Deep learning as an individual, conditional, and contextual influence on first-year student outcomes

Robert D. Reason
Penn State University

Bradley E. Cox
Florida State University

Kadian McIntosh and Patrick T. Terenzini
Penn State University

A paper presented at the Annual Forum of the Association for Institutional Research, Chicago, IL. May 31, 2010

Abstract

For years, educators have drawn a distinction between deep cognitive processing and surface-level cognitive processing, with the former resultin in greater learning. In recent years, researchers at NSSE have created DEEP Learning scales, which consist of items related to students' experiences which are believed to encourage deep processing. In this study we examine the predictive power of NSSE's DEEP Learning scale as it relates to four learning outcomes. Further, we explore whether a contextual effect is present. That is, we explore if students receive a benefit simply from being a part of a deeply engaged community.

Deep learning as an individual, conditional, and contextual influence on first-year student outcomes

Deep cognitive processing, studying material with a focus on learning its significance and meaning (Marton & Saljo, 1976), has been shown to result in greater learning outcomes in college students (Nelson Laird, Shoup, Kuh, & Schwarz, 2008; Ramsden, 2003). Finding ways to encourage deep processing among college students has become an essential pursuit of higher educators, who believe that engagement in certain activities and environments likely encourage students to pursue deep processing. This belief, however, is not well tested.

In this study, we explore whether creating situations in which students engage in activities believed to be related to deep processing results in increased learning. In so doing, we examine the predictive power of NSSE's DEEP Learning scale. We thus conduct analyses seeking answers to two research questions:

> Does engagement in activities related to deep approaches to learning, as operationalized by NSSE's DEEP Learning Scale, result in more positive student learning outcomes?
>
> Is there a contextual effect from the deep learning approaches?

To answer these questions, we use three different measures of students' self-reported gains as well as an objective measure of students' critical thinking ability at the end of the first year. This research thus attempts to confirm and extend the findings of Nelson Laird and his colleagues (2008) who found that students' scores on the DEEP Learning scale were predictive of self-reported learning.

Deep Processing as an Individual Influence

In 1976, Marton and Saljo described two distinct processes by which college students read a particular passage of text. Students who employed "surface-level processing" focused on “rote-learning” (p. 7) and memorization of the text. In contrast, those who employed “deep-level” processing sought to understand the text’s purpose, meaning, and significance. In the decades since, the terms have been refined, with Entwistle (2000) neatly articulating the concepts:

> In the deep approach, the intention to extract meaning produces active learning processes that involve relating ideas and looking for patterns and principles on the one hand (a holist strategy - Pask, 1976, 1988), and using evidence and examining the logic of the argument on the other (serialist). The approach also involves monitoring the development of one’s own understanding (Entwistle, McCune & Walker, 2000). In the surface approach, in contrast, the intention is just to cope with the task, which sees the course as unrelated bits of information which leads to much more restricted learning processes, in particular to routine memorization. (n.p.)

The use of deep processing is not restricted to a particular subject matter, context, or student. Rather, deep approaches to learning have been studied within a variety of contexts, including traditional academic courses (Elias, 2005; Holschuh, 2000), computer-based classes (Aharony, 2006; Greene, 2003), and work environments (Geertshuis, 2006; Kirby, 2003). Deep cognitive processing occurs in all subject areas, including engineering (Case, 2004), science (Beckwith, 1991; Chin, 2000), reading (Marton & Saljo; 1976), and business (Elias, 2005). Nor is deep processing restricted to particular people; the phenomenon has been studied with pre-college students (Aharony, 2006; Chin, 2000) and adults (Geertshuis, 2006; Kirby, 2003), but most commonly with college students (e.g.,Beckwith, 1991; Case, 2004; Elias, 2005; Marton & Saljo, 1976; Watkins, 1983). Although not uniform in their findings (see, for example, Beckwith,

(1991) and Murphy & Alexander (2002), studies of deep processing have found consistently that students who used deep (as opposed to “surface”) approaches demonstrated improved learning outcomes.

Although educational psychologists typically focus on how specific learning strategies applied to specific subject matters in specific circumstances affect specific outcomes (Ramsden, 2003; Trigwell & Prosser, 1991), we extrapolate from these findings and the work of Biggs (1979), Entwistle (2000), and Marton and Saljo (1976) to argue that, because deep approaches to learning yield benefits in each specific instance (across subjects, contexts, and students), more frequent deep processing in general should be associated with improvements in generalized learning outcomes. On the whole and all other things being equal, those who have a tendency to use deep processing should achieve higher scores on measures of broad, generalized learning outcomes (e.g., critical thinking). We, therefore, hypothesize that students’ engagement with deep learning activities will be positively related to their end-of-first-year outcomes.

Deep Processing as a Conditional Influence

There is some reason to believe that the effect of engagement in deep processing on positive learning outcomes may be conditional. Kuh and his colleagues (Kuh, Cruce, Shoup, Kinzie, & Gonyea, 2008; Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006) found that engagement in educationally purposeful activities affected students with lower pre-college achievement more positively than students with higher pre-college achievement. Although their measure of educationally purposeful activities is not directly analogous to deep processing, it does include measures that were incorporated into the DEEP Learning scale. The authors concluded that “engagement appears to have a conditional, compensatory effect on grades (Kuh, et al., p. 35)," a conclusion that supports a hypothesis that deep processing may have a similar conditional and compensatory effect on outcomes.

Deep Cultures as Contextual Influences

Studies of individual learning processes note that students' information processes are influenced by a variety of contextual factors. Beattie and her colleagues (1997), in summarizing the research related to deep approaches to learning, concluded that early assumptions of the fixed nature of an individual's approach to learning have been replaced by an understanding that students can employ either deep or surface-level approaches as necessary based on the context of the task. Murphy and Alexander (2002) similarly concluded that students can employ *both* approaches intermittently, and that the choice of deep vs. surface approaches to learning is influenced by both the individual student and the context in which the learning takes place. A particular student might use primarily deep approaches with familiar material (e.g., within his/her major) but more surface-level strategies when completing coursework he/she finds unimportant or disinteresting. Likewise, students can be primed or prompted to use a specific approach by certain characteristics of the questions on a test (Biggs, 1979). Students are more likely to use surface-level strategies when asked to respond to true/false or multiple-choice questions on a test, but employ more deep approaches when preparing essays or analyzing a case-study on that same exam. In a study using measures similar to those in the current study, Nelson Laird et al (2008) found that, while the effects of deep learning activities were consistent for students across majors, the extent to which students used deep approaches to learning varied significantly by disciplinary fields.

