JoSoTL

Journal of the Scholarship of Teaching and Learning

*Volume 12, Number 4, December 2012* *josotl.indiana.edu*

JOSOTL
JOURNAL OF THE SCHOLARSHIP OF TEACHING AND LEARNING

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 1 - 11.

# The influence of involvement with faculty and mentoring on the self-efficacy and academic achievement of African American and Latino college students

**Stacie Craft DeFreitas[1] and Antonio Bravo Jr.**

*Abstract: African American and Latino college students were surveyed to examine the influence of involvement with faculty and mentoring on self-efficacy and academic achievement. It was hypothesized that involvement with faculty and mentoring were related to greater academic achievement. It was suggested that the relationship of these factors was mediated by self-efficacy. Involvement with faculty and self-efficacy were significantly related to academic achievement. The relationship between involvement with faculty and better academic achievement was partially explained by higher self-efficacy. Possible explanations for mentoring not being predictive of academic achievement in this sample were provided and the significance of faculty-student interactions was discussed.*

*Keywords: Involvement with faculty, self-efficacy, academic achievement, African American, Latino, college students*

## I. Introduction.

Despite increases in college enrollment for African American and Latino students (Aud, Fox & Kewal Ramani, 2010), they still lag behind European American students in terms of their academic performance (Fletcher & Tienda, 2010). Two methods that have been used to alleviate these deficits are mentoring programs (e.g. Santos & Reigados, 2002) and increasing general involvement with faculty (e.g. Komarraju, Musulkin, & Bhattacharya, 2010). These practices may be particularly important for ethnic minorities who may not have people in their family or social network that can support their academic endeavors (Alvarez, Blume, Cervantes & Thomas, 2009). If these practices are going to continue to be utilized by institutions to improve the performance of ethnic minority students, then it behooves us to examine them more closely to determine their effectiveness and how they influence achievement. This study examines mentoring and involvement with faculty as predictors of academic achievement as well as the role of academic self-efficacy in explaining these relationships.

### *A. Mentoring, Involvement with Faculty and Academic Achievement.*

Both mentoring and involvement with faculty are related to higher academic achievement. A group of formally mentored primarily African American and Latino college students had higher academic performance than a control group as measured by GPA and higher retention (Thile & Matt, 1995). Another group of primarily Latino and African American mentored students obtained more college credits when compared to a control group (Campbell & Campbell, 2007). Improvements in GPA have even been found when mentors are other students and not faculty

[1] Assistant Professor of Psychology, University of Houston Downtown, defreitass@uhd.edu

(Sorrentino, 2006-2007). These findings suggest that mentoring relationships are related to better academic achievement for college students.

Research has suggested that interacting with faculty in a variety of ways is related to better academic achievement (e.g. Komarraju, et al., 2010). Sax and colleagues (2005) found that merely interacting with faculty outside of class was related to higher academic achievement, especially for male college students. Lundberg and Schreiner (2004) suggested that students who have positive interactions with their faculty have better learning outcomes in college. Their finding is particularly relevant to this study because it was conducted with a large ethnically diverse sample. Interactions, such as conducting research with faculty, have been related to higher academic achievement for students as well, particularly African American students (Kim & Sax, 2009). Furthermore, Tinto (1975) states that students who are more involved with campus activities, such as interacting with faculty, will be more likely to persist academically.

*B. Mentoring, Involvement with Faculty and Self-Efficacy.*

Bandura (1997) defined self-efficacy as one's belief that they can obtain a specific goal. He discussed the sources of self-efficacy which include vicarious learning or modeling and verbal persuasion. Both mentoring and involvement with faculty are likely to impact these two sources of self-efficacy. Mentors are by definition intended to be role models to their mentees and often offer advice and suggestions for success. In addition, when students interact with faculty outside of the classroom, they are likely to learn from observation in various settings in addition to receiving advice as well. Furthermore, faculty members are likely to encourage (verbal persuasion) students with whom they interact. In turn this interaction will increase student's self-efficacy. Because faculty members are usually perceived as credible and knowledgeable, their encouragement and faith in students' academic abilities will be believed. Once students have some success after being encouraged by a mentor, they are likely to have greater increases in self-efficacy.

Among Latino college students, a mentoring program found that greater frequency of contact with one's faculty mentor was related to higher levels of self-efficacy (Santos & Reigados, 2002). Of note, one study did not find a relationship between mentoring and self-efficacy, however this is likely due to the fact that the sample included elementary school children who may not have benefited as much from their mentoring intervention (Lee & Crammond, 1999).

Faculty interactions, other than mentoring, also influence self-efficacy. Vogt (2008) found that faculty-student interactions such as sharing with the students, advising students on research, and being accessible had a positive effect on student self-efficacy. Komarraju and colleagues (2010) found that having off campus contact with faculty, feeling respected by them, and perceiving them as being approachable were all related to higher self-concept—note that their measurement of self-concept was very similar to self-efficacy. They stated that these factors are particular important for ethnic minority students who may not have other academic role models (Komarraju, et al.). Cokely (2000) found similar results with those feeling supported by faculty also having higher academic self-concept. It is clear that in some ways mentoring and involvement with faculty overlap when the mentoring is being conducted by faculty however research demonstrates a relationship between both factors and the development of self-efficacy in college students.

*C. Self-efficacy and Academic Achievement.*

Self-efficacy has a well-established influence on academic achievement (e. g. Bandura, 1990, 1997; DeFreitas, 2012). Higher self-efficacy has been directly linked to higher grades for college students (e.g. Choi, 2005), even when ability levels (Kitsantas, Winsler & Huie, 2009) and past performance (Elias & MacDonald, 2007) were controlled. A meta-analysis found that academic self-efficacy and college GPA are related also when controlling for socioeconomic status and high school performance (Robbins, Lauver, Le, Davis, & Langley, 2004). Due to the powerful influence of self-efficacy on academic achievement, it is important to consider it as a mediator for the influence of mentoring and other forms of student-faculty involvement on achievement. When students are mentored and have interactions with faculty, this is likely to improve their self-efficacy because they are encouraged and provided with a role model. This elevation in self-efficacy is related to improved academic achievement because students now believe that they can do well academically, therefore they perform better.

*D. The Current Study.*

Few studies, if any, have examined the academic achievement of African American and Latino students at a Hispanic serving institution therefore this study examines a population often overlooked. This study proposes the following hypotheses.

1. African American and Latino college students who are mentored will have better academic achievement (higher GPA) than those who are not mentored.
2. African American and Latino college students who are involved with faculty will have better academic achievement (higher GPA) than those who are not involved with faculty.
3. Furthermore, we suggest that self-efficacy mediates the relationships between these factors and academic achievement such that:
    a. Mentoring is related to higher self-efficacy which in turn is related to better academic achievement
    b. and involvement with faculty is related to higher self-efficacy which in turn is related to better academic achievement

## II. Methods.

A. Participants.

This study included 249 African American (N = 105) and Latino (N = 144) undergraduate students attending a Hispanic serving four-year institution. The average age of the sample was 24.3 (SD = 7.17). There was a large first generation college student population (Latino N=74 and African American N=27) and 57% of the sample had family incomes below $30,000. See table 1 for means and standard deviations for the sample.

B. Measures.

A demographics and academic history measure was included in which the participants reported their age, gender, ethnicity and other information about their academic experiences. GPA was taken from institutional records.

Academic self-efficacy was assessed using the Self-Regulated Learning scale of the Multidimensional Scales of Perceived Self-Efficacy (Bandura, 1990). Students reported how well they were able to complete 11 specific academic tasks on a 7-point likert scale from "not very well at all" to "very well." Questions included "how well can you use the library to get information for class assignments" and "how well can you organize your school work." The Academic self-efficacy scale was reliable for both African Americans ($\alpha = .89$) and Latinos ($\alpha = .87$).

Involvement with faculty (Millem & Berger, 1997) was assessed using the faculty subscale from the Involvement Behavior Scales and examines student interaction with faculty outside of the classroom. Students reported their experiences with faculty on a 4-point likert scale ranging from "almost never" to "often." It included questions such as how often have you "met with faculty during office hours" or "had coffee or a soft drink with a professor." This six item scale had sound reliability for African Americans ($\alpha = .75$) and Latinos ($\alpha = .79$).

The Mentoring scale assessed students mentoring relationships within the university (Gloria, Robinson Kurpius, Hamilton, & Wilson, 1999). The scale was four items and included questions such as "there is someone at [university name] that you consider a mentor or role model" and "There is someone at [university name] that cares about your educational success." Students reported on a 4-point likert scale from "strongly disagree" to "strongly agree." This scale had reasonable reliability for African Americans ($\alpha = .70$) and Latinos ($\alpha = .69$).

C. Procedure.

Students were given course credit or extra credit for their participation in this study. Students completed survey measures online after signing up and agreeing to participate in the study. Students took approximately 30 to 45 minutes to complete the survey. GPA was taken from institutional records 1 year after self-report measures were completed.

## II. Results.

A. Preliminary analysis.

Data cleaning was the first step, including analyzing the distributions of variables for factors such as normality and univariate outliers. One key variable, Involvement with Faculty, was positively skewed therefore a log transformation was conducted. After the transformation, Involvement with Faculty was normally distributed therefore in this analysis, the log of Involvement of Faculty is used.

Income data was missing for 51 participants who declined to respond. A t-test was conducted to determine whether respondents who reported income were different from those who did not report income on GPA and academic self-efficacy. There were no differences between the groups therefore in analysis missing data for income were replaced with the mean.

A correlation matrix was calculated for all variables to be included in the regression analyses (see table 2). Self-efficacy was positively related to mentoring and involvement with faculty for both African American and Latino students. Of note, GPA was only related to involvement with faculty and self-efficacy for African American students. Results of the correlation matrix were also utilized to spot multicollinearity however none was relevant. Means for all study variables are also reported in table 1. The only significant difference between

African Americans and Latinos on study variables was with age such that African Americans in the sample were significantly older, $F(1, 230) = 8.58$, $p < .01$.

**Table 1.** Means and standard deviation for all study variables reported separately for African Americans and Latinos.

| | African Americans N = 105 | | Latinos N= 144 | |
|---|---|---|---|---|
| | M | SD | M | SD |
| Involvement with faculty | .19 | .12 | .20 | .13 |
| Mentoring | 2.66 | .64 | 2.65 | .65 |
| Academic self-efficacy | 4.88 | 1.12 | 4.67 | 1.05 |
| GPA | 2.41 | .84 | 2.54 | .95 |
| Age (years)* | 26.13 | 7.93 | 23.32 | 6.64 |
| Yearly income[+] | 1.93 | 1.61 | 2.09 | 1.58 |

* $p < .01$

[+] 2 = $20,000 to $29,000

B. Primary Analysis.

The main study hypotheses were tested using hierarchical linear regression analysis. Entered at the first step for each regression equation were the background variables (Income and age). In order to determine whether involvement with faculty and mentoring were related to GPA, hierarchical regression analyses were used (see table 3). In the first step, age was a statistically significant predictor of GPA, $F(2, 229) = 3.57$, $p < .05$ but income was not suggesting that older students were more likely to have a higher GPA. It is likely that income was not a statistically significant predictor of GPA of this sample due to the fact that the university is an open enrollment institution that has students primarily from working class families. This restricted range in income was likely the cause of this lack of finding. In the second step, mentoring and involvement with faculty were entered. Only involvement with faculty was significantly related to GPA, $F(2, 227) = 8.09$, $p = .01$, so that those with more involvement were likely to have higher GPAs.

