readings or classes with others outside of class (students, family members, co-workers, etc.)

Reflective Learning^{b,c} ($\alpha_{2004} = .89, \alpha_{2005} = .81^d$)

- RL1 Examined the strengths and weaknesses of your own views on a topic or issue^e
- RL2 Tried to better understand someone else's views by imagining how an issue looks from his or her perspective^e
- RL3 Learned something that changed the way you understand an issue or concept^e
- RL4 Learned something from discussing questions that have no clear answers
- RL5 Applied what you learned in a course to your personal life or work
- RL6 Enjoyed completing a task that required a lot of thinking and mental effort

^a Component items measure on a 4-point scale (1=Very little, 2=Some, 3=Quite a bit, 4=Very much).

^b Component items measured on a 4-point scale (1=Never, 2=Sometimes, 3=Often, 4=Very often)

^c Component items were additional items asked of online responders during the 2004 NSSE administration.

^d In 2005, the reflective learning sub-scale consisted of only three items, the first three items listed.

^e Item added to the core survey for the 2005 NSSE administration. Other items were not retained.

Table 2.
NSSE 2004 Deep Learning Exploratory Factor Analysis

Items	Factor 1	Factor 2	Factor 3
<i>Structure (Pattern) Matrix</i>			
Higher Order Learning			
HL1		.80 (.80)	
HL2		.83 (.79)	
HL3		.78 (.76)	
HL4		.79 (.81)	
Integrative Learning			
IL1			.72 (.75)
IL2			.69 (.70)
IL3			.69 (.66)
IL4			.66 (.66)
IL5			.64 (.53)
Reflective Learning			
RL1	.85 (.86)		
RL2	.83 (.85)		
RL3	.84 (.86)		
RL4	.80 (.79)		
RL5	.77 (.70)		
RL6	.73 (.65)		
Percent Variance Explained	40.99	10.95	6.96
<i>Component Correlations</i>			
Factor 1	1.00		
Factor 2	.44	1.00	
Factor 3	.55	.50	1.00

N = 110,886

Note: Principal Components Analysis with Oblimin rotation with Kaiser normalization; cross-loadings were all less than .50.

Table 3.
Summary of Fit Indices for Confirmatory Factor Analysis Models

Model	χ^2	<i>df</i>	NFI	NNFI	CFI	RMSEA
Single Factor	52,596.72	55	.71	.65	.71	.14
Second-Order Factor	5,186.28	52	.97	.96	.97	.05
Correlated Three Factor	5,164.58	51	.97	.96	.97	.05

N = 41,966

Note: NFI = normed fit index; NNFI = non-normed fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation. All χ^2 probabilities < .001.

Table 4.
Standardized Factor Loading and Reliability Estimates for Second-Order Factor Model

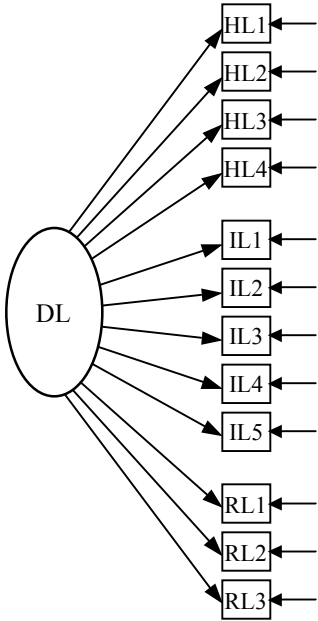
Factors and Items	Factor Loadings				R ²
	HL	IL	RL	DL	
Higher Order Learning				.70	.49
HL1	.80				.56
HL2	.84				.65
HL3	.77				.51
HL4	.75				.49
Integrative Learning				.99	.99
IL1		.64			.31
IL2		.65			.32
IL3		.72			.41
IL4		.64			.30
IL5		.69			.36
Reflective Learning				.91	.45
RL1			.84		.58
RL2			.87		.66
RL3			.89		.50

Note: R² refers to the amount of variance accounted for in a factor by a particular indicator and is a reliability estimate. DL = Deep Learning; HL = Higher Order Learning; IL = Integrative Learning; RL = Reflective Learning

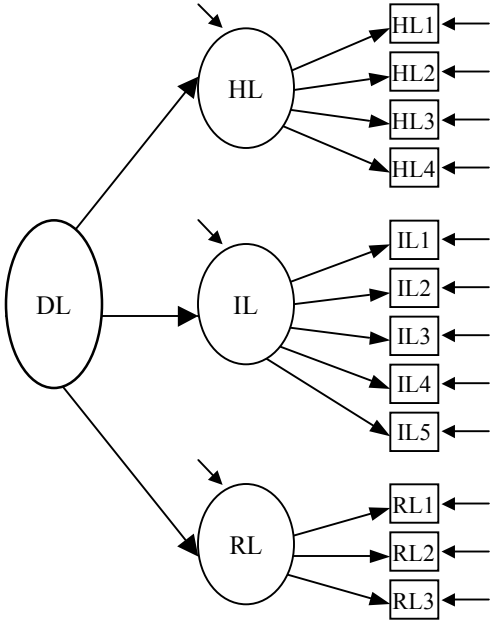
Figure Captions

Figure 1. Confirmatory Factor Analysis Models

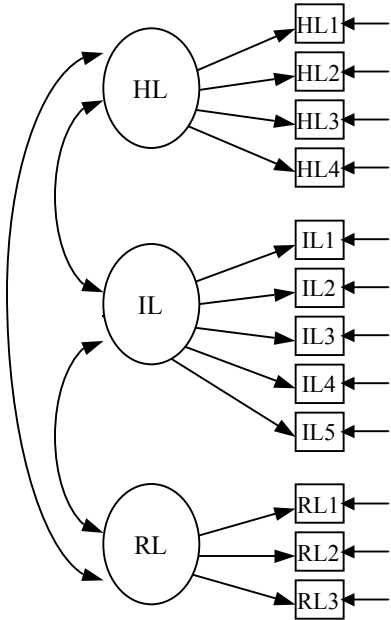
Single Factor



Second-Order Factor



Correlated Three Factor



Note: Circles represent factors, squares represent observed variables. DL = Deep Learning; HL = Higher Order Learning; IL = Integrative Learning; RL = Reflective Learning.