In noting these disciplinary differences, Nelson Laird and colleagues (2008) reinforced the common understanding that students are influenced by the normative behaviors of those with whom they most frequently interact. The current study's conceptual framework (see Figure 1; Terenzini & Reason, 2005) likewise stipulates that students' experiences both affect and are affected by the experiences of their peers. Nearly three decades of research provides considerable evidence of the peer effect (Pascarella & Terenzini, 2005). Astin and Panos, as early as 1969, noted a tendency for students to move toward the dominant values and belief structures of other students, labeling this phenomenon "progressive conformity." Acknowledging the general tendency for students to progressively conform to the norms of their peer environment, and noting the extent to which contextual factors influence students' approaches to learning, we hypothesize that students will derive benefits from their mere presence within a student body that tends to engage in deep learning activities.

NSSE and Measures of Deep Learning

Typically, studies involving deep learning processes are smaller in scale and context-specific, which does not allow for large-scale analyses of the effects of deep processing on learning outcomes. Because measurement of students' internal thought processes would be impossible to do at the scale on which NSSE operates (more than 80,000 students at more than 500 schools annually), NSSE researchers created their DEEP Scales using reflective measures of experiential proxies to represent students' overall tendency to employ deep learning processes. The Nelson Laird study (2008), however, is perhaps the only large-scale, nation-wide analysis published thus far that explicitly links deep learning activities with college student outcomes.

NSSE's approach to measurement parallels one of education's most fundamental assumptions: although engaging in the type of effort required to learn for meaning and integrate new information requires a personal commitment on the part of the student, environmental and curricular interventions can be designed to encourage students to employ deep approaches to learning (English, Luckett, & Mladenovic, 2004; Hall, Ramsay, & Raven, 2004). Faculty members can engage students in group work which holds students accountable for their own learning and requires understanding of the subject matter in order to function in the group (Hall, et al., 2004). Institutions can also include first-year seminars that require reflection and integration of knowledge (English, et al., 2004), or plan for such integration through intentionally designed learning communities (Cole, Mccormick, & Kinzie, 2009). Incorporating assessments of learning that require students to demonstrate and apply understanding also motivates students to assume deep approaches to learning (Ramsden, 2003). Recent efforts by the Association of American Colleges and Universities (AAC&U) to identify “high-impact practices” that improve students’ educational attainment focus almost entirely on interventions meant to encourage deep approaches to learning: writing intensive courses, collaborative assignments, undergraduate research, service learning, and capstone courses and projects – all practices Kuh (2008) suggests encourage student to adopt deep approaches to learning.

Employing NSSE’s DEEP Learning scale, Nelson Laird and his colleagues (2008) studied self-reported student outcomes and found that students’ scores on the engagement with deep learning approaches vary significantly by the type/category of their major (e.g., hard v. soft). Even with the variance in engagement by discipline, the researchers found that, regardless of discipline, more frequent engagement with deep approaches resulted in higher levels of student self-reported learning. Accordingly these authors concluded that engagement in deep learning, as measured by the NSSE scale, resulted in student learning outcomes regardless of disciplinary context, although context certainly mediated the degree of engagement. However, when these authors tested the relationship between NSSE DEEP Learning scale and students GPA, presumably a more objective measure of academic performance, the relationship was weak. The current study extends this line of research by testing three distinct ways in which deep processing could influence any of four student outcomes, including an objective measure of students' critical thinking skills.

Research Methods

The study's purpose was to explore the relationship between engagement in activities believed to encourage deep cognitive processing and both self-reported and objectively-measured student learning outcomes for first-year students at 33 institutions. Based on the conceptual framework that guided the study, individual level variables were examined within the institutional context and the context of the students' environments. This study designed required that we use multilevel modeling techniques. The following section explains our research methods in greater detail.

Conceptual Framework

Data come from a larger study of the wide array of forces shaping first-year student outcomes. The conceptual framework for that study (see Figure 1) expands upon Astin's Inputs-Environment-Outputs approach (Astin, 1993) and Terenzini, Springer, Pascarella, and Nora's (P. Terenzini, Springer, Pascarella, & Nora, 1995a, 1995b) model of college effects on student outcomes. These conceptual frameworks hypothesize that students come to college with a range of demographic, personal, and academic background characteristics and experiences that shape students' engagement with various aspects of their institution. Those involvements are themselves influenced by a variety of curricular, classroom, and out-of-class experiences and conditions. The framework for the current study suggests that all of these dynamics occur within, and are mediated by, an often-overlooked fourth domain, the institutional context, comprising an institution’s internal organizational characteristics, structures, practices, and policies, as well as the campus’s faculty and peer cultures and environments (Terenzini & Reason, 2005).

[Insert Figure 1 about here]

Sample

This study uses data drawn from 5,905 first-year students at 33 institutions. This project drew student participants from a pool of all baccalaureate degree-seeking first-year students at one of the 33 participating institution during the Fall 2006 term. To be included in our final sample, students must have completed both the ACT college entrance exam and at least one of four assessments in the Spring 2007 term, near the end of their first college year: the National Survey of Student Engagement, the Critical Thinking and/or Writing Skills module(s) of the Collegiate Assessment of Academic Progress (CAAP), and a supplemental survey designed specifically for this study. Data from these assessments were augmented by information provided by participating institutions and drawn from students’ ACT college entrance exam. Collectively, data from these sources provide a multi-dimensional profile of participating students’ secondary school background, academic preparation, and pre-college activities, as well as their collegiate experiences, reported gains in their knowledge, skills, and personal development, and tested skill levels in critical thinking.

