Deep Approaches and Learning Outcomes: An Exploration Using Indirect Measures

Steven S. Graunke

The purpose of this study was to explore the relationship between students’ engagement in deep approaches to learning (DAL) and measures of affective learning outcomes. Students completed 13 items designed to represent indirect assessments of their effectiveness on three institutionally defined learning outcomes. Each scale was regressed on gender, class level, enrollment in a soft discipline, and score on an overall DAL scale. Results showed statistically significant positive effect between engagement in DAL and scores on each of the three indirect assessment measures. Further research exploring the link between DAL, student academic ability, and direct measures of student learning is suggested.

Introduction

Ramsden (2003) defined approaches to learning as the ways in which individuals, usually students, interact with the material that they are trying to learn. Approaches to learning are usually divided into surface and deep strategies. Deep approaches to learning (DAL) are used in order to foster a more thorough understanding of material, while surface approaches are designed mainly for specific recall or reproduction of content, usually to meet basic requirements (Marton and Säljö, 1976; Biggs, 2003). In other words, students who use deep approaches are for more likely to remember and apply information further into the future than students who use surface approaches that might only facilitate passing a test. The use of deep strategies has been linked to higher GPA (e.g. Zhang, 2000; Zeegers, 2001), gains in higher order reasoning skills (Chapman, 2001), and possession of characteristics associated with critical thinkers (Nelson Laird, Garver, Niskodé-Dosset, & Banks, 2008). Such outcomes are consistent with the student learning and development outcomes advocated by the Council for the Advancement of Standards in Higher Education (CAS) (Council for the Advancement of Standards, 2008).

While several sources suggest that engagement in DAL improve student performance on direct measures, there has been little examination of the relationship between engagement in deep approaches and affective outcomes. For example, there are no studies to date exploring the relationship between engagement in deep approaches and students’ self-perceptions of their own learning. Biggs (1993) indicated that further understanding of the affective outcomes of assessment might provide greater insight into students’ learning processes, which could benefit higher education professionals looking to enhance student learning. The following study seeks to close this gap by examining the relationship between students’ engagement in deep approaches and self-rated abilities on three institutionally defined learning outcomes.

Deep Approaches to Learning

Marton and Säljö (1976) were the first to describe differences between deep and surface approaches. In a series of studies, they found that students who had searched for the underlying meaning of a reading passage provided more complicated descriptions of the author’s arguments, while students who focused only on
memorizing provided simple explanations. Students who searched for underlying meaning were said to be engaging in deep level processing. Subsequent researchers further explored the strategies and conditions that would inspire the use of DAL. Biggs (1987, 2003) articulated a three-stage model of learning he termed the 3P model. The first stage, the presage, encompassed factors that influence student’s likelihood to select an approach before a learning event occurs. These included students’ knowledge, motivation, and conflicting demands. The second stage, process, described students’ actual approach to learning, be it a deep or surface approach. Processes interact with assorted presage factors in order to determine the outcome, which Biggs (1987, 2003) termed the product. Biggs (1993) defined products broadly to include quantitative outcomes that assess “how much” the student has learned, qualitative outcomes that assess “how well” learning occurred, and affective outcomes that reflect students’ feelings about their learning (p. 76).

More recent literature has explored the implementation of DAL across disciplines. Nelson Laird, Shoup, Kuh, and Schwartz (2008) found that senior students in soft disciplines (disciplines in which there were less consensus about knowledge and methods) were more likely to apply DAL strategies than seniors in hard disciplines. Other studies explored the association between students’ use of DAL and increases in skills development or personal growth. Nelson Laird, Seifert, Pascarella, Mayhew, and Blaich (2011), for example, found a relationship between students use of reflective learning and gains in critical thinking as well as a significant positive effect between overall use of DAL and students’ desire to engage in more cognitively complex tasks (otherwise defined as need for cognition).