Hierarchical regression was then used to predict whether self-efficacy is predictive of GPA. The first step of the regression is noted above. In the second step, self-efficacy was significantly predictive of GPA, $F(1, 228) = 4.84$, $p<.05$, so that those with higher self-efficacy had higher GPAs.

Finally to test whether, self-efficacy mediated the relationship between involvement with faculty and GPA, another hierarchical regression was conducted to determine whether involvement with faculty predicted GPA beyond the influence of self-efficacy. This method of determining mediation was developed by Baron and Kenny (1986). As noted above both involvement with faculty and academic self-efficacy predicted GPA. The next step is to demonstrate that involvement with faculty predicts self-efficacy, which it did, $F(1, 230) = 31.68$, $p<.01$ (see table 4). The final step to determine mediation requires testing whether involvement with faculty predicts GPA beyond academic self-efficacy. Involvement with faculty was a significant predictor of GPA, $F(1, 226) = 4.84$, $p<.01$, even with self-efficacy in the model. This

indicates that self-efficacy is only a partial mediator of the relationship between involvement with faculty and GPA.

**Table 2.** Correlation matrix with African Americans above the diagonal and Latinos below the diagonal.

| | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* |
|---|---|---|---|---|---|---|---|---|
| 1. Involvement with faculty | | .42** | .35** | .30** | .22* | -.11 | -.16 | -.03 |
| 2. Mentoring | .57** | | .28** | .19 | -.10 | -.09 | -..07 | .17 |
| 3. Academic self-efficacy | .36** | .32** | | .24* | .02 | .22* | .07 | .06 |
| 4. GPA | .03 | -.05 | .13 | | -.10 | .07 | -.11 | .11 |
| 5. Gender | .02 | .12 | .14 | -.06 | | -.12 | .00 | -.03 |
| 6. Age | .08 | .11 | .17 | .25** | -.15 | | .27** | .04 |
| 7. Yearly income | -.09 | -.12 | .14 | .06 | -.16 | .27** | | .01 |
| 8. First generation student | -.11 | .12 | -.02 | -.06 | -.04 | -.02 | -.12 | |

*p<.05, **p<.01

**Table 3.** Regression analyses for prediction of GPA by mentoring and involvement with faculty.

| | | *B* | *SE B* | *95% CI B* | *β* | *t* | $\Delta R^2$ | **ΔF** |
|---|---|---|---|---|---|---|---|---|
| Model 1 | | | | | | | .03 | 3.57* |
| | Age* | .01 | .01 | [.00, .027] | .14 | 2.03 | | |
| | Income | .04 | .03 | [-.25, .10] | ..08 | 1.16 | | |
| 2 | | | | | | | .07 | 8.09*** |
| | Age | .01 | .00 | [.00, .03] | .13 | 1.98 | | |
| | Income | .05 | .03 | [-.03, .10] | .11 | 1.70 | | |
| | Mentoring | .12 | .08 | [-.05, .28] | .10 | 1.37 | | |
| | Involvement w Faculty** | 1.13 | .44 | [.27, 2.00] | .19 | 2.59 | | |

*p<.05, **p=.01, ***p<.01

## III. Discussion.

The study hypotheses were partially supported in that involvement with faculty was related to better academic achievement in African American and Latino college students. In addition, that relationship was partially explained by higher self-efficacy. However, mentoring was not predictive of academic achievement.

It is possible that mentoring was not predictive of academic achievement because we assessed an informal type of mentoring relationship that did not include information about the duration or intensity of the mentoring relationship. Best practices in mentoring suggest that a formal, structured mentoring program is likely to result in better outcomes for students than an informal one (Campbell, 2007). Due to the fact that this study did not discern the amount and

quality of mentoring that was occurring nor whether mentorships were formal, a relationship between mentoring and improvement in academic achievement may have been difficult to ascertain. In addition, this study included a unique sample of primarily first generation college students attending a Hispanic serving institution. It is possible that this group of students may require more intense mentoring to result in a substantial impact on their academic performance due to the lack of information that they may have about college. This is because first generation students often lack the knowledge that other students have about the best way to successfully navigate the academic and social demands of college, generally entering college unprepared in many ways (Engle, 2007).

**Table 4.** Regression analyses to test meditation by examining Involvement with faculty as a predictor of GPA beyond Self-efficacy.

| | | *B* | *SE B* | *95% CI B* | *β* | *t* | $\Delta R^2$ | **ΔF** |
|---|---|---|---|---|---|---|---|---|
| Model 1 | | | | | | | .03 | 7.38** |
| | Age* | .01 | .01 | [.002, .03] | .14 | 2.23 | | |
| | Income* | .07 | .15 | [.01, .12] | .15 | 2.39 | | |
| 2 | | | | | | | .03 | 7.95** |
| | Age | .01 | .06 | [-.002, .02] | .10 | 1.60 | | |
| | Income* | .06 | .03 | [.01, .12] | .14 | 2.27 | | |
| | Self-efficacy** | .12 | .04 | [.04, .21] | .17 | 2.82 | | |
| 3 | | | | | | | .02 | 6.39* |
| | Age | .01 | .01 | [-.001, .02] | .11 | 1.73 | | |
| | Income** | .07 | .03 | [.02, .13] | .16 | 2.64 | | |
| | Self-efficacy | .08 | .05 | [-.01, .17] | .11 | 1.75 | | |
| | Involvement with Faculty* | .99 | .39 | [.22, 1.76] | .16 | 2.53 | | |

*p<.05, **p<.01

As suggested by previous research (e.g. Lundberg & Schreiner, 2004), involvement with faculty was related to better academic achievement in this study. When students felt able to discuss academics and other subjects with faculty outside of the classroom they performed better in the classroom. This finding is particularly important in this study as it was conducted with African Americans and Latino students at a Hispanic serving institution. This suggests that even when ethnic students are not minorities at a university, they can still benefit from involvement with faculty. When students feel that they are respected and can be heard by faculty this is likely to impact them positively in many ways. This may be particularly important for ethnic minority students who are more likely than European American students to perceive that faculty have negative views of them and their academic potential (Museus, Nichols & Lambert, 2008). When ethnic minority students have positive interactions with faculty, this counteracts this belief such that they feel a sense of belonging in the academic environment and embrace the idea that they can have a successful academic career.

In addition, this study went beyond previous studies by suggesting that academic self-efficacy was a mediator of this relationship between faculty-student interaction and improved academic achievement. When faculty interacts with students, they increase student's belief in

their ability to achieve academically which is related to better actual performance. It is likely that faculty do this in a number of ways. Faculty may give direct information about assignments or general information about study habits. It is also likely that students are encouraged during these one on one times with professors so that they feel better about their own academic abilities. Though self-efficacy was only a partial mediator, this finding does give us some idea as to an important element of the faculty-student relationship that is related to improving the GPA of African American and Latino college students.

## IV. Limitations and Future Research.

One limitation of this study was that mentoring was not assessed more strictly to determine whether it was formal or informal. Future studies should collect information on formal mentoring with inclusion of the quality and the duration of the mentoring relationship. This could be beneficial to help us determine a relationship between formal mentoring and academic achievement. In addition, this study would be improved by having African American and Latino samples from more than one Hispanic serving institution. This would assist in generalizing the findings to other institutions of this type. Further, as with much of the extant research (e. g. Kim & Sax, 2009; Lundberg & Schreiner, 2004; Santos & Reigados, 2002) this study was correlational and therefore cannot demonstrate a causal relationship. Unfortunately, it is very difficult to develop an experimental study in which mentoring programs or involvement with faculty is assigned to a group of students through random assignment. Typically, students who do not want to participate would merely drop out of the study therefore it may not be useful to attempt to develop such a study. However, future studies should focus on the development of quasi-experimental designs such as Campbell and Campbell (2007) and Sorrentino (2006-2007) in which students that are participating in mentoring programs or involved with faculty would be matched then compared with other peers who are not involved in these activities. Finally, since self-efficacy was only a partial mediator of the relationship between faculty-student interaction other factors should be examined that may fully mediate the relationship such as increased knowledge about course requirements and materials.

## V. Implications.

Knowing the significance of faculty-student interactions, it is vital for faculty to make conscious efforts to have more positive relationships with their students especially outside of class. The first step is assisting students in knowing that faculty is available to them outside of the classroom. Students should be repeatedly invited to office hours by faculty, particularly early in the semester as well as before and after major assignments. Lundberg and Schreiner (2004) suggest using techniques like communicating more positively by smiling, making direct eye contact, avoiding direct criticism, and encouraging students to speak up in class in order to make students feel more comfortable talking to the professor. This increased comfort level should help students to request advice or help when they need it. This may be particularly important when students are first generation college students and do not know what to expect from college professors to whom they may feel subordinate. Furthermore, faculty members who show respect are also instrumental in helping students to feel motivated and capable of achieving academically (Komarraju et al., 2010). Students are likely to feel respected if professors take time to talk with them about issues that are of concern to them in a more egalitarian fashion. If faculty members

utilize this knowledge it can be the first step in beginning significant faculty-student relationships. In essence, when faculty members interact with students, particularly outside of the classroom, this is related to higher levels of academic self-efficacy for these students. They are more likely to have confidence in their ability to achieve their academic goals due to being respected by faculty that they find approachable and available (Komarraju et al.; Vogt, 2008). In addition, conscious efforts by faculty to increase self-efficacy by praising and encouraging students can also assist in improving academic achievement. If the university supports these endeavors, by doing things such as including information about building student relationships in new faculty orientation and other trainings, then faculty will put more of an effort into making these important connections with students. Strong relationships between faculty and students are particularly important for ethnic minority students who may not have others who can guide them through academic life.

## References

Alvarez, A. N., Blume, A. W., Cervantes, J. M & Thomas, L. R. (2009). Tapping the wisdom tradition: Essential elements to mentoring students of color. *Professional Psychology: Research and Practice, 40(2)*, 181-188. doi:10.1037/a0012256

Aud, S., Fox, M., and KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups*(NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington, DC:U.S. Government Printing Office. http://nces.ed.gov/pubs2010/2010015.pdf

Bandura, A. (1990). *Multidimensional Scales of Perceived Self-efficacy*. Stanford, CA: Stanford University.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 52*, 1173-1182. doi: 10.1037/0022-3514.51.6.1173 Retrieved from http://www.public.asu.edu/~davidpm/classes/psy536/Baron.pdf

Campbell, C. D. (2007). Best practices for student-faculty mentoring programs. In T. D. Allen & L. T. Eby (Eds.). *The Blackwell Handbook of Mentoring: A Multiple Perspective Approach* (pp. 325-343). Malden, MA: Blackwell publishing.