Although all participants and protocols met universal project guidelines, institutional variations in student recruitment procedures precludes any assumption that the final sample was drawn completely at random; such variation also prevents calculation of a conventional response rate. Nonetheless, estimates based on institutional data collection plans predicted that 9,877 students would participate. A total of 5,905 students (59.8% of those predicted) eventually provided information on one or more of the study’s instruments. Missing data were imputed using the estimation-maximization algorithm in SPSS 17, an iterative maximum-likelihood procedure that yields unbiased correlation and regression coefficients. Students' responses were weighted to be representative of the first-year student population at their institution with regard to race, gender, and ACT Composite score; weights also adjusted for differing response rates across institutions.

Measures

Independent variable of primary interest. The NSSE DEEP Learning scale was the independent variable of primary interest for this study (see Table 1 for details). Based on the efficacy of deep approaches to learning, researchers at NSSE operationalized four measures of deep approaches to learning (Nelson Laird, Shoup, & Kuh, 2005). According to Nelson Laird and his colleagues, the NSSE scales are grounded in the extensive literature of deep learning. NSSE created three subscales: *higher order learning*, *integrative learning*, and *reflective learning*. A second-order scale, *Deep approaches to learning* scale, is a combination of the three subscales and serves as an overall measure of deep approaches on a campus. It is this last scale that serves as our independent variable of primary interest.

The DEEP Learning scale score enters our analysis at both the individual level and institution level. At the individual level, a student's DEEP Learning scale score represents the student's level of engagement in activities meant to encourage deep processing. At the institutional level, an institution's aggregate DEEP Learning scale score (the average of all students' scores in the institution) is used as an environmental variable to assess the contextual effects of being part of a deeply engaged environment.

Control variables. Our models consist of control variables at both the individual and institutional level. At the individual level, four control variables were included: race (white/nonwhite), sex (male/female), high school rank (top quartile), and ACT Composite score. At the institutional level, we included institutional control (public/private), level (doctoral/other), and median ACT Composite score for the incoming first-year cohort during the year of our survey administration.

Outcome variables. This paper explored the effects of deep learning processes, as measured by the DEEP Learning scale score, on four different outcome variables: CAAP Critical Thinking Score as well as students' self-reported gains in general education, practical competence, and personal and social development. The CAAP Critical Thinking module is one of six standardized, nationally normed modules measuring student progress in the acquisition of core academic skills typically developed in the first two years of college (ACT, Inc. nd). The Critical Thinking module measures students' abilities to clarify, analyze, evaluate, and extend arguments, and is nationally normed to have an average score of 61.65, and a standard deviation of 5.39.

The other outcome variables were students' self-reported gains in three areas from the NSSE instrument. Students' self-reported gains in general education included items related to communication skills, breadth of knowledge, and critical/analytical thinking. The practical competence scale items related to on-the-job and problem-solving skills. The personal and social development scale involved items related to developing self-awareness, civic responsibility, and ethical/spiritual development. All items on these three scales asked students to rate the extent to which the institution contributed to the student's growth in the area under consideration. The scales and their component items are detailed in Table 1.

[Insert Table 1 about here]

Analytic Models

Our analyses involved the development of three sets of random-intercept hierarchical linear models (HLM or mixed models) predicting students' scores on each of the four outcome variables measured. All level-1 (student specific) variables were centered around the grand-mean for the entire sample; level-2 (institutional) variables are the weighted average of scores for students at a given institution. We first estimated a null model without any level-1 or level-2 variables. Beginning with a null model allowed an estimation of the influence of the institutional-level variables and the influence of the individual-level variables. This model also tests the assumption that at least some of the variance in the outcome measure is attributable to institutional differences (Raudenbush & Bryk, 2002).

For the second set of models, we added the block of level-1 variables, including control variables and the DEEP Learning scale. The resulting coefficient for the DEEP variable reflects the influence of deep learning activities *net* of key differences in students' pre-college characteristics. In addition to the level-1 main effects, our models include a variable interacting students' ACT score with their DEEP Learning scale score. We do so for two reasons. First, because preliminary results indicated that ACT scores and DEEP Learning scale scores had opposite effects on all three of the self-reported gain scales, we hoped the interaction term would add some clarity regarding this unexpected finding. Second, Kuh and his colleagues (2006, 2008) have reported that the benefits of student engagement vary based on students' pre-college preparation. Specifically, "student engagement in educationally purposeful activities had a small, compensatory effect on first-year GPA of students who entered college with lower levels of academic achievement" (2008, p. 549). Our interaction term allows us to examine whether DEEP learning activities have a compensatory effect on four different first-year student outcomes.

Finally, the third set of models built upon the completed level-1 models by adding all four level-2 variables simultaneously: the three control variables (i.e., median ACT Composite, public/private, and doctoral/other) and the aggregated DEEP Learning scale for each institution. Because all level-1 variables are centered around their grand mean, the coefficient for the level-2 DEEP variable can be interpreted as a direct measure of the “contextual” or “compositional” influence of attending an institution in which the student body, as a whole, engages in deep learning activities (Raudenbush & Bryk, 2002, p. 139).

The inclusion of students’ ACT score has important implications for the interpretation of our models. Because ACT scores were strongly correlated with students’ scores on the CAAP-CT ($r = .817$; $p < .01$), there remained a comparatively small amount of variance that could be explained by other variables in the models. Recognizing that the effects of DEEP activities might be mitigated by the inclusion of other college-experience variables we include the DEEP Learning scale as the *only* measure of students’ college experiences, thereby maximizing the possibility for the DEEP Learning scale to reach statistical significance.