Learning Outcomes Assessment

This final line of research into the impact of DAL may be especially important given the increased scrutiny of learning outcomes assessment. In 2006, a commission convened by then Secretary of Education Margret Spellings indicated that “more and better information on the quality and cost of higher education is needed by policymakers, researchers and the general public” (U.S. Department of Education, 2006, p.22). This recommendation led to the adoption of several efforts to make attainment of learning outcomes more transparent, such as the Voluntary System of Accountability (VSA). More recently, Arum and Roksa (2011), in the book Academically Adrift: Limited Learning of College Campuses, displayed evidence from the Collegiate Learning Assessment (CLA) and the National Survey of Student Engagement (NSSE) that seemed to indicate students in higher education were making few gains in critical thinking and were not engaging in educationally purposeful activities.

Efforts such as the VSA and studies such as those published in Academically Adrift used tests of critical thinking in order to measure gains in learning outcomes. Other groups have advocated the use of authentic assessment of student products. The American Association of Colleges and Universities (AAC&U), for example, sponsored the development of the Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics as an alternative to standardized testing (AAC&U, 2009). These rubrics are designed as a guide for assessing coursework, portfolios, and other student artifacts.

Standardized tests and authentic assessment of student artifacts using rubrics constitute examples of direct assessment, as they call for students to demonstrate what
they have learned (Polumba & Banta, 1999). However, indirect methods of assessment, which ask students to reflect on their learning, are also valuable in that they provide information about students’ learning experiences as well as data for policy assessment (Ewell & Jones, 1991). Indirect assessment can be thought of as reflective of Biggs’ (1993) affective outcomes, in that they measure student’s current feeling about their learning and can provide more detail about the learning experience rather than simply examining the results of tests or other direct measures.

One important way in which indirect assessment data is collected occurs via regular student surveys. For example, since 1993, Indiana University-Purdue University Indianapolis (IUPUI) has used its Continuing Student Survey (CSS) as a method for gathering data on students’ satisfaction with their academic and social experiences (Graunke, 2011). In 2010, the survey was updated to include an indirect assessment of this institution’s comprehensive learning outcomes for all undergraduate students (Graunke & Brown, 2010). Like the AAC&U’s Essential Learning Outcomes, these learning outcomes were articulated as a set of skills integrated throughout the undergraduate experience. The indirect assessment included as part of the CSS represents one half of a comprehensive learning outcomes assessment program that also included faculty ratings of student attainment demonstrated through performance in individual courses. Pike (1999) found such self-reports to be accurate representations of student learning when aggregated at the group level.

Given the increased emphasis on measuring student learning outcomes, higher education institutions would receive considerable benefit from understanding which practices are likely to facilitate positive results on outcomes assessments. The Principles of Undergraduate Learning (PULs) at IUPUI provide a particularly useful framework for the exploration of this relationship because of the conceptual similarity between the skills emphasized at IUPUI and typical DAL strategies. For example, IUPUI defines Intellectual Depth, Breadth, and Adaptiveness (IDBA) as students’ ability to “examine and organize disciplinary ways of knowing and to apply them to specific issues and problems” (Academic Affairs, 2007). Similarly, Biggs (2003) indicated that when students are motivated to gain knowledge, “they automatically try to focus on underlying meaning, on main ideas, themes, principles or successful applications.” (p.16)

Likewise, Integration and Application of Knowledge (IAK) is defined as “The ability of students to use information and concepts from studies in multiple disciplines in their intellectual, professional, and community lives” (Academic Affairs 2007). Ramsden (2003) contended that students who engage in deep approaches to learning are able to form connections between concepts being taught and knowledge acquired in previous classes, constructs in other classes, and experiences from students’ daily lives. Finally, IUPUI also includes critical thinking amongst the PULs. Existing evidence on the impacts of DAL suggests that students who use deep approaches may make gains in areas such as higher-order reasoning and problem solving (Chapman, 2001). In addition, though tenuous, there is evidence to suggest that a relationship exists between engagement in DAL and critical thinking abilities, especially for high ability students (e.g. Nelson Laird, Garver et al., 2008).

Most existing studies, however, explore only the quantitative products associated with student learning outcomes. To date, there has been little exploration of the
affective products measured by indirect assessment methods. Using Biggs’ conceptualization of affective learning outcomes resulting from deep learning approaches, this study explores the relationship between students’ processes of studying as reported by their results on the National Survey of Student Engagement (NSSE) and subsequent indirect measures of student learning. This study might serve as a model for institutions seeking to draw connections between student activities, including engagement in DAL in both in-class and out-of-class settings, and specific learning outcomes. It will also contribute to the scholarly literature on DAL by linking engagement in DAL to students’ perceptions of their own skills and abilities.