Campbell, T. A., & Campbell, D. E. (2007). Outcomes of mentoring at-risk college students: Gender and ethnic matching effects. *Mentoring & Tutoring: Partnership in Learning, 15*(2), 135-148. doi:10.1080/13611260601086287

Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools, 42*, 197-205. doi: 10.1002/pits.20048 Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/pits.20048/pdf

Cokley, K. (2000). Perceived faculty encouragement and its influence on college students. *Journal of College Student Development*, 41(3), 348-352.

DeFreitas, S. C. (2012). Ethnic differences in the relationship between self-efficacy, outcome expectations, and academic achievement for first-year college students. *Social Psychology of Education, 15(1),* 109-123. *DOI 10.1007/s11218-011-9172-0*

Elias, S. M. & MacDonald, S. (2007). Using past performance, proxy efficacy, and academic self-efficacy to predict college performance. *Journal of Applied Social Psychology, 37 (11),* 2518-2531. doi:10.1111/j.1559-1816.2007.00268.x Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1559-1816.2007.00268.x/pdf

Engle, J. (2007). Postsecondary access and success for first generation college students. *American Academic, 3(1),* 25-48. Retrieved from http://www.aft.org/pdfs/highered/academic/january07/Engle.pdf

Fletcher, J., & Tienda, M. (2010). Race and ethnic differences in college achievement: Does high school attended matter? *The ANNALS of the American Academy of Political and Social Science, 627*(1), 144-166. doi:10.1177/0002716209348749 Retrieved from http://ann.sagepub.com/content/627/1/144.full.pdf+html

Gloria, A. M., Robinson Kurpius, S. E., Hamilton, K. D., & Wilson, M. S. (1999). African American students' persistence at a predominately white university: Influences of social support, university comfort, and self-beliefs. *Journal of College Student Development, 40*(3), 257-268.

Kim, Y. K., & Sax, L. J. (2009). Student-faculty interaction in research universities: Differences by student gender, race, social class, and first generation status. *Research in Higher Education, 50*(5), 437-459. Retrieved from http://www.chapman.edu/students/housing-and-residence/_files/housing-pdfs/Kim-Sax-2009.pdf

Kitsantas, A., Winsler, A. & Huie, F. (2009). Self-regulation and ability predictors of academic success during college: A predictive validity study. *Journal of Advanced Academics 20(1)*: 42-68.

Komarraju, M., Musulkin, S. & Bhattacharya, G. (2010). Role of student-faculty interactions in developing college students' academic self-concept, motivation and achievement. *Journal of College Student Development, 51*(3), 332-342.

Lee, J. & Cramond, B. (1999). The positive effects of mentoring economically disadvantaged students. *Professional School Counseling, 2*(3), 172-178.

Lundberg, C. A., & Schreiner, L. A. (2004). Quality and Frequency of Faculty-Student Interaction as Predictors of Learning: An Analysis by Student Race/Ethnicity. *Journal of College Student Development, 45*(5), 549-565. doi:10.1353/csd.2004.0061 Retrieved from http://faculty.uca.edu/kevinh/Lundberg%20article.pdf

Milem, J. F., & Berger, J. B. (1997). A modified model of college student persistence: Exploring the relationship between Astin's theory of involvement and Tinto's theory of student departure. *Journal of College Student Development, 38*(4), 387-400.

Museus, S. D., Nichols, A. H., & Lambert, A. D. (2008). Racial differences in the effects of campus racial climate on degree completion: A structural equation model. *Review of Higher Education: Journal of the Association for the Study of Higher Education*, *32*(1), 107-134. doi:10.1353/rhe.0.0030 Retrieved from http://muse.jhu.edu/journals/review_of_higher_education/v032/32.1.museus.html

Robbins, S. B., Lauver, K., Le, H., Davis, R. & Langley, A. (2004). Do psychosocial and study skills factors predict college outcomes? *Psychological Bulletin, 130*, 261-288.

Santos, S.J., & Reigadas, E. T. (2002). Latinos in higher education: An evaluation of a university faculty mentoring program. *Journal of Hispanic Higher Education, 1*(1), 40-50. Retrieved from http://jhh.sagepub.com/content/1/1/40.full.pdf+html

Sorrentino, D. M. (2006-2007). The Seek Mentoring Program: An application of the goal-setting theory. *Journal of College Student Retention, 8(2),* 241-250. Retrieved from http://baywood.metapress.com/media/1fpnnmqqxqf5ac159dcg/contributions/7/d/9/t/7d9td30yn9t08dwl.pdf

Thile, E. & Matt, G. (1995). The Ethnic Mentor Undergraduate Program: A brief description and preliminary findings. *Journal of Multicultural Counseling and Development, 23(2),* 116-126.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89-125. doi: 10.3102/00346543045001089 Retrieved from http://rer.sagepub.com/content/45/1/89.full.pdf+html

Vogt, C. M. (2008). Faculty as a critical juncture in student retention and performance in engineering programs. *Journal of Engineering Education, 97*(1), 27-36.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 12 – 28.

# A qualitative examination of connections between learner-centered teaching and past significant learning experiences

**Tim Brackenbury**[1]

*Abstract: Learner-centered teaching is a collection of instructional practices that shift the emphasis of courses from the instructors' goals and methods of delivery to the knowledge and skills that the students develop. This study examined potential commonalities between features of learner-centered teaching and the past significant learning experiences of current faculty. A phenomenological analysis of written essays revealed eight dominant themes: 1) Student responsibility for learning, 2) Learning through direct experience or example, 3) Responsive instructors, 4) Difficult activities that took time, 5) Connections to previous knowledge and experiences, 6) Direct research experience, 7) Challenging initial ideas and assumptions, and 8) Rich in content. These themes are discussed in terms of their connections to features of learner-centered teaching and potential implications for educators.*

*Keywords: Learner-centered teaching, significant learning experiences, faculty reflections, student learning, active learning.*

Learner-centered teaching is a paradigm of challenge. It challenges students to become active agents in their learning. Students must face complex problems in order to acquire new knowledge and skills, while also developing new ways of thinking and acting. Learner-centered teaching challenges instructors to release some of their control over the class and what happens. They must care about more than just content; treat student errors as learning opportunities; and change their role from distributors of knowledge to facilitators of learning. Learner-centered teaching challenges both students and instructors (as well as administrative and discipline stakeholders) to carefully consider the kinds of professionals/people that graduates should be, and the optimal practices for achieving these desired outcomes (e.g., Weimer, 2002).

These challenges can be daunting to educators and students who are new to learner-centered teaching because they appear to be dramatically different from the methods of education that most individuals have experienced before (i.e., traditional, instructor-centered models). But, are they significantly different from all previous educational experiences or just certain common ones (such as lecture)? As Barr and Tagg (1993) stated, "… not all elements of the new paradigm are contrary to corresponding elements of the old: the new includes many elements of the old within its larger domain of possibilities" (p. 15). The primary objective of the present study was to examine commonalities and differences between features of learner-centered teaching and past learning experiences that individuals identified as particularly meaningful and valuable to their development (i.e., significant learning experiences). To do so, applied features of learner-centered teaching, defined as the characteristics that are directly experienced by educators and students (as opposed to the theoretical constructs behind them),

[1] Department of Communication Sciences and Disorders, Bowling Green State University, 200 Health Center Building, Bowling Green, OH 43403, tbracke@bgsu.edu

were compared with themes derived from faculty reflections on their past significant learning experiences.

*A. Applied Features of Learner-centered Teaching.*

Learner-centered teaching includes a number of different methods (such as problem-based learning, service learning, and team learning) that are based on research in cognitive development and effective teaching practices. This literature review focuses on three applied features of learner-centered teaching: a constructive basis for learning, the acquisition of knowledge and skills that are purposeful and transferable, and instructor changes that occur when adopting learner-centered approaches.

Learner-centered teaching emphasizes knowledge and skills that are constructed by students, rather than directed by instructors. Based on the collective writings of educators and psychologists such as Bruner, Dewey, Piaget, and Vygotsky, constructivism posits that learning is created by individuals and groups as the result of their current knowledge/thoughts/beliefs interacting with new experiences (e.g., Hinchliffe, 2011; Richardson, 1997; Savery & Duffy, 1995). In other words, new knowledge and skills are created as learners attempt to make sense of incongruences between their current knowledge and new experiences. Constructivism does not presuppose that learning cannot occur from direct transmission (e.g., attending a lecture). Rather, it suggests that such learning is not as complex, meaningful, or enduring as knowledge that is constructed by the individual. As a result, learner-centered teaching emphasizes learning activities and experiences in which instructors facilitate, rather than dictate, students' construction of knowledge.

A second applied feature of learner-centered teaching is that the knowledge and skills that are acquired/constructed are purposeful, relevant, and transferable. The terms "surface" and "deep" describe different types of knowledge and skills, as well as the approaches that students and faculty take toward learning (e.g., Marton & Säljö, 1976; Rhem, 1995). Surface learning describes the acquisition and recall of facts from experiences. The knowledge achieved here is considered surface because there is little beyond identification and recall that students can do with it. Facts learned while watching Jeopardy or playing Trivial Pursuit are examples of surface knowledge. Deep learning also includes the acquisition of information, but emphasizes students' abilities to apply these ideas to new and varied contexts. Perkins (2008) described two types of deep learning/knowledge, performative and proactive. Performative knowledge addresses students' abilities to use their current knowledge (including surface learning) to solve complex and atypical problems. Proactive knowledge goes further by applying current knowledge in wholly novel contexts.

The notions of surface and deep learning/knowledge are also reflected in Bloom's taxonomy (1956). Developed specifically to assist teachers/instructors with educational objectives, this taxonomy identified six levels of cognitive functioning. They are (in order from simple to complex) knowledge, comprehension, application, analysis, synthesis, and evaluation. The knowledge, comprehension, and application levels appear to be reflections of surface learning, while the remaining levels address deep learning. Although Bloom's taxonomy has been criticized and modified over the years (e.g., Marzano & Kendall, 2007), it endures within learner-centered teaching because it identifies multiple aspects of knowledge and their applications to student development.

Weimer (2002) provided a detailed description of a third applied feature of learner centered teaching; the changes that occur when an instructor moves from a traditional to learner-centered approach. These include changes to the balance of power, the function of content, the role of the teacher, the responsibility for learning, and the purpose and process of evaluation (see Table 1). These five changes reflect the dominant aspects of learner-centered teaching as experienced by educators and students, and as presented in the literature (although Weimer's terms are not used consistently across authors).