Results

Our analyses involved the building of three models for each of four outcome variables. The null models simply partition the outcome measure’s variance among individuals (level 1) and institutions (level 2). As the “Null” column in Tables 2 indicates, roughly one third (32.0%; [9.081/(9.081+19.252)]) of the variance in critical thinking scores occurred between institutions (level 2). Institutional variance accounted for a much smaller percent of the overall variance for each of the three self-reported gain scales (Table 3): General Education (4.6%), Practical Competence (3.1%), and Personal and Social Development (6.2%). Nonetheless, the level 2 component of the variance was statistically significant for all four outcome measures, thereby indicating that a multi-level modeling approach was appropriate (Porter, 2005).

[Insert Tables 2 and 3 about here]

The remaining models were developed to assess the influence of DEEP learning activities in three ways: as an individual influence, as a conditional influence, and a contextual influence. Results related to each of these potential modes of influence are addressed separately.

DEEP Activities as Individual Influences

The effect of students' engagement in deep learning activities differed sharply between outcome measures. Like Nelson Laird and colleagues (2008), we found the DEEP Learning scales to be positively related to students' self-reported gains during the first year of college (see "Level 1" column in Table 3). The positive effects of deep learning activities remain even amid controls for students' demographic characteristics and pre-college academic performance. However, the DEEP Learning scale did *not* have a statistically significant relationship to students' scores on the CAAP Critical Thinking Module (Table 2).

DEEP Activities as Conditional Influences

In addition to the main DEEP variable, we included a term interacting students' DEEP Learning scale scores with their ACT Composite scores. This interaction term was not statistically significant in our models predicting students' CAAP-CT scores, or their self-reported gains in General Education or Personal and Social Development. However, it was a significant, and negative, predictor of students' reported gains in Practical Competence. That is, the positive influence of DEEP activities decreased as students' ACT Composite scores increased. Nonetheless, because the DEEP Learning scale and ACT Composite are in different metrics, and because their main effects have opposite signs for three of the four criterion measures, we created four graphs to further clarify the ACT-conditional nature of students' engagement with DEEP activities. In these graphs, we plot students' *predicted outcomes* against students' DEEP Learning scales, with different color lines differentiating between those with different ACT scores. Thus, in Figure 2, we depict the relationship between DEEP activities and critical thinking for students with ACT Composite scores of 16, 20, 24, 28, and 32.

[Insert Figure 2 about here]

As Figure 2 indicates, students who entered with higher ACT Composite scores also had higher scores on the Critical Thinking assessment – regardless of the level of engagement with deep learning activities. Moreover, the largely horizontal and parallel lines suggest that, for students with a given ACT Composite score, engagement in deep learning activities has no meaningful effect on their critical thinking skills at the end of the first year.

Figure 3 tells quite a different story. For all three self-reported gain scales, more frequent engagement in deep learning activities is associated with higher reported gains during the first year of college. Students with *higher* ACT Composite scores report *smaller* gains. Nonetheless, relative to Figure 1, the graphs in Figure 2 depict much less distinction between students with different ACT Composite scores. Despite vastly differing entering ACT Composite scores (of 20 and 28 in Figure 2), there is considerable overlap between students' reported gain scores. Moreover, and more specifically addressing the conditional effects of deep processing, all three graphs suggest that students with all levels of prior achievement receive roughly comparable benefits from engagement in deep learning activities. Even in terms of gains in practical competence, the one set of models in which the ACT by DEEP interaction effect was statistically significant, the differences in slopes is subtle and of questionable practical significance.

Thus, Figure 3 provides graphical confirmation of the findings presented in Table 3: there appears to be little, if any, compensatory conditional effect of deep experiences, and engagement in deep learning activities (i.e., the DEEP Learning scale), but not prior academic achievement (i.e., the ACT Composite), positively contributes to students' self-reported gains during their first year of college.

[Insert Figure 3 about here]

DEEP Activities as Contextual Influences

In our final set of analyses, we explored the possibility that students derived benefits simply from their attendance at institutions in which there was a student culture of engagement with deep learning activities. Such a culture, in contrast to more traditional institutional descriptors (e.g., Carnegie classification, selectivity, or public/private control), might be intentionally and (relatively) quickly cultivated by institutional policies and faculty practices. However, as shown in the bottom half of Tables 2 and 3, an institution's average DEEP Learning scale score has no significant relationship with any of the four outcomes measured. In contrast, an institution's median ACT Composite score was positively and significantly related to three of the four outcomes.

Discussion and Implications

Several of the findings from this study are surprising and, presumably, challenge long-held beliefs about engagement and learning. NSSE's operationalization of DEEP learning activities was related to self-reported outcomes (also from the NSSE instrument), but unrelated to the objective measure of critical thinking. Further, there were no contextual effects identified, meaning that students reaped no benefit from simply being in an "engaged" environment. This latter finding challenges pervious research and the conceptual framework upon which this study was based (Terenzini & Reason, 2005).

Deep Activities as Individual Influences

Individual student's scores on the DEEP Learning scale were significantly and positively related to student's self-reported gains in general education, practical competence, and personal and social development. This finding reinforces Nelson Laird and colleagues' finding (Nelson Laird, et al., 2008) and the belief that deeper learning results from greater engagement.

In contrast, we found no relationship between students' scores on the DEEP Learning scale and their scores on the CAAP Critical Thinking module. Of course, this finding could indicate that no such relationship exists—that DEEP learning activities, as measured by NSSE, are not related to students' critical thinking abilities at the end of their first-year of college. This interpretation calls into question the efforts of higher education professionals to influence students' critical thinking skills within the first year of college. Further the finding calls into question a deeply held assumption that higher education officials can influence such outcomes through environmental intervention, those interventions designed to promote students engagement in deep processing.

There are, however, others explanations for this challenging, and ostensibly counterintuitive, finding. Of particular note is the difference between the self-reported outcome measures and the critical thinking measure. The self-reported gains scores ask students to indicate how much they have improved or strengthened a particular skill; that is, how much they have learned. The CAAP Critical Thinking score measures a student's critical thinking skills at a point in time. Even with the controls for pre-college academic abilities (i.e., ACT Composite) in place, we cannot consider the critical thinking score to be a gain score. This explanation highlights a limitation in our ability to define and assess learning on this scale. This discussion has played out in the literature, related both to self-reported learning (Kuh, 2005; Pike, 1992b, 1995) and the methodological issues of predicting gains (Pascarella & Wolniak, 2004; Pike, 2004a, 2004b), and remains unresolved.