Methods

Sample
Data used in this analysis was derived from three separate survey administrations. The first was a local administration of NSSE conducted during the Spring 2009 semester. NSSE is designed to collect data associated with students’ undergraduate educational experiences for the purposes of assessment, benchmarking, accreditation, and other institutional improvement initiatives (National Survey of Student Engagement, 2009). The NSSE was administered to a randomly selected sample of first-year and senior students. The initial survey was sent 3 weeks after the beginning of the spring semester, with 4 reminders sent over the course of 8 weeks. The final response rate for this institution was 29% in 2009. Data was also obtained from administrations of the CSS in Spring 2010 and Spring 2011. An initial invitation and 3 reminders were sent to students selected to participate in 2010, while an initial e-mail and 2 reminders were sent to potential participants in 2011.

In 2010, 22% of those sampled responded to the survey, and 18% responded in 2011. A total of 133 students responded to NSSE in 2009 and to the CSS in either 2010 or 2011. Of those, 111 completed all of the DAL items found in NSSE and could therefore be used in this analysis. The appendix provides a summary of the descriptive statistics for each variable using all 111 eligible responses.

Dependent Variables
Thirteen items from the CSS were used to create the three scales used as indirect measures of learning in this study. The first scale was designed to measure students’ self-ratings of their intellectual breadth, depth, and adaptiveness (IDBA; α = 0.816), which describes the students’ ability to understand information in their discipline, compare and contrast strategies across disciplines, and select an appropriate strategy based on the requirements of a situation (Academic Affairs, 2007).

Similarly, the second scale was designed to measure students’ self-ratings of their ability to integrate and apply knowledge (IAK; α = 0.816). IAK is characterized by students’ ability to use what they learned in college to their personal and professional lives as well as apply knowledge from different disciplines in a wide variety of settings. Both IDBA and IAK correspond with the AAC&U’s Essential Learning Outcome of Integrative and Applied Learning (AAC&U, 2010), defined as “an understanding and a disposition that a student builds across the curriculum and co-curriculum, from making simple connections among ideas and experiences to synthesizing and transferring learning to new, complex situations within and beyond the campus” (p. 1). In addition, a five item scale was used to assess students’ ratings of their critical thinking ability (CT; α = 0.920). A complete list of the items
used to create these scales is available from the author.

**Independent Variables**

The number of independent variables was limited to 4 in order to maintain adequate statistical power (Green, 1991). The first is a scale derived from students’ responses to the NSSE survey which was designed to represent an estimate of students’ use of DAL (Nelson Laird, Garver et al., 2008). The overall DAL scale is a combination of three smaller subscales which assess students’ general approach to learning. These scales can be conceptualized as the extent to which students’ generally engage in a process that would encourage a more thorough understanding of the material being learned. Nelson Laird, Garver et al. (2008) found a strong positive relationship between students’ total DAL score and their tendency to employ critical thinking skills as measured by the California Critical Thinking Dispositions Inventory. A full list of the items that comprise each subscale as well as question stems and Cronbach’s alpha for the entire scale obtained from this sample can be found in Table 1.

Three additional variables were also included. These variables were related to students’ self-ratings of these learning outcomes or had been related to one or more of the DAL measures in peer-reviewed studies. These data represent key presage characteristics which might influence students’ choices to adopt a DAL approach or affect their subsequent learning (Biggs, 2003). Data were extracted from official university records and merged with survey responses. Sex was coded as a flag for female, (female =1, male =0). The analysis also included both first-year and senior respondents to the NSSE survey in order to maximize the size of the sample. Pike (1999) observed that a halo effect was more prominent in self-rating data provided by first-year students than similar data provided by seniors, possibly because seniors had received a greater number of accurate evaluations on college-level assignments. Therefore, for this study, class level was dummy coded in the sample with students who were seniors in 2009 being coded as 1.