**Table 1.** Weimer's (2002) list of changes that occur between traditional, instructor-centered and learner-centered teaching models.

| **Change** | **Traditional, Instructor-centered Models** | **Learner-centered Models** |
|---|---|---|
| The balance of power | Instructors design courses with no student input and are in charge of everything. | Instructors and students work together to design and implement the course. |
| The function of content | Content is the primary force driving the course because it provides the foundation upon which skills can be built. | Content is a tool to develop a knowledge base, new ways of thinking, and self-awareness of learning. |
| The role of the teacher | Knowledge and skills are directly passed from the teacher to the students. | Students develop knowledge and skills via discovery; indirectly from the teacher. |
| The responsibility for learning | Rules and regulations are developed by the instructor to direct student learning. | Course design accentuates student learning and students' roles in the process. |
| The purpose and process of evaluation | The measurement of student progress to date. | Assessment is directed at improving instruction and providing students additional learning opportunities. |

*The balance of power.* In traditional, instructor-centered classrooms, all of the decisions about the course (including the course goals, learning experiences, assignments, assessments, and topics discussed) are developed and implemented by the instructor. In a learner-centered classroom, the design of the course is developed through collaboration between the instructor and students. Learner-centered instructors do not abdicate all of the decision making to their students. Rather, they include the students as colleagues within the development process. In shifting the balance of power, learner-centered instructors are attempting to increase student engagement (as students now become part of the process), and facilitate their path towards constructing their knowledge and becoming independent learners (see also Doyle, 2008).

*The function of content.* Content is one of the primary forces in traditional teaching. Because disciplines are consistently adding new content to their fields, instructor-centered courses are typically designed to cover as much content as possible. This restricts instructors from a) going into much depth about the content and how it can be used, and b) allowing students to explore and apply the content. As a result, their students focus on surface-level knowledge of the content, often memorizing the information and forgetting it shortly after the experience is over. In learner-centered teaching, content is only one aspect of courses. One example of this is Fink's (2008) taxonomy of significant learning. This taxonomy is built around six aspects of course design that would purportedly result in learning that is significant, lasting,

and valued later in life. In other words, deep learning. The categories within this taxonomy were: foundational knowledge (including content), application, integration, the human dimension, caring about one's own learning, and learning how to learn. Thus, content is not forgotten or ignored in learner-centered teaching. Rather, it is placed within a larger context by using content to motivate learning in the other categories and vice versa. For example, students within a learner-centered class might be given an application problem that is somewhat outside their current knowledge. After doing some initial work on the problem, they could then be directed toward a source of content (e.g., a book or article) as an aid to solving the problem. As a result, the students would not only be exposed to the content but learn it within a meaningful context. They also learn how to use informative sources to solve problems.

*The role of the teacher.* In the traditional classroom, the role of the instructor is that of the knowing professional who dispenses her/his knowledge directly to the students. This is shown through methods such as lectures that describe new and complicated topics, stories from their past relevant experiences, and summaries of the material that is most relevant to examinations. In learner-centered teaching, instructors provide the architecture for learning but do not directly state all of content to be learned. Instead, they design class activities that help students discover the important information. As a result, students learn more from the experience and each other than directly from the teacher. For example, rather than telling students about the speed of sound and having them solve equations on the topic, an instructor may have them measure and calculate the speed of sound for themselves using an oscilloscope and tape measure (Stoll, 2008). To paraphrase Alison King's (1993) often cited article title, adapting a learner-centered approach changes an instructor, "from sage on the stage to guide on the side."

*The responsibility for learning.* Weimer (2002) points out that in response to students who are often ill prepared for college education, many faculty have developed strict rules and guidelines. Course syllabi are typically the place where such regulations are listed, and a good illustration of how learner-centered classrooms emphasize the responsibility differently. The majority of syllabi focus on information about the instructor, descriptions of the course and its objectives, the topics to be addressed, and grading policy (Doolittle & Siudzinski; 2010; Eberly, Newton, & Wiggins, 2001; Garavalia, Hummel, Wiley, & Huitt; 1999). Instead of focusing on rules and regulations, the syllabi from learning-centered courses are designed as a "spring-board for the course experience" (Eberly et al., p.68). These syllabi establish the class environment as stimulating and collaborative, emphasize what will be done to facilitate student learning and who is responsible for the actions that take place, and are developed with student input (e.g., Bain, 2004; O'Brien, Millis, & Cohen, 2008). By doing so, learner-centered syllabi demonstrate from the start of a course that the primary responsibility for learning belongs to the students and that the faculty member is there to facilitate the process.

*The purpose and process of evaluation.* The changes discussed thus far have emphasized the instructor and students. Weimer's (2002) final change focuses on the purpose and process of evaluation. Evaluation is a critical part of the teaching process because it identifies if learning has taken place and, if so, what types of learning occurred. The primary purpose of evaluation in traditional classrooms is as a measure of achievement to date. It is typically summative; in that once the evaluation task has been completed there are no opportunities to show further learning (except, perhaps, from a cumulative evaluation at the end of the course). Evaluations in learner-centered classrooms are both formative and summative (e.g., Black & Wiliam, 2009; Fink, 2003; Rushton, 2005; Yorke, 2003). Formative evaluations differ from summative ones because their purposes are a) to assess current knowledge and skills in order to improving learning, while b)

providing additional opportunities for student learning. Feedback is an important part of formative evaluation. It extends beyond the identification of correct and incorrect responses to address the lines of thought that lead to errors and promote learning. As Bain (2004) stated, "the primary goal [of this type of assessment] is to help students learn to think about their own thinking so they can use the standards of the discipline or profession to recognize shortcomings and correct their reasoning as they go (p. 160)."

*B. Significant Learning Experiences and the Present Study.*

In his 2003 book on designing college courses, L. Dee Fink emphasized *Creating Significant Learning Experiences*. He described significant learning experiences as an engaged, high-energy process that results in meaningful and lasting change that has value in students' lives. Although Fink included detailed instructions for creating learner-centered courses aimed at providing significant learning experiences, he did not report any data on whether or not such lasting experiences occurred. He is not alone. Although the scholarship of teaching and learning literature includes a large number of qualitative, quantitative, mixed-methods, survey, and experiential research in support of learner-centered teaching (e.g., Black, 1993; Hake, 1998; Lambert & McCombs, 1997; Prince, 2004; Weimer, 2006), very little research exists on the long-term effects of learner-centered teaching (i.e., years after the learning experience). Despite this, significant learning experiences do occur, as evidenced by instructors and students who can easily recall past learning experiences that were engaging and lead to meaningful and lasting change.

The primary research objective of the present study is to determine if there are commonalities between features of learner-centered teaching and past significant learning experiences. If so, what are they and what differences exist? This is a retrospective, exploratory study. No a priori predictions were made as to which features of learner-centered teaching would connect with individuals' past significant learning experiences and which would not.

## I. Method.

The method of investigation was that of phenomenological research, specifically psychological phenomenology (e.g., Creswell, 1998). This form of qualitative research examines the experiences of multiple people in regards to a particular event or phenomenon (Creswell, 2011; Miles & Huberman, 1994; Moustakas, 1994). In this investigation, the individuals were twenty-four faculty members and administrators in the discipline of communication sciences and disorders. The events/phenomena were their past significant learning experiences. In general, the procedures of phenomenology involve identifying the phenomenon and individuals who have experienced it, collecting data from the individuals about their experience, identifying individual idea statements and grouping them into codes and themes, and reflecting on the meanings of the experience.

*A. Participants.*

Demographic information was available, via self-report, for 22 of the 24 participants. The participants included 16 faculty members, eight administrators, and two clinical instructors/supervisors (participants were allowed to select more than one category). They were

from 15 different states in the United States and one was from New Zealand. Their college-level teaching experiences ranged from 2 to 40 years (median = 15) and their typical teaching load within a 9-month school year was from 0 to 10 courses (median = 4). The programs that they worked in graduated an estimated 6 to 400 undergraduate (median = 30) and 12 to 65 master's students (median = 20) per year. Information on the gender or ages of the participants was not collected (feminine pronouns will be used below when referring to specific participants).

*B. Data Collection.*

All of the participants were self-selected attendees at a presentation during the 2010 Council on Academic Programs in Communication Sciences and Disorders Annual Convention titled *Everything I need to Know About Teaching I Learned From Speech-Language Therapy* (Brackenbury & Shaughnessy, 2010). The stated goals of this session were to 1) introduce attendees to aspects of learner-centered teaching, 2) model and demonstrate learner-centered teaching, and 3) entice attendees to learn more about and increase their use of learner-centered teaching. The session began by stating these goals and providing background information about the presenters. The attendees were then briefly introduced to the idea that many aspects of learner-centered teaching are rooted in past experiences (although no specific features of learner-centered teaching had yet been addressed).

Next, the attendees were directed to take ten minutes and write an essay about a significant learning experience from their own college education. They were given sheets of paper and instructed to:

> Identify an experience through which you really learned the heck out of something (either a class, an experience within a class, or a clinic/research event). What was it? How did this experience improve you (knowledge, skills, insight, caring, ability to learn…)? What did the instructor do to facilitate your learning? Why does this experience stand out over all of the others?

The session then focused on five features of client-centered therapy. After each feature was introduced, the attendees were instructed to write about if/how the feature related to the significant learning experience (see Brackenbury, 2011). The session concluded with one of the presenters sharing her results from engaging in this same process, a description of how the method of the presentation demonstrated aspects of learner-centered teaching, and suggested references for further learning. Of the approximately 80 people who attended the session, 24 turned in a signed consent form and written essay.

*C. Analysis.*

The data analysis followed procedures for psychological phenomenology as outlined by Creswell (1998, 2011). It began with the author and two graduate research assistants identifying their own thoughts, experiences, and biases on the subject of learner-centered teaching and past significant learning experiences. The purpose of this was for each member of the research team to identify and limit the potential effects of their own bias on the analysis and interpretation (i.e., bracketing). The author was the developer and principal investigator of the study. He was one of the co-authors of the conference presentation where the data were collected. At the time of the data collection, he had been a certified speech-language pathologist for 19 years and taught at the university level for nine years. He had used client-based therapy throughout his clinical work and

had transitioned his teaching from the traditional model to learning-centered approximately four years earlier. He was well read on learner-centered teaching, but did not directly use either the general literature on the scholarship of teaching and learning or the references cited above during the analysis process. At the outset of the study, he had suspected that there were some connections between past significant learning experiences and aspects of learner-centered teaching (based on his own college experiences), and was curious to examine which features did and did not relate. The two graduate students were first-year master's students in communication sciences and disorders. They had both experienced aspects of learner-centered teaching during their undergraduate educations, but were not familiar with its literature. They were given the handout from the conference session and discussed it with the author once, before the analysis process. They reported that they could envision connections between past significant learning experiences and learner-centered teaching, but were unsure what those connections might be.

In the second step of the analysis, the author and research assistants inputted the participants' responses into a spreadsheet. The responses were typed in their original form, with adjustments made only for spelling and minor grammatical errors. Questions about responses that were illegible or difficult to understand were resolved through discussion between the research team members. Once all of the responses were entered, the author read through the entire spreadsheet and double-checked it with the written essays to make sure that all of the entries were accurate and comprehensible.