Deep Activities as Conditional Influences

This study's results suggest that there exists little, if any, ACT-conditional effect of deep engagement, regardless of the outcome measured. Counter to expectations and in contrast to other studies involving NSSE and first-year student outcomes (Kuh et. al, 2006; 2008), engagement in deep learning activities does not appear to provide any meaningful compensatory advantage for students who enter with lower levels of prior academic achievement. For, although students at all levels of ability may improve their critical thinking skills (for example) during the first college year, it should come as no surprise that eight months of college cannot erase the *gaps* in ability that took eighteen (or more) years to develop.

Of course, our statistical analysis *did* include one set of models in which the ACT by DEEP variable had a statistically significant negative coefficient. Although we believe this finding warrants additional future research, we hesitate to read too much into this particular result for two reasons. First, the finding related to gains in practical competence runs counter to results for all three other outcome measures. Second, when the effect is depicted graphically, it is hard to discern and appears to be of marginal practical significance.

Deep Cultures as Contextual Influences

Our analysis indicated that no contextual effects existed for the DEEP Learning scale scores and any of the criterion variables we explored. We hypothesized, based on previous peer-effects research and environmental theories, that students would benefit simply from being a part of an environment in which deep approaches to learning were emphasized. This was not the case. It appears that the effects of deep learning activities are solely at the individual levels—it matters that students engage in deep learning activities, but students get no benefit from a deeply engaged learning environment.

Our conceptual framework assumes that the student environment, defined as the aggregate of individual student's experiences, would affect (directly or indirectly) individual student learning outcomes (Terenzini & Reason, 2005). Results related to the contextual effects of the DEEP Learning scales call into question this assumption. We found no evidence of a direct contextual effect above and beyond the effects of an *individual* student's engagement in deep processing. However, our approach to exploring contextual effects may have masked a potential indirect effect. Decades of research (Astin & Panos, 1969; Pascarella & Terenzini, 2005) has shown that students' behaviors are influenced by the behavior of their peers. Our research design and current analytic procedures do not allow us to ascertain the extent to which a "deep" environment encourages individual students to engage in deep processing at higher levels of then they otherwise would have in a different environment.

Conclusion

Our research was guided by two questions: does engagement in deep approaches to learning, as operationalized by NSSE's DEEP Learning scales, result in more positive student learning outcomes; and, is there a contextual effect from the deep learning approaches? Results from our analyses indicate that the answer to our first question is "maybe." Certainly the strong relationships between DEEP Learning scale scores and self-reported gains in general education, practical competence, and personal and social development would lead one to conclude a positive relationship exists. This conclusion, however, must be tempered in light of the lack of relationship between the DEEP Learning scale and critical thinking. Although plausible explanations for this counterintuitive finding exist, the finding keeps us from concluding a strong, clear relationship exists between DEEP score and student learning.

The answer to our second question is much clearer. In none of our analyses did students receive benefits from a contextual effect related to deep learning. The effects of deep learning approaches, as measured by NSSE's DEEP Learning scale, appear to be only at the individual student level. That is to say, simply being a part of an environment in which deep learning is emphasized does not benefit students. If a benefit exits, students must put forth the effort to engage in deep learning activities in order to benefit.

Although our findings should not call into question the belief that engagement in deep cognitive processing results in more learning outcomes, they certainly lead us to conclude the higher education administrators cannot expect a panacea effect from environmental interventions meant to increase student engagement in deep processing, particularly as it relates to objective measures of learning outcomes. While the positive effects of engagement are well-founded, the assumption that engagement is the mechanism through which deep processing needs further study.

References

ACT, Inc. (nd). Information retrieved, May 28, 2010 from http://www.act.org/.

Aharony, N. (2006). The use of deep and surface learning strategies among students learning English as a foreign language in an internet environment. *British Journal of Educational Psychology, 76*(4), 851-866.

Astin, A. W. (1993). *What matters in college? Four critical years revisited*. San Francisco: Jossey-Bass.

Astin, A. W, & Panos, R. (1969). *The educational and vocational development of college students*. Washington, DC: American Council on Education.

Beckwith, J. B. (1991). Approaches to learning, their context and relationship to assessment performance. *Higher Education, 22*(1), 17-30.

Biggs, J. (1979). Individual differences in study processes and the quality of learning outcomes. [10.1007/BF01680526]. *Higher Education, 8*(4), 381-394.

Case, J., & Marshall, D. (2004). Between deep and surface: Procedural approaches to learning in engineering education contexts. *Studies in Higher Education, 29*(5), 605-615.

Chin, C., & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*(2), 109-138.

Cole, J., McCormick, A., & Kinzie, J. (2009, January). *Supporting first-generation students: The role of learning communities to promote deep learning.* Paper presented at the Annual Meeting of the Association of American Colleges and Universities, Washington, DC.

Elias, R. Z. (2005). Students' approaches to study in introductory accounting courses. *Journal of Education for Business, 80*(4), 194-199.

English, L., Luckett, P., & Mladenovic, R. (2004). Encouraging a deep approach to learning through curriculum design. *Accounting Education, 13*, 461-488.

Entwistle, N. 2000, "Promoting deep learning through teaching and assessment: conceptual frameworks and educational contexts", Paper presented to the Teaching and Learning Research Programme (TLRP) Conference, Leicester, England. November. URL: http://www.ed.ac.uk/etl/docs/entwistle2000.pdf

Geertshuis, S. A., & Fazey, J. A. (2006). Approaches to learning in the workplace. *Journal of Workplace Learning, 18*(1), 55-65.

Greene, B. A., Dillon, C., & Crynes, B. (2003). Distributive learning in introductory chemical engineering: University students' learning, motivation, and attitudes using a CD-Rom. *Journal of Educational Computing Research, 29*(2), 189-207

Hall, M., Ramsay, A., & Raven, J. (2004). Changing the learning environment to promote deep learning approaches in first-year accounting students. *Accounting Education, 13*, 489-505.