### Table 1

**Descriptive Statistics on Dependent and Independent Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Full sample</th>
<th>IDBA model</th>
<th>IAK Model</th>
<th>CT Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>Standard error</td>
<td>N</td>
</tr>
<tr>
<td>IDBA</td>
<td>104</td>
<td>3.15</td>
<td>0.057</td>
<td>105</td>
</tr>
<tr>
<td>IAK</td>
<td>111</td>
<td>2.76</td>
<td>0.049</td>
<td>104</td>
</tr>
<tr>
<td>CT</td>
<td>111</td>
<td>0.59</td>
<td>0.047</td>
<td>104</td>
</tr>
<tr>
<td>Deep Approaches to Learning</td>
<td>111</td>
<td>0.35</td>
<td>0.045</td>
<td>104</td>
</tr>
<tr>
<td>Female</td>
<td>111</td>
<td>0.62</td>
<td>0.046</td>
<td>104</td>
</tr>
<tr>
<td>Enrolled in a soft discipline</td>
<td>111</td>
<td>0.62</td>
<td>0.046</td>
<td>104</td>
</tr>
</tbody>
</table>

*Note: IDBA = Intellectual Depth, Breadth, and Adaptiveness, IAK = Integration and Application of Knowledge, CT = Critical Thinking. Standard error adjusted using population correction factor for a population of 10,748.*
Deep Approaches and Learning Outcomes

Doing so acknowledges the qualitative differences while simultaneously limiting comparisons between first-year and senior respondents. Finally, Nelson Laird, Shoup et al. (2008) found that students who were enrolled in soft disciplines reported using DAL approaches more often than students who were enrolled in hard disciplines. Students’ primary major or intended major as of the 2009 NSSE administration was therefore coded using the same coding scheme employed in that study (soft discipline = 1).

Analysis

Separate ordinary least squares regression analyses were conducted on each of the PUL scales using the independent variables, thus producing a different model for each scale. Only respondents who had completed every item associated with all three DAL scales as well as every item associated with the PUL scale being modeled were used in the analysis. In order to adjust for a limited population of potential respondents, a finite population correction factor was used to adjust standard errors. The correction was based on the total numbers of first-year and senior students counted during the fall 2008 university census (10,748).

Results

Means and standard errors for all variables included in the IDBA, IAK, and CT models can be found in Table 2. Variance inflation factors for all variables were below 10 in each model, suggesting that multicollinearity was not likely. Table 3 presents each full model. For IDBA, F value obtained for the OLS regression (F = 5.73) was statistically significant at the $p \leq 0.05$ level with 103 degrees of freedom. Only the DAL scale was found to be significantly related to IDBA ($B = 0.536, p \leq 0.05$). Standardized coefficient reveal for DAL suggests that one standard deviation change in students’ engagement in DAL would account for a 0.48 of a standard deviation change in students’ self-rating of their IDBA abilities. In the IAK model, the weighted combination of independent variables was statistically significant at the $p \leq 0.05$ level (F = 3.36) with 104 degrees of freedom. Like IDBA, only total DAL score ($B = 0.361$) was found to have a significant relationship to IAK at the $p \leq 0.05$ level. Standardized beta suggested that a 1

Table 2

<table>
<thead>
<tr>
<th></th>
<th>IDBA</th>
<th>IAK</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>β</td>
<td>B</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.680</td>
<td>*</td>
<td>2.027</td>
</tr>
<tr>
<td>Female</td>
<td>0.060</td>
<td>0.050</td>
<td>0.113</td>
</tr>
<tr>
<td>Senior</td>
<td>-0.094</td>
<td>-0.078</td>
<td>0.077</td>
</tr>
<tr>
<td>Soft discipline flag b</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.017</td>
</tr>
<tr>
<td>Deep Approaches to Learning</td>
<td>0.551</td>
<td>0.477*</td>
<td>0.361</td>
</tr>
<tr>
<td></td>
<td>F = 5.73</td>
<td>3.36</td>
<td>6.17</td>
</tr>
</tbody>
</table>

Note: a IDBA = Intellectual Depth, Breadth, and Adaptiveness, IAK = Integration and Application of Knowledge, CT = Critical Thinking.
b Unstandardized coefficient for IDBA soft discipline less than 0.0005. Asterisk (*) represents statistically significant at $p \leq 0.05$.
c Standard error adjusted using population correction factor for a population of 10,748.
standard deviation change in DAL score was associated with a 0.32 standard deviation change in IAK. This effect is smaller than was observed in the IDBA model. The CT model was statistically significant at the at the $p \leq 0.05$ level ($F = 6.17$) with 104 degrees of freedom. Like the other models, DAL score was variable found to have a significant relationship with students’ perceptions of their CT ability in this model ($\beta = 0.415$, $p \leq 0.05$).