Next, each participant's essay was separated into individual ideas, which were then sorted into codes and themes. Each member of the research team, working separately, divided the responses into individual ideas and placed each idea in its own spreadsheet cell (i.e. horizontalization of the data). Every idea was then considered for the potential code it represented and copied into another spreadsheet, organized by theme. Thus, the themes were derived from the data themselves and not any previously articulated decisions. Once completed, each member of the research team shared their codes and themes (along with the idea statements within them) with the other members of the team. A dialogue process was then used to reach consensus on the final set of themes.

As with other forms of research, phenomenological investigations are concerned with the credibility and transferability of their findings. Of particular concern are the potential negative influences of bias from the examiners and the degree to which the information presented represents the actual experiences of the participants (Creswell, 2011; Maxwell, 2005; Miles & Huberman, 1994). Credibility and transferability were established in a number of ways, including the above outlined procedures of bracketing, horizontalization, and triangulation during the code/theme development. In addition, having the participants write their own descriptions minimized potential interferences between their thoughts and reflections and the data analyzed. The median length of the participants' descriptions of their significant learning experiences was 169 words (with a range of 59 to 459). Thus, the 10 minutes they were given to write appears to have been sufficient for the participants to reflect and describe their experiences. Because they held onto their original essays throughout the presentation, the participants were able to make revisions at any time during the session (although they were not instructed to do so). Visual inspection of the written essays and the flow of ideas within them suggest that few such changes were made.

## II. Results.

The participants wrote about a number of different types of significant learning experiences. Some presented aspects of design for an entire course. These included courses that were based on direct "hands on" experiences, discussions, factual information, research experiences, research evidence, self-direction, service experiences, and student accountability. Other participants focused on specific activities within a class, such as case presentations, clinical experiences, designing their own final examination, and interpersonal interactions. Research was a third topic that was addressed. Participants also wrote about their master's thesis, doctoral dissertation, and research projects that were part of other courses. The other experiences described included direct clinical work, a workshop attended after graduating, and studying with other students.

The research team identified eight themes within the participants' descriptions of their past significant learning experiences. Table 2 presents a summary list of the themes, along with a sample statement for each one. Each theme is described in detail below. In both the table and text, the themes are presented in order from those with the greatest number of supporting idea statements to those with the fewest idea statements. It is important to note that the number of supporting idea statements is not considered a measure of a theme's importance. Themes with many idea statements may, for example, reflect common, cursory thoughts; while themes with few idea statements may reflect concepts that are innovative and deep.

### *A. Theme I: Student Responsibility for Learning.*

The participants included many statements that identified themselves as active agents in their learning. These occurred in a variety of activities, including identifying a researcher and following their "question trail" back in the literature, self advising, creating case studies, designing and conducting research, developing a final assessment activity, and teaching their fellow students. Their descriptions of these activities included action words such as "back-track/reanalyze," "choose," "connect," "control," "create," "develop," "dig," "discover", "engage," "experience," "identify," and "learn."

Although this theme focuses on students' responsibility for their own learning, instructors were reported to have had important roles in this process. One participant stated, "My professor did very little, yet her "minimal" teaching was significant in my learning. She allowed me to be in control of my own learning. She was supportive and strived as a guide." Another said,

> He [the instructor] required each of us to be the expert at one reading/book each; to write a short summary for him, present the summary to the class and prepare discussion topics… He simultaneously honored our experience and assumed that we would be responsible.

In addition to describing their role in the learning process, some of the participants also made statements regarding the importance of these experiences. For example, one participant reported, "This advising method helped me to become independent and in time [lead me] to what I desired out of my schooling and how I wished to change." Another wrote, "I felt empowered like I could make a difference in the world. I didn't feel like a student. I felt like a professional and I acted like a professional." A third wrote, "As a student, I did "more" than what I thought my professor expected. Hence, I expect more of my students. When given the opportunity to construct their own knowledge they always go above and beyond my expectations."

**Table 2.** The 8 themes identified from the participants' reflections on their past significant learning experiences and sample statements of them.

| Theme | Sample Statement |
|---|---|
| I. Student Responsibility for Learning. | My professor did very little, yet her "minimal" teaching was significant in my learning. She allowed me to be in control of my own learning. She was supportive and strived as a guide. |
| II. Learning through direct experience or example. | This activity gave me hands-on experience with preschool children. It required me to apply my classroom-based knowledge of child language development to preschoolers. It improved my knowledge of the course content and my skills of directing activities with children. |
| III. Responsive Instructors. | The instructor provided guidance, engaged in conversation allowing me to openly share what I had learned (or thought I learned), and asked thoughtful questions to further my understanding and learning. |
| IV. Difficult activities that took time. | It was frustrating in a way because there were no "right" or simple answers and there were numerous ways to get to the conclusion. |
| V. Connections to previous knowledge and experiences. | I learned to figure out relationships of existing knowledge to what I was doing and figure out how it fit or did not fit into my research. |
| VI. Direct research experience. | Reading the literature and interpreting the data in novel ways piqued my interest and enhanced my confidence for research. |
| VII. Challenging initial ideas and assumptions. | The learning style by Dr. V was confrontational, debating each premise, sometimes caused frustration, but by the end of each class, we had the satisfaction of knowing that we had learned something meaningful, or we needed to dig a little deeper to find the answer. |
| VIII. Rich in Content. | I was in this course (which really could've been an Intro to speech-language-pathology course) where I couldn't just BS my way through. There were facts, for the first time, not just opinions (even learned opinions). There were answers, not just suppositions and points of view, and I was enamored in it. |

*B. Theme II: Learning Through Direct Experience or Example.*

The participants described a number of experiences in which they were directly involved in a professional setting. These included clinical rotations, hands-on experiences with preschoolers, a treatment workshop, a brain dissection lab, a conference workshop, research, volunteering at a homeless shelter, and taking a class as an adult. These experiences were described as "focused," "hands-on," "intense," "meaningful," and offering an "invaluable perspective."

Numerous lessons were reported from these experiences. Some related to clinical management and improving understanding of course materials, such as "I learned two things that day, appreciation of client motivation and never give a kid something you have to take back without warning him first" and "It improved my knowledge of the course content and my skills of directing activities with children." Other experiences increased the participants' understanding of communication disorders, "[The instructor] taught us that day, better than anything we read in our text, what a fluency disorder can do to the person who has it." Along with increased clinical understanding, one participant wrote about how the experience changed her self-image.

> I loved constructing this assignment. I could be creative. I could do "more" than complete an assignment. I could make a difference by volunteering my time. I felt really good about myself as a learner, but more importantly as a human being. I felt connected to something larger than myself.

*C. Theme III: Responsive Instructors.*

As reported above, some of the comments made regarding student responsibility for learning included remarks about the instructor. Along with these were a number of other comments directed specifically at the instructor and his/her actions. Specific instructor activities that were mentioned included case studies, lectures, Socratic dialogue, stating and challenging assumptions, student created assignments/assessments, and two-minute quizzes. Across activities, the instructors were described as "accessible, encouraging, and responsive" and "reflective, even-handed, non-judgmental, but always organized, logical, goal-directed, serious and building momentum toward deep understanding and knowledge." They "provided guidance, engaged in conversation allowing me to openly share what I had learned (or thought I learned), and asked thoughtful questions to further my understanding and learning." The abilities to reason with students and make "great connections" were also listed. One participant described her instructor and his impact on the experience as,

> He was able to assess our learning based on our ability to apply and make sense of the abstract concepts within the framework of our institution. Feedback was seldom about right and wrong. Grades were irrelevant. It was about untangling concepts and we could keep trying until we had our own aha moment.

*D. Theme IV: Difficult Activities That Took Time.*

High levels of difficulty and significant amounts of time on task were described throughout the participants' significant learning experiences. Course features that were identified along with this theme included "a very large, very intimidating text," "extensive reading," and presentations that were scrutinized "until the presenter could justify the reason why that approach had, or had not, been successful." One participant summarized her experience by saying that the class "was like taking a trip through unknown and mysterious territories."

Along with descriptions of difficulty, some of the participants also commented on the amount of time they spent on the research. For example, one participant stated, "I spent hundreds of hours analyzing this data set. When I started I had never looked at a spectrogram. By the end I could almost identify the word just by looking at the screen." Another participant said, "I worked harder on this project than any other assignment in my undergraduate career. Yet, it was the most enjoyable project I completed across my studies." Finally, one participant described the benefits of having spent time on the task, while also expressing concerns about not having had as much time as she wanted: "[I] felt I could absorb anything that I read, if I could only dedicate the time to do it."

*E. Theme V: Connections to Previous Knowledge and Experiences.*

One of the common benefits described was being able to see connections between ideas and information within the experience and prior knowledge. Some of the connections were reported within a course, as shown by the statement "… each class period he would ask for us to relate about our previous cases." Other connections were made between clinical and classroom experiences, such as "It required me to apply my classroom-based knowledge of child language development to preschoolers." Still others, connected research and prior experiences, "I learned

to figure out relationships of existing knowledge to what I was doing and figure out how it fit or did not fit into my research."

*F. Theme VI: Direct Research Experience.*

As shown by the quotes in the previous themes, research was a common activity within the significant learning experiences. Three types of research were discussed. The first was research within a class; which included creating literature reviews of prior research, developing and refuting points of view based on the literature, and designing and implementing small experiments. The other two types of research experiences discussed were master's theses and doctoral dissertations. In both of these experiences the participants talked about developing research questions, analyzing data, having to "back-track and re-analyze data files when I learned something new." The stated benefits of these research experiences included increased understanding of the connection between research questions and methodology, increased abilities to absorb new information, and enhancing confidence for doing more research.

Some of the faculty members discussed how their instructor / research advisors facilitated their learning. For example, one faculty described an instructor who

> provided guidance, engaged in conversation allowing me to openly share what I had learned (or thought I learned), & asked thoughtful questions to further my understanding & learning. This experience stands out because I was actively engaged in my own learning.

Another faculty stated, "My mentor would ask probing questions and continually bring new research articles into the mix. After I became accustomed to this process, it was natural to enter into conversation with my mentor about my observations in the data set."

*G. Theme VII: Challenging Initial Ideas and Assumptions.*

Some of the significant learning experiences described contexts in which the participants' initial ideas and assumptions were directly challenged. One case reported on an instructor who

> was confrontational, debating each premise, sometimes causing frustration, but by the end of each class, we had the satisfaction of knowing that we had learned something meaningful, or we needed to dig a little deeper to find the answer. This approach probably would not be tolerated by students of today, but it rewarded good thinking and case building/management skills.

In another experience, the instructor would spend the first half of the class laying out a body of research that lead to a logical conclusion.

> After the break, he went through some same and new research that systematically pulled the rug out from under our beliefs…The process pulled us in. We reasoned along with the instructor. [It] remains the prime example of logical reason and learning in my background.