Holschuh, J. P. (2000). Do as I say, not as I do: High, average, and low-performing students' strategy use in biology. *Journal of College Reading and Learning, 31*(1), 94-108.

Kirby, J. R., Knapper, C. K., Evans, C. J., Carty, A. E., & Gadula, C. (2003). Approaches to learning at work and workplace climate. *International Journal of Training and Development, 7*(1), 31-52.

Kuh, G. D. (2005). Imagine asking the client: Using student and alumni surveys for accountability in higher education. In J. C. Burke (Ed.), *Achieving accountability in higher education: Balancing public, academic, and market demands* (pp. 148-172). San Francisco: Jossey-Bass.

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the effects of student engagement on first-year college grades and persistence. *Journal of Higher Education, 79,* 540-563.

Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). What matters to student success: A review of the literature *National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success*. Washington, D.C.: National Postsecondary Education Cooperative.

Marton, F., & Saljo, R. (1976). On qualitative differences in learning: Outcome and process. *British Journal of Educational Psychology, 46*(1), 4-11.

Nelson Laird, T. F., Shoup, R., & Kuh, G. D. (2005, May). *Measuring deep approaches to learning using the national survey of student engagement.* Paper presented at the Annual Meeting of the Association for Institutional Research, Chicago, IL.

Nelson Laird, T. F., Shoup, R., Kuh, G. D., & Schwarz, M. J. (2008). The effets of discipline on deep approaches to student learning and college outcomes. *Research in Higher Education, 49,* 469-494.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students, v.2: A third decade of research*. San Francisco: Jossey-Bass.

Pascarella, E. T., & Wolniak, G. C. (2004). Changre or not to change--is there a question? A response to Pike. *Journal of College Student Development, 45,* 353-355.

Pike, G. R. (1992b). Using mixed-effect structural equation models to study student academic development. *Review of Higher Education, 15,* 151-177.

Pike, G. R. (1995). The relationships between self-reports of college experiences and achievement test scores. *Research in Higher Education, 36,* 1-22.

Pike, G. R. (2004a). Change or not to change: A rose by any other name: A response to Pascarella and Wolniak. *Journal of College Student Development, 45*, 355-356.

Pike, G. R. (2004b). Lord's Paradox and the assessment of change during college. *Journal of College Student Development, 45*, 348-353.

Porter, S. R. (2005). What can multilevel models add to institutional research? In M. A. Coughlin (Ed.), Applications of intermediate/advanced statistics in institutional research (Resources in Institutional Research, no. 16) (pp. 110-131). Tallahassee, FL: Association for Institutional Research.

Ramsden, P. (2003). *Learning to teach in higher education*. London: RoutledgeFalmer.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA.: Sage.

Terenzini, P., Springer, L., Pascarella, E., & Nora, A. (1995a). Influences affecting the development of students' critical thinking skills. *Research in Higher Education, 36*, 23-29.

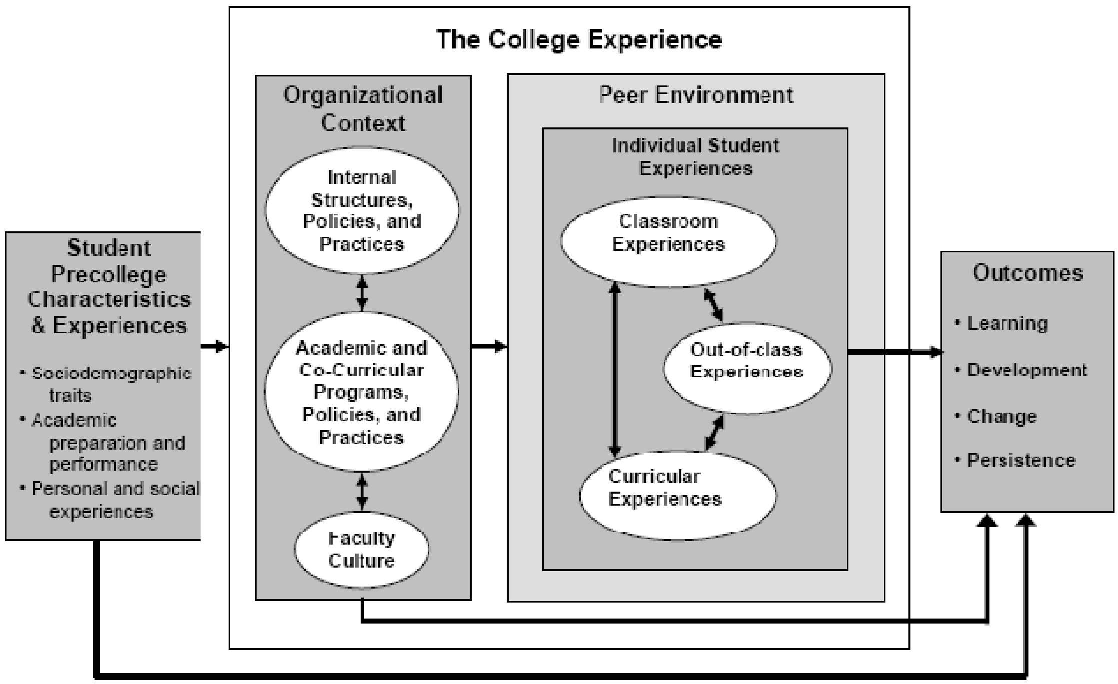
Terenzini, P., Springer, L., Pascarella, E., & Nora, A. (1995b). Academic and out-of-class influences affecting the development of students' intellectual orientations. *Review of Higher Education, 19*, 23-44.

Terenzini, P. T., & Reason, R. D. (2005, November). *Parsing the first year of college: Rethinking the effects of college on students.* Paper presented at the the Annual Meeting of the Association for the Study of Higher Education, Philadelphia, PA.