**Limitations**

Several studies have demonstrated the importance of including proxy measures for academic ability (such as standardized test scores) or past academic performance (such as high school grades) in regression models describing students’ use of DAL (e.g. Nelson Laird, Garver et al., 2008; Nelson Laird et al., 2011). Because of university policies that allow transfer students the option of submitting standardized test scores and high school grades, only 76 eligible participants had an SAT score included in university records. Likewise, high school grade point average (GPA) was only available for 70 participants in the data set. These variables were excluded in order to increase the number of observations used in the analysis, thus increasing the statistical power of the model. Subsequent researchers would be well advised to consider standardized test scores and previous academic performance when conducting assessments on student learning. Another emerging line of research concerned the use of DAL by students of differing socio-economic backgrounds. For example, using parental education as a proxy for socio-economic status, Ribera (2012) found that students who had at least one parent with a graduate degree were significantly more likely to use DAL than students whose parents did not. Students who had at least one parent with a Bachelor’s degree were also significantly more likely to use DAL than first-generation students. Future research may include demographic variables such as eligibility for a Pell Grant when accounting for students’ use of DAL.

**Discussion**

The purpose of this study was to explore the relationship between engagement in DAL and three indirect measures of attainment of institutional learning outcomes. Use of indirect assessment measures is appropriate as they represent an example of the affective products of a learning experience described by Biggs (1993). This study did find evidence of a relationship between all of the indirect measures included in this study, which would suggest that engagement in DAL could produce the types of affective outcomes described by Biggs (2003). The results suggest that encouraging students’ use of DAL is an effective means of improving the results of learning outcomes assessments. Both Biggs and Ramsden (2003) provide examples of pedagogical strategies faculty could employ to promote students’ use of DAL. Institutions seeking to identify strategies to improve performance on learning outcomes assessments, and therefore vicariously obtain greater evidence of attainment of learning outcomes, would do well to explore the relationship between DAL and affective measures.

Several authors (e.g. Ewell & Jones, 1991; Pike, 1999; Polumba & Banta, 1999, Volkwein, 2011) have suggested combining multiple assessment measures to gain a comprehensive perspective on the effectiveness of learning activities. Engagement in DAL has been shown to be related to some direct measures, such as GPA (e.g. Zhang, 2000; Zeegers, 2001). Measuring the impact of DAL through
Multiple methods might provide a broad range of evidence to identify opportunities for improving teaching on campus. However, with the increased emphasis on authentic assessment methods, such as e-portfolios or artifacts evaluated via the VALUE rubrics, it would be important to establish the relationship between engagement in DAL and performance on more authentic assessment measures. Future researchers could consider exploring this relationship more in-depth.

The results of this study do suggest that engagement in DAL may be related to self-rated performance on IDBA and IAK. However, the weights associated with the IDBA and IAK equations suggest that DAL may impact gains in IDBA differently than gains in IAK. Both IDBA and IAK were related to the Essential Learning Outcome Integrative and Applied Learning (AAC&U, 2009) and were related to study strategies associated with DAL (Biggs, 2003, Ramsden, 2003). Differences in the effect of DAL on self-ratings of IDBA and IAK suggest that engagement in DAL may have differential effects depending on how learning is measured. The VALUE rubric for Integrative and Applied Learning emphasizes a wide variety of skills and abilities, including students’ ability to connect knowledge and experience, observe relationships across disciplines, and apply concepts across multiple dimensions of experience (AAC&U, 2010). Future studies using authentic assessment methods may wish to explore differences in the effect of DAL on students’ depth of disciplinary knowledge as well as their ability to integrate knowledge from multiple disciplines or across settings. Also, as previously mentioned, the interpretation of the results is hindered by the inability to include standardized test scores or high school GPA in the model. The relationship between engagement in DAL and academic ability has been well documented in previous studies (e.g. Nelson Laird, Garver, et.al, 2008; Nelson Laird et al., 2011). Previous studies, combined with the results of the current study, suggest that the relationship between DAL and learning outcomes linked to IDBA, IAK, and Integrative and Applied Learning may be impacted by students’ academic ability. Further exploration of this point is warranted.