*H. Theme VIII: Rich in Content.*

Most of the comments that were made about content were related to developing connections between content and aspects of learning (see Themes V and VII). Two of the participants discussed content richness as central features of their significant learning experiences. In one

case, the participant had previously been a high achieving English literature major who perceived that field to be over-reliant on opinions. Her description of her first course in communication sciences and disorders included the following.

> There were facts, for the first time not just opinions (even learned opinions). There were answers, not just suppositions and points of view, and I was enamored in it…Ever since that experience, I crave content myself and I try to make sure there is a great amount of content in all the classes I teach.

The other participant who directly wrote about content was unhappy with a discussion-based course because the bulk of the knowledge did not come from the instructor, there was not as much content as she was expecting, and she felt that she needed more structure. She summarized her comments by saying,

> Now I think that course was probably problematic because as a service learner I needed more structure to help me grapple with how to handle content as well as more would-be content to handle. Content-to-handle was probably there in the class but seemed less clear to me because I expected it to come in a particular format; a detailed set of readings [or] a textbook. In fact on reflection, this sense of having rules for how to handle content and ideas about where to find that content are enduring concern for me as a teacher and for my students. What I learned the heck out of was about my own learning.

## III. Discussion.

This investigation examined potential connections between features of learner-centered teaching and past significant learning experiences. Current university faculty members and administrators in the discipline of communication sciences and disorders wrote short essays about a past college-level experience where they "really learned the heck out of something." The participants wrote about a variety of experiences, from individual classroom assignments, to whole courses, to research and clinical experiences.

A phenomenological analysis revealed eight recurring themes within these experiences: Student Responsibility for Learning, Learning Through Direct Experience or Examples, Responsive Instructors, Difficult Activities That Took Time, Connections to Previous Knowledge and Experiences, Direct Research Experiences, Challenging Initial Ideas and Assumptions, and Rich in Content.

### *A. Comparisons with Applied Features of Learner-centered Teaching.*

The primary research questions addressed in this study were: Are there commonalities between features of learner-centered teaching and past significant learning experiences? And, if so, what are they and what differences exist? These questions can now be answered by comparing the eight themes found in this study with the applied features of learner-centered outlined earlier.

There appear to be many direct connections between the applied methods of learner-centered teaching and the eight themes identified in this study. For example, the constructivist nature of learner-centered teaching promotes learning that is created by individuals and groups, as the result of their current knowledge/thoughts/beliefs interacting with new experiences. This intersection was demonstrated by the participants through the themes of Connections to Previous Knowledge and Experiences, and Challenging Initial Ideas and Assumptions. The themes of

Student Responsibility for Learning and Learning Through Direct Experience or Examples may also be reflective of constructivism, by allowing the participants to have had some ownership of their development and providing direct professional experiences.

Learner-centered teaching's goal of deep, performative, and proactive knowledge also appears to connect with the significant learning experiences. The participants' comments suggested that the many complexities that come with difficult activities, especially those that provide direct professional experiences and examples (such as research), prompted them to deeply consider the content and how it could be applied to a variety of problems. Likewise, the challenges to their initial ideas and assumptions made them "dig a little deeper to find the answer."

Weimer's (2002) list of changes that occur when an instructor moves from a traditional to learner-centered approach provided another framework for comparison. The themes identified in this study appear to directly reflect almost all of these changes. For example, learner-centered teaching changes the balance of power from the teacher only to a shared responsibility among the teacher and students. The theme of student responsibility for learning addressed this through comments about how the participants were allowed to develop and/or select specific assignments and assessments, research topics, and advising. The participants reported that having some say in their classroom experience was empowering. It allowed them to think and do more, and helped them to feel what it is like to be a professional. Likewise, statements within the themes of Responsive Instructors and learning through direct experience described faculty who included student input as part of the design of the educational experiences.

The function of content also appeared to have a direct connection with the significant learning experiences, although there were important differences within comments made regarding content. Most of these statements supported learner-centered teaching's use of content as a tool for developing the students' knowledge, skills, and self-awareness. These centered around the experiences enlightening prior content knowledge, causing reconsideration of prior content knowledge, and/or requiring the acquisition of new content knowledge. Similar sentiments were also expressed in the themes of connections to previous knowledge and experiences, and challenging initial ideas and assumptions. In other words, the participants showed an appreciation for deep learning, through the application of content to new and varied situations. Two participants wrote specifically about experiences in which they had wanted more content. Even in these cases, their comments were directed towards content that was relevant and would have helped them learn more, as opposed to wanting more content in order to learn more facts.

The role of the instructor in learner-centered teaching is to facilitate the students' development of knowledge and skills. Comments within the theme of Responsive Instructors demonstrated this feature. The participants, for example, described teachers who "provided guidance", allowed open sharing of student learning, and focused on student conceptual development.

Along with changing the balance of power, learner-centered teaching creates a greater sense of student responsibility for learning. Although observed across many of the themes, this feature was demonstrated the most prominently within Student Responsibility for Learning. The participants described themselves as being actively involved in their experiences. They described their learning as intentional and purposeful. They saw this as helping them to go beyond what they would have before and to become independent learners.

Weimer's (2002) fifth change with learner-centered teaching, the purpose and process of evaluation, was the feature of her list that was not evident as a theme. Only two participants made direct comments about the purpose and process of evaluation. One reported that she was given the freedom in a course on manual communication (e.g., sign language and communication boards) to design her own final exam experience. The other described her instructor's contribution to her learning by stating that, "Feedback, was seldom about right and wrong. Grades were irrelevant. It was about untangling concepts and we could keep trying until we had our own aha moment." It is unclear why more of the participants did not comment on the evaluation components of their experiences. It may be that the evaluation part of their experience was a) not important to them because they were focused on their own learning, b) typical of other courses, c) unmemorable, or d) simply not a part of the experience that they thought to include in their descriptions (as their instructions did not identify specific aspects of the features to write about).

Taken together, the themes identified in this study support a connection between applied features of learner-centered teaching and past significant learning experiences. These connections, however, should be considered as preliminary for a number of reasons. First, the data were based on the participants' recollections of events that occurred multiple years in the past (in some cases decades later). Their memories may have been influenced by their other experiences that have occurred since then, including encounters with scholarly teaching. Second, the participants had self-selected to attend the conference session on teaching and learning. Although features of learner-centered teaching had not yet been addressed in the session, these may have been faculty who were primed to think in this direction. Third, because all of the participants were within the same discipline, they may have been prone to think in terms of learner-centered teaching. This is not likely to be the case, however, because learner-centered teaching is not the dominant teaching method being used in the discipline (Brackenbury, Folkins, & Ginsberg, 2011) and the discipline's research on it is developing, but not substantial. In addition, potential connections between client-centered therapy and learner-centered teaching had not been previously discussed (either within the published literature or previous conference presentations), suggesting that the participants were not highly predisposed to make these associations. Finally, the author and research assistants were all predisposed, to varying degrees, to consider connections between past significant learning experiences and learner-centered teaching. Given the procedures taken to address issues of credibility and transferability, however, these potential confounders do not appear to have had a significant negative influence on the results of this study.

*B. Implications and Conclusions.*

The results of this study demonstrate that features of learner-centered teaching can be directly connected with past significant learning experiences. The connections identified between learner-centered teaching and the significant learning experiences suggest that these principles of learning have been around for a long time. Although these connections are not entirely surprising, as the cognitive underpinnings of learner-centered teaching have been around for numerous decades, they suggest that learner-centered teaching can result in learning that lasts for many years after the experience. However, not all applied features of learner-centered teaching were connected to past significant learning experiences (most notable the purpose and of

evaluation). Further research into learner-centered teaching and past significant learning experiences should help to specify these commonalities and differences.

The observed connections between learner-centered teaching and past significant learning experiences may prove to be helpful for faculty who are new to this type of teaching. Rather than considering learner-centered teaching as a new, challenging way of educating students, their first steps towards being more learner-centered could be to consider their own past significant learning experiences. They could then identify the aspects that made these experiences so powerful and then look within the scholarship of teaching and learning literature to find evidence of connected features. It would also provide a familiar platform from which they can build learner-centered teaching into their own instruction. Additionally, the types of activities and instructors that the faculty described in their experiences could be used to inspire teachers to develop similar experiences and teaching personas.

A similar case can be made for using the data and results from this study to assist students who are new to learner-centered teaching. Many students have difficultly when they first encounter learner-centered teaching, especially if they had not experienced it before the college level (e.g., Doyle, 2008). It may be beneficial for learner-centered instructors to have their students do the same reflection that the participants in this study completed. Discussing the students' experiences and connecting features of them with the methods of the course may help facilitate their understanding of why the class was designed differently than their previous courses.

The results of this study suggest that applied principles of learner-centered teaching can be connected to significant learning experiences. Although there are no guarantees that the learner-centered experiences that faculty provide will be later identified by their students as significant, educators can use the idea of creating significant learning experiences as a guiding principle within and across their courses. The connections demonstrated here suggest that learner-centered teaching can provide both a framework and specific methods for doing so.

## References

Bain, K. (2004). *What the Best College Teachers Do*. Cambridge, MA: Harvard University Press.

Barr, R.B., & Tagg, J. (1993). A new paradigm for undergraduate education. *Change*, 27, 12-25.

Black, K.A. (1993). What to do when you stop lecturing, *Journal of Chemical Education, 70,* 140-144.

Black P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability, 21*, 5-31.

Bloom, B. S. (Ed.) (1956). *Taxonomy of Educational Objectives: Book 1 Cognitive Domain*. New York: Longman.

Brackenbury, T. (2011, November). How past learning experiences compare with learner-centered teaching. Poster presented at the annual convention of the American Speech-Language-Hearing Association, San Diego, CA.

Brackenbury, T., Folkins, J., & Ginsberg, S. (2011, November). The signature pedagogy of communication sciences and disorders. Presentation at the annual convention of the American Speech-Language-Hearing Association, San Diego, CA.

Brackenbury, T., & Shaughnessy, P. (2010, April). Everything I Need to Know About Teaching, I Learned from Speech and Language Therapy. Paper presented at the CAPCSD Convention, San Antonio, TX.

Creswell, J.W. (1998). *Qualitative Inquiry and Research Design*. Thousand Oaks, CA: Sage.

Creswell, J.W. (2011). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed.). Upper Saddle River, NJ: Pearson.

Doolittle, P.E., & Siudzinski, R.A. (2010). Recommended syllabus components: What do higher education faculty include in their syllabi? *Journal on Excellence in College Teaching, 20,* 29-61.

Doyle, T. (2008). *Helping Students Learn in a Learner-centered Environment*. Sterling, VA: Stylus.

Eberly, M.B., Newton, S.E., & Wiggins, R. A. (2001). The syllabus as a tool for student-centered learning. *Journal of General Education, 50,* 56-74.

Fink, L.D. (2008). *Creating Significant Learning Experiences*. San Francisco, CA: Jossey-Bass.