Trigwell, K., & Prosser, M. (1991). Relating learning approaches, perceptions of context and learning outcomes. *Higher Education, 22*, 251-266.

Watkins, D. (1983). Depth of processing and the quality of learning outcomes. [10.1007/BF00120900]. *Instructional Science, 12*(1), 49-58.

Figure 1. Conceptual framework for the Parsing the First Year of College Project



(Terenzini & Reason, 2005)

Table 1. Specifications for Variables Included in Analysis.

Independent Variable of Interest

NSSE DEEP Scale. A 12-item scale that assesses the level of students' engagement with activities believed to encourage deep learning. Respondents were asked how often they were required to or engaged in: analyzing elements of an idea, experience, or theory; synthesizing ideas, information, or experiences; evaluating and making judgments about information and arguments; and applying theories or concepts to new situations; included diverse perspectives in discussion or assignments; put together ideas from different courses for discussions and assignments; discussed ideas from class with faculty members outside of class; discussed ideas from classes with other students, family members, or co-workers; examined their own views on a topic; tried to better understand someone else's views on a topic; and learned something new that changed their understanding of a issue. (4-point, Likert-type scale: 1 = very little/never; 2 = some/sometimes; 3 = quite a bit/often; 4 = very much/very often)

Control Variables

Race: individual-level variable indicating students race (1 = White; 0 = other)

High school rank: individual-level variable indicating if the student graduated in the top quartile of his/her high school class.

ACT Composite score: individual-level variable indicating student's overall score on the ACT entrance assessment, scale ranges from 1 to 36.

Institutional Control: institutional-variable indicating whether the institution was public or private.

Level: institutional-level variable indicating whether the institution was a doctoral-granting institution or not.

Median ACT Composite: institution-level variable indicating selectivity of the institution; median ACT Composite score for the incoming first-year cohort in 2006.

Outcome Variables

CAAP Critical Thinking Score: A 32-item, standardized and nationally-normed assessment of students' skills in clarifying, analyzing, evaluating, and extending arguments. Scores are scaled to a national norm (based on data from 69,690 undergraduate students – not just first-year students) with a mean score of 61.65, and a standard deviation of 5.39. The range for the sample in this study was 47 to 67.

Gain in General Education. A four-item scale measuring the development of general education skills. Respondents were asked "To what extent has your experience at this institution contributed to your knowledge, skills, and personal development in the following areas: acquiring a broad general education; speaking clearly and effectively; writing clearly and effectively; and thinking critically and analytically. (4-point, Likert-type scale: 1 = very little, 2 = some, 3=quite a bit; and 4 = very much)

Gain in Practical Competence. A five-item scale measuring the development of work-related and real-world skills. Respondents were asked “To what extent has your experience at this institution contributed to your knowledge, skills, and personal development in the following areas: acquiring job or work-related knowledge and skills; working effectively with others; using computing and information technology; analyzing quantitative problems; and solving complex real-world problems. (4-point, Likert-type scale: 1 = very little, 2 = some, 3=quite a bit; and 4 = very much)

Gain in Personal and Social Development. A seven-item scale measuring the development of work-related and real-world skills. Respondents were asked “To what extent has your experience at this institution contributed to your knowledge, skills, and personal development in the following areas: developing a personal code of values and ethics; understanding yourself; understanding people of other racial and ethnic backgrounds; voting in local, state, or national elections; learning effectively on your own; contributing to the welfare of your community; developing a deepened sense of spiritually. (4-point, Likert-type scale: 1 = very little, 2 = some, 3=quite a bit; and 4 = very much)

Table 2. Models predicting students' score on the CAAP Critical Thinking module.

Model	Null	Level 1	Level 2
Intercept	61.645**	61.857**	54.354**
Level 1			
Female		-0.072	-0.073
White Race/Ethnicity		0.346**	0.352**
High School Rank		0.213*	0.213*
ACT Composite		0.861**	0.858**
DEEP Learning scale (individual student)		0.002	0.002
ACT x DEEP		0.001	0.001
Level 2			
DEEP Learning scale (school average)			0.052
Median ACT Composite			0.216*
Private Institution			-0.446
Doctoral Institution			-0.475
Residuals			
Level 1 (Sigma Squared)	19.252**	8.384**	8.383**
Level 2 (Tau)	9.081**	1.354**	1.155**