This study also provides evidence supporting a relationship between engagement in DAL and students’ impressions of their improvements in critical thinking. Taken together, this study and studies that directly measure critical thinking suggest that students who engage in DAL have a stronger impression of their critical thinking abilities and are able to demonstrate such an improvement via standardized tests. Faculty seeking to improve students critical thinking skills would be well served to engage in pedagogical strategies designed to facilitate the use of DAL in class. Student affairs practitioners seeking to enhance student learning in compliance with CAS standards may also look to identify ways to integrate DAL in programs they design or services they provide.

**Conclusion**

Students engagement in deep approaches to learning appear to be related to self-reports of critical thinking ability, as well as reports of ability to integrate and apply knowledge and their intellectual depth, breadth, and adaptiveness. These institutionally defined learning outcomes are similar to two of the AAC&U Essential Learning Outcomes (2009). The results of this study support Biggs’ (1993) contention of a relationship between deep approaches to learning and affective outcomes. This study may be used as a model demonstrating how assessments of students’ use of deep approaches...
approaches to learning can be combined with internal indirect measures of assessment to see if students using effective learning strategies perceive educational benefits.

References


Steve Graunke is Director of the Office of Student Data, Analysis, and Evaluation at Indiana University-Purdue University Indianapolis and a student in the HESA ED.D. program. Steve received his BA in Psychology from Butler University, his MA in Counseling from Ball State University, and a Graduate Certificate in College and University Teaching from Ball State. Steve’s research interests include teaching and learning, research methods, and exploring the factors associated with college student retention and graduation.
Appendix

Deep Approaches to Learning (DAL) Scale Items.

| Overall Deep Approaches to Learning Scale | $\alpha = 0.653$ |
| Higher Order Learning subscale | $\alpha = 0.821$ |
| Coursework emphasized: ANALYZING the basic elements of an idea, experience, or theory, such as examining a particular case or situation in depth and considering its components | |
| Coursework emphasized: SYNTHESIZING and organizing ideas, information, or experiences into new, more complex interpretations and relationships | |
| Coursework emphasized: MAKING 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 | |
| Coursework emphasized: APPLYING theories or concepts to practical problems or in new situations | |

| Integrative Learning subscale | $\alpha = 0.601$ |
| Worked on a paper or project that required integrating ideas or information from various sources | |
| Included diverse perspectives (different races, religions, genders, beliefs, etc.) in class discussions or writing assignments | |
| Put together ideas or concepts from different courses when completing assignments or during class discussions | |
| Discussed ideas from your readings or classes with faculty members outside of class | |
| Discussed ideas from your readings or classes with others outside of class (students, family members, co-workers, etc.) | |

| Reflective Learning subscale | $\alpha = 0.863$ |
| Examined the strengths and weaknesses of your own views on a topic or issue | |
| Tried to better understand someone else’s views by imagining how an issue looks from his or her perspective | |
| Learned something that changed the way you understand an issue or concept | |

*Note.* $^a$ Respondents are asked “During the current school year, how much has your coursework emphasized the following mental activities?” Scale is 4 = “Very much”, 3 = “Quite a bit”, 2 = “Some”, 1 = “Very little”.

$^b$ Respondents are asked “In your experiences at your institution during the current school year, about how often have you done each of the following?” Scale is 4 = “Very often”, 3 = “Often”, 2 = “Sometimes”, 1 = “Never”.

$^c$ Respondents are asked “During the current school year, how often have you done each of the following?” Scale is 4 = “Very often”, 3 = “Often”, 2 = “Sometimes”, 1 = “Never”.

59