Garavalia, L.S., Hummel, J.H., Wiley, L.P., & Huitt, W.G. (1999). Constructing the course syllabus: Faculty and student perceptions of important syllabus components. *Journal on Excellence in College Teaching, 10*, 5-21.

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

Hinchliffe, G. (2011). What is a significant educational experience? *Journal of Philosophy of Education, 45*, 417-431.

King, A. (1993). From sage on the stage to guide on the side. *College Teaching, 41*, 30-35.

Lambert, N.M., & McCombs, B.L. (1997). *How Students Learn.* Washington, DC: American Psychological Association.

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

Marzano, R.J. & Kendall, J.S. (2007). *The New Taxonomy of Educational Objectives (2ond ed.)*. Thousand Oaks, CA: Corwin Press.

Maxwell, J.A. (2005). *Qualitative Research Design: An Interactive Approach (2nd ed.).* Thousand Oaks, CA: Sage Publications.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative Data Analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Moustakas, C. (1994). *Phenomenological Research Methods*. Thousand Oaks, CA: Sage Publications.

O'Brien, J.G., Millis, B.J., & Cohen, M.W. (2008). *The Course Syllabus* (2nd ed.). San Francisco, CA: Jossey-Bass.

Perkins, D. (2008). Beyond understanding. In R. Land, J. H. F. Meyer, & J. Smith (Eds.) *Threshold Concepts Within the Disciplines*, 3-19. Rotterdam: Sense Publishers.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 223-231.

Rhem, J. (1995). Deep/Surface approaches to learning: An introduction. *The National Teaching and Learning Forum, 5,* 1-5.

Richardson, V. (1997). *Constructivist Teacher Education.* New York: Routledge.

Rushton, A. (2005). Formative assessment: A key to deep learning? *Medical Teacher, 27,* 509-513.

Savery, J.R., & Duffy, T.M. (2001). *Problem Based Learning: An Instructional Model and its Constructivist Framework* (Tech. Rep. No. 16-01). Indiana University, Center for Research on Learning and Technology.

Stoll, C. (2008). *18 Minutes With an Agile Mind*. Presentation at TED2006. Retrieved March 30, 2012, from http://blog.ted.com/2008/03/27/clifford_stoll/

Weimer, M. (2002). *Learner-centered Teaching*. San Francisco, CA: Jossey-Bass.

Weimer, M. (2006). *Enhancing Scholarly Work on Teaching and Learning*. San Francisco: Jossey-Bass.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45,* 477-501.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 29 – 42.

# Application essays as an effective tool for assessing instruction in the basic Communication course: A follow-up study

**Joseph P. Mazer[1], Cheri J. Simonds[2], and Stephen K. Hunt[3]**

*Abstract: The assessment of student learning in general education courses is of critical importance in higher education. This study examines the utility of a writing assignment (application essays) in a basic communication course as an effective assessment tool. The authors conducted a content analysis of student portfolios to determine the extent to which application essays provide evidence of student learning in the basic course. The present study extends the findings from recent assessment efforts (Jones, Simonds, & Hunt, 2005) to explore types of mass media events students address in application essays and assess the revisions made to the assignment based on findings from Jones et al. (2005). Results reveal (a) the various communication events that students write about in application essays, (b) the communication concepts that students address, (c) that students typically, but not always, make appropriate connections when they write application essays, and (d) after revising the assignment based upon data from recent assessment efforts, more students made appropriate connections between the communication event and concept. Implications for classroom pedagogy and course management are discussed.*

*Keywords: assessment, application essays, basic communication course*

Accurate assessment of students' comprehension of course concepts is an essential component of the instructional process. Over the years, educators and scholars from a variety of disciplines have introduced and implemented diverse assessment methods in an attempt to produce a reliable procedure to assess instruction and student learning (Fallon, Hammons, Brown, & Wann, 1997; Sforzo, 2005; Sircar, Fetzer, Patterson, & McKee, 2009; Stefanou, Hood, & Stefanou, 2001). Recently, portfolio assessment has been used and evaluated through a wide range of disciplines including language arts (Black, Daiker, Sommers, & Stygall, 1994; Crouse, 1994; Gill, 1993; Reyes, 1991; Voth & Moore, 1997), math and sciences (Barrow, 1994; Chapman, 1996; Slater, 1995), and education (Farris & Fuller, 1996; Gipe & Richards, 1992; Patzer & Pettegrew, 1996; Vizyad, 1994).

The need to accurately assess classroom instruction and student learning was required of us as basic communication course administrators of a large (about 1,500 students a semester) multi-section, general education course focused mainly on public speaking skills. As part of our initial efforts in this arena, we examined whether or not the use of student portfolios in our basic communication course was an effective, authentic tool for assessment and concluded that student portfolios are an effective mechanism for gathering data with respect to student accomplishment of course goals and student learning (Hunt, Simonds, & Hinchliffe, 2000).

Emerging from this initial research, as a component worthy of further examination, is the application essay assignment. Application essays require students to apply theoretical concepts

[1] Department of Communication Studies, Clemson University, 407 Strode Tower, Clemson, SC, jmazer@clemson.edu
[2] School of Communication, Illinois State University, Campus Box 4480, Normal, IL 61790, cjsimon@ilstu.edu
[3] School of Communication, Illinois State University, Campus Box 4480, Normal, IL 61790, skhunt2@ilstu.edu

they learn in the basic communication course to communication events they experience in the "real world." This assignment asks students to write a one-page paper, describe a particular communication event in their life, link it to a communication concept from the course material, and analyze how the event is related to the communication concept (Hunt et al., 2000). Typically, students are instructed to complete a series (three to five) of these theoretical application essays—focusing on a variety of course concepts—over the course of a semester.

Arguments have been advanced which support the use of application essay assignments as an effective tool for assessing instruction in the basic course. First, the application essays from our course have been used previously to assess critical thinking. Hunt et al. (2000) found that the application essay promotes student classroom participation by helping them apply concepts learned in the classroom to their own personal experiences. Hunt et al. (2000) conclude that application essays "allow teachers to assess each student's level of critical thinking by judging the description of the application essay, the link to the specified concept, and the analysis of each communication concept" (p. 93). It is reasonable to assume that if the application essay can be used to assess student critical thinking, it may also serve as an effective measure of student learning.

In a recent study (Jones, Simonds, & Hunt, 2005), we extended our initial assessment efforts (Hunt et al., 2000) to explore and solidify both the strengths and limitations of the application essay as an effective tool for assessing student learning in the basic communication course. In this recent assessment effort, we designed a coding instrument to record communication events (e.g., mass media, public speakers, one-on-one encounters), concepts (e.g., communication process, ethical communication, nonverbal communication), and appropriate or inappropriate analysis of the event and concept in 369 application essays. In our analysis of communication events, we found that the vast majority of application essays ($n$ = 155, 42%) analyzed communication in the mass media. This was followed by one-on-one encounters ($n$ = 61, 16.5%), public speakers ($n$ = 34, 10.6%), classroom examples ($n$ = 35, 9.5%), and the students' own speeches ($n$ = 31, 8.4%).

In order to simplify our analysis of communication concepts, we collapsed student responses into the five major units of the course: immersion, message clarity, message responsiveness, persuasion, and synthesis. This allowed us to identify points in the semester where students wrote the majority of application essays. Students wrote 123 application essays (33.3% of the total) addressing topics (e.g., communication process, critical thinking, and ethical communication) in the immersion unit—a unit in which students begin to identify critical elements present in most communication situations and begin to establish criteria for recognizing communication competence.

In the message clarity unit, students practice creating messages, develop skills in producing concise, well-formed, and listener-adapted messages, and practice skills in listening for the main points of messages to separate the content of the message from biases of the speaker. Students wrote 121 application essays (32.8% of the total) addressing topics (e.g., audience analysis, language, delivery) in this unit.

Students wrote 35 application essays (9.5% of the total) addressing topics (e.g., group communication, cultural diversity, listening, and managing conflict) in the message responsiveness unit—which emphasizes the related functions of perspective-taking, empathy, seeking and providing comfort and social support, managing conflict, and moving competently through the various group roles that facilitate the decision-making process.

The persuasion unit is designed to help students understand the persuasive process, both as speakers who wish to influence others and as listeners who wish to resist persuasive attempts when necessary. Students wrote 32 application essays (8.7% of the total) addressing topics (e.g., ethos, pathos, logos) in this unit. Students wrote two application essays (.5% of the total) addressing topics in the synthesis unit, a segment of the course that provides students with the opportunity to synthesize and reflect upon what they have learned and experienced throughout the semester.

We found that students typically, but not always, made appropriate connections between the communication event and concept in their application essays. We found that some students established no clear link in 22.8% ($n$ = 84) of the application essays. More importantly, our analysis revealed that some students advanced an incorrect/inappropriate link between the communication event and course concept 6.5% ($n$ = 24) of the time. Indeed, this was a troubling trend. We argued that one possible reason this trend began to develop was due to the application essay assignment instructions being unclear and inconsistent for the students throughout the semester. To rectify this apparent problem, we made significant changes to the instructions for the application essay (See Appendix A).

The students in this sample focused the bulk of their attention on the first two units of the course (immersion and message clarity). This finding was troubling for several reasons. Initially, less than 10% of the application essays dealt with the concepts presented in the message responsiveness unit. This was problematic because it is in this unit that students read about, discuss, and develop an understanding of cultural diversity, listening, and conflict management. The fact that fewer students wrote application essays on these topics did not necessarily mean that they did not develop an appreciation for these topics; however, it did suggest that many students were possibly missing a valuable opportunity to extend the learning occurring in the classroom. This pattern also held for the persuasion unit. Again, it appeared that students were not taking full advantage of the application essay as a vehicle for extending learning opportunities regarding persuasive communication. To address this issue, we revised the instructions for the application essay assignment (See Appendix A).

We argued that because so many application essays were concentrated at the start of the semester instructors might not have been doing enough to reinforce to students that they should be writing application essays. Specifically, instructors may have been focusing on the importance of the application essay assignment early in the semester and then turned their attention to other course assignments such as the group and persuasive speeches. This finding suggested a need to better train instructors in terms of having the students write application essays on all of the major units of the course. As course directors, we encouraged instructors to make specific application essay assignments during these units and explain the assignments using the instruction sheet (See Appendix A).

The present study extends our most recent assessment efforts (Jones et al., 2005) to explore the types of mass media events students address in application essays. We also assess the revised assignment instruction sheet that encourages students to complete application essay assignments throughout the course units and explain how to go about making stronger connections between course concepts and communication events. The following research questions guided our efforts in this study:

$RQ_1$: What communication events do students write about in the application essays?
$RQ_2$: What types of mass media do students address in the application essays?
$RQ_3$: What communication concepts are students addressing within the application

essays?

$RQ_4$: Are students making appropriate connections between communication events and communication concepts in the basic communication course?