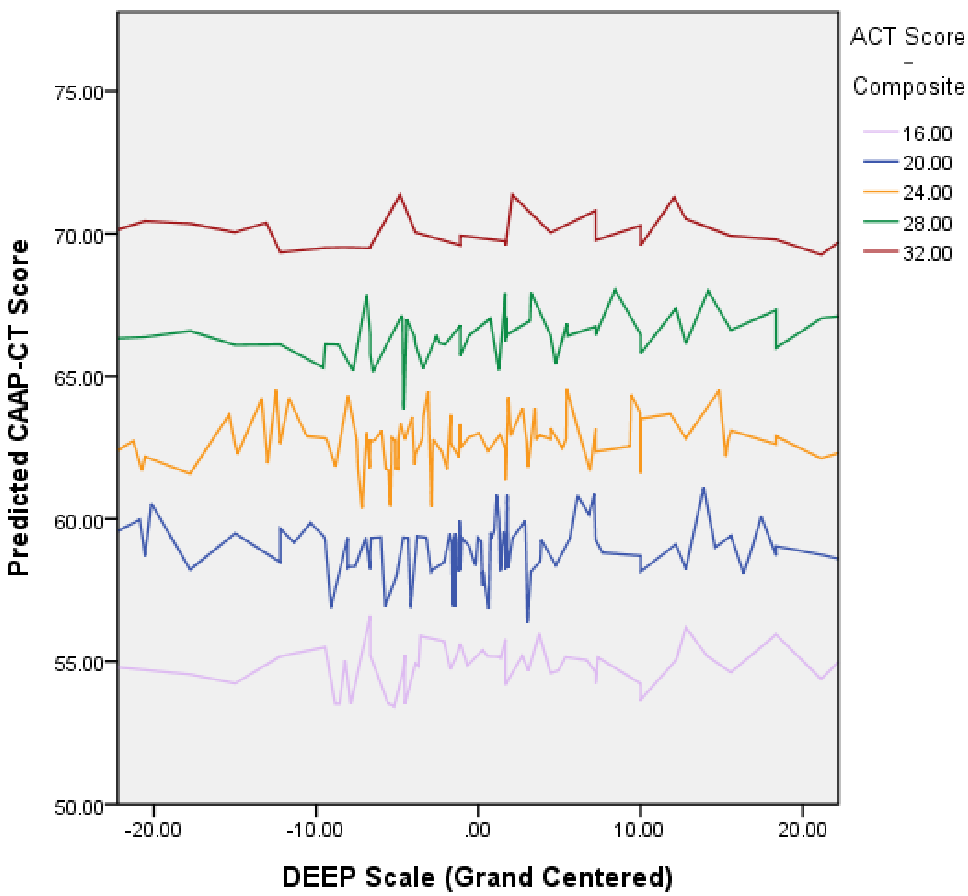
*p < .05; **p < .01

Table 3. Models predicting students' self-reported gains.

	General Education			Practical Competence			Personal & Social Development		
	Null	Level 1	Level 2	Null	Level 1	Level 2	Null	Level 1	Level 2
Intercept	67.813**	67.424**	38.055**	61.204**	60.739**	59.347**	51.154**	50.431**	46.702*
Level 1									
Female		2.543**	2.562**		1.095*	1.079*		1.890**	1.911**
White Race/Ethnicity		-0.912	-0.734		-2.288**	-2.277**		-3.109**	-3.039**
High School Rank		2.678**	2.671**		-0.250	-0.263		1.032	1.036
ACT Composite		-0.370**	-0.407**		-0.552**	-0.567**		-0.866**	-0.887**
DEEP Learning scale (individual student)		0.607**	0.603**		0.609**	0.609**		0.631**	0.630**
ACT x DEEP		0.002	0.002		-0.010**	-0.011**		-0.004	-0.004
Level 2									
DEEP Learning scale (school average)			0.317			-0.261			-0.272
Median ACT Composite			0.574*			0.803*			0.870
Private Institution			-0.047			-1.504			2.773
Doctoral Institution			-3.526**			-3.086			-2.987
Residuals									
Level 1 (Sigma Squared)	392.74**	302.59**	302.53**	393.08**	299.30**	299.40**	430.93**	323.02**	323.16**
Level 2 (Tau)	18.87**	10.03**	4.76*	12.39**	13.08**	11.03**	28.29**	28.58**	20.32**

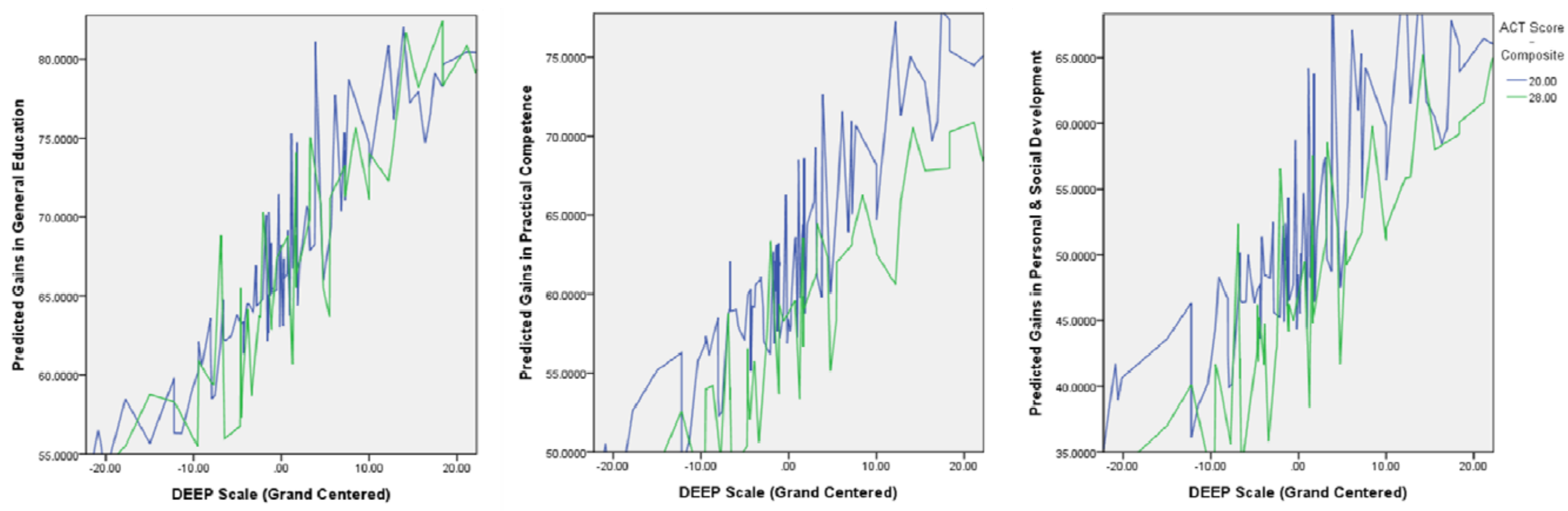
*p < .05; **p < .01

Figure 2. Average predicted critical thinking (CAAP-CT) scores for students with varying ACT Composite scores and DEEP Learning scale scores.



DEEP Learning scale scores are centered around the overall sample mean DEEP score. The graph depicts results for the middle 80% (approximately) of student scores on the DEEP Learning scales.

Figure 3. Average predicted self-reported gain scores for students with varying ACT Composite scores and DEEP Learning scale scores.



Note. The use of scores from only those students with ACT Composite scores of 20 and 28 is done for (relative) visual clarity. Graphs for students with other ACT Composite scores yield similar plots. DEEP Learning scale scores are centered around the overall sample mean DEEP score. The graph depicts results for the middle 80% (approximately) of student scores on the DEEP Learning scales