## I. Method.

### *A. Sample.*

We extracted 115 application essays from a set of student portfolios. The portfolios for this data set were collected at the end of the fall semester at a large midwestern university. We extracted from the portfolios a systematic random sample of student application essays for each instructor. Students voluntarily submitted their portfolios and signed informed consent forms.

### *B. Category Definitions.*

To answer the three research questions, we utilized the same coding instrument from Jones et al. (2005). This coding instrument was used to record communication events, concepts, and appropriate or inappropriate link analysis. We scanned each essay and generated a list of every possible category for communication events and concepts. For example, communication events included mass media (internet, magazines, movies, radio, and television), classroom examples, live entertainment, one-on-one encounters, public speakers, special events speeches, and self-analysis. Concepts included, but were not limited to audience analysis, communication apprehension, communication process, critical thinking, ethical communication, persuasion, small group processes, nonverbal communication, and visual aids.

### *C. Procedures.*

We trained three coders who were graduate teaching assistants for the course. Each coder then independently analyzed 12 essays from the sample to assess intercoder reliability. Based on Pearson's R formula, overall reliabilities were .1. Upon completion of the coding for the entire sample, we calculated frequency counts and percentages to answer our research questions.

## II. Results.

The purpose of this study was to extend our most recent assessment efforts (Jones et al., 2005) to explore the use of application essays as a formative assessment tool in the basic communication course. By using frequency distributions, we determined (a) the various types of communication situations that students in this study wrote about in their application essay assignment, (b) the concepts the students applied to these communication situations, (c) that students in this study typically, but not always, made appropriate connections when they wrote application essays, and (d) after revising the assignment based upon data from our most recent assessment efforts, more students made appropriate connections between the communication event and concept in the present study.

The first research question examined the communication situations that students wrote about in their essay assignments. Specifically, we were interested in what situations or events were commonly written about in the essay assignments (see Table 1). The vast majority of essay

assignments ($n$ = 36, 31.3%) analyzed communication students observed in the mass media. This was followed by one-on-one encounters ($n$ = 33, 28.7%), public speaker ($n$ = 11, 9.6%), and the student's own speech ($n$ = 7, 6.1%). We then extended the findings from Jones et al. (2005) by expanding on the mass media application essays (see Table 2). The second research question asked what types of mass media ($n$ = 36, 31.3%) students addressed in the application essays. Of this 31.3%, we were interested in examining the frequency of each type of media. The vast majority of essay assignments ($n$ = 24, 20.9%) analyzed a television show, followed by movies ($n$ = 10, 8.7%) and magazines ($n$ = 1, .9%).

**Table 1.** Application essay communication events.

| | Frequency | Percent |
|---|---|---|
| Classroom Examples | 6 | 5.2 |
| Live Entertainment | 4 | 3.5 |
| Mass Media | 36 | 31.3 |
| Not Applicable | 1 | .9 |
| One-On-One | 33 | 28.7 |
| Public Speaker | 11 | 9.6 |
| Small Group | 6 | 5.2 |
| Student's Own Speech | 7 | 6.1 |
| Other | 11 | 9.6 |
| Total | 115 | 100.0 |

**Table 2.** Type of media used in application essays.

| | Frequency | Percent |
|---|---|---|
| Not Present | 79 | 68.7 |
| Internet | 1 | .9 |
| Magazine | 1 | .9 |
| Movies | 10 | 8.7 |
| Television | 24 | 20.9 |
| Total | 115 | 100.0 |

Research question three addressed the communication concepts students wrote about in their application essay assignments. We identified 29 different concepts students utilized throughout the course of the semester (see Table 3) and alphabetized the list to simplify the process for the coders. Analysis revealed that students tended to write more application essays in certain units in the course, $\chi^2(3) = 27.27$, $p < .05$. Students wrote 48 application essays (41.7% of the total) in the immersion unit on topics such as the communication process ($n$ = 10, 8.7%), perception ($n$ = 9, 7.8%), and critical thinking ($n$ = 7, 6.1%); 16 application essays (13.9% of the total) in the message clarity unit on delivery ($n$ = 6, 5.2%) and visual aids ($n$ = 4, 3.5%); 30 application essays (26.1% of the total) in the message responsiveness unit on topics such as small group roles ($n$ = 10, 8.7%) and stereotypes ($n$ = 6, 5.2%); and 20 application essays (17.4% of the total) in the persuasion unit on audience centeredness ($n$ = 8, 7.0%), pathos ($n$ = 4, 3.5%), ethos ($n$ = 3, 2.6%), and logos ($n$ = 2, 1.7%).

**Table 3.** Concepts used in application essays.

| | Frequency | Percent |
|---|---|---|
| Immersion Unit | | |
| Communication Process | 10 | 8.7 |
| Perception | 9 | 7.8 |
| Critical Thinking | 7 | 6.1 |
| Ethical Communication | 6 | 5.2 |
| Credibility | 4 | 3.5 |
| Communication Apprehension | 3 | 2.6 |
| Frame of Reference | 2 | 1.7 |
| Self-Concept | 2 | 1.7 |
| Self-Image | 2 | 1.7 |
| Nonverbal Communication | 1 | .9 |
| Plagiarism | 1 | .9 |
| Self-Esteem | 1 | .9 |
| Subtotal | 48 | 41.7 |
| | | |
| Message Clarity Unit | | |
| Delivery | 6 | 5.2 |
| Visual Aids | 4 | 3.5 |
| Evidence | 2 | 1.7 |
| Language | 2 | 1.7 |
| Problem-Solution Order | 1 | .9 |
| Main Points | 1 | .9 |
| Subtotal | 16 | 13.9 |
| | | |
| Message Responsiveness Unit | | |
| Small Group Roles | 10 | 8.7 |
| Small Group (general) | 7 | 6.1 |
| Stereotypes | 6 | 5.2 |
| Listening | 4 | 3.5 |
| Cultural Diversity | 3 | 2.6 |
| Subtotal | 30 | 26.1 |
| | | |
| Persuasion Unit | | |
| Audience Centeredness | 8 | 7.0 |
| Pathos | 4 | 3.5 |
| Ethos | 3 | 2.6 |
| Monroe's Motivated Sequence | 2 | 1.7 |
| Logos | 2 | 1.7 |
| Persuasion | 1 | .9 |
| Subtotal | 20 | 17.4 |
| | | |
| Other | 1 | .9 |
| Total | 115 | 100.0 |

The final research question examined whether students were making appropriate links in their application essay assignments. Coders were asked to make these determinations given their knowledge of the course terms. We determined that any essay with a score over 80% did an adequate job linking the event to the concept. In our course, an 80% is a B grade (above average). We found that 63% ($n$ = 72) of the participants in the present study made a direct link between the communication events and communication concepts. Strong and direct links are illustrated in the following application essay excerpts (see italics for language coded as a direct link):

*Example 1:*

We decided to play a game that Jason's little brother had gotten for Christmas: *Think Blot*. To summarize, the game incorporates inkblots that you would see if you went to see a psychologist. Each player has a minute to examine a picture and write down every possible person, activity, place, or item that they can see in the inkblot shown. After a minute, each player shares something he or she was able to see in the picture; if another person saw the same thing, each receives a point. If only one player sees an object, he or she has to get the other players to see it, and then he or she receives two points.

We all played the same game, the four of us using our imaginations to see as many different objects as possible in the same picture. This game relates to our discussion in class on Friday, January 25, concerning perception, and more specifically subjective perception. As stated in the text, "perception is the process of becoming aware of objects and events from the senses" (Lucas, 2004, p. AA-6). To add, perception is subjective because we interpret what we sense and make it our own; in turn, adding and subtracting what we see, hear, smell, and touch. More specifically, subjective perception is giving our own, uniquely constructed meaning to stimuli. *My experience playing Think Blot is an example of subjective perception, because we were each viewing the same picture, yet had our own interpretations of it. We did not just see a generic interpretation of the inkblot and say that we saw a spot on a page, but rather we also tied in the elements of closure and figure and ground to see figures. This could also make Think Blot interpretative stimuli, because it uses a blend of internal states and external stimuli to interpret the inkblot. The differences in our perceptions also deal with our backgrounds, experiences, and also our states of mind. For example, while we were playing, I kept seeing food (i.e. an apple), which could have been because I was hungry at the time. Also, I was the only girl playing, so the males would see more masculine objects such as guns and knives, and I would see them as a golf club or candlestick.*

*Example 2:*

In the middle of watching *Boston Public*, one of my favorite television shows, a commercial for The Gap came on the screen. Sarah Jessica Parker, an actress famous for her role in the television show *Sex and the City*, appeared in the commercial. She was dressed very nicely in khaki capris, pink tee shirt, and matching shoes. Her hair was curled, and when she smiled her teeth literally sparkled. In this commercial Sarah Jessica Parker was singing a new interpretation of the Broadway song "Enjoy Being a Girl."

> While she was singing, Sarah was running up and down the aisles of The Gap store, picking up every item she could and dropping it into her cart. In an instant, many other girls joined her in the store, and started singing and shopping as well. The commercial finally ended with Sarah saying that every girl should shop at Gap.
>
> This commercial frightened me because it was so corny. After watching the advertisement, I noticed that the commercial had many different fallacies in it. The first fallacy I noticed was the bandwagon fallacy. The bandwagon fallacy "invites you to join the group and do something because everyone is doing it" (Lucas, 2004, p. FF-84). *The Gap used this fallacy by including many different girls at the end of the advertisement. By adding in so many girls at the end of the commercial, a teenage girl watching the advertisement may be compelled to drive to The Gap and buy some clothes. She would feel the need to do that because all the cool girls on the commercial were doing the same thing.* In addition, I also noticed this commercial had the appeal to authority fallacy. The appeal to authority fallacy "occurs when a person offers information that is outside his or her area of expertise" (Lucas, 2004, p. FF-85). *The Gap commercial used this fallacy by having Sarah Jessica Parker tell every girl to shop at Gap. Sarah Jessica Parker is an actress, not a fashion consultant, so it is ridiculous that she is telling other people what clothes to buy.*

Application essay assignments with a score less than 80% and all essays without scores were analyzed as to why an appropriate link was not made. That data indicates that students in this study typically, but not always, made appropriate connections between the communication situation and the course concept (see Table 4). We found that of the 49 students who scored 79% or lower, some students ($n$ = 24, 21%) established no clear link between the communication situation and concept (see Table 5). This finding highlights an important improvement from the Jones et al. (2005) study, which revealed that 25% failed to link the communication event to the course concept. A weak link is illustrated in the following application essay excerpt (see italics for language coded as a weak link):

> This past weekend I went home for my sister's sixteenth birthday party. At the party, my three year old cousin, Sam, was just discovering that people have last names. He went from person to person asking, "What is your last name?" Then he would stand there and repeat it and say, "neat." Finally my grandma asked Sam what his last name was. He said, "My name is Samuel Boeckman-get-over-here-right-now!" Everyone thought this was hilarious, but Sam was confused. After all, your "last" name was the one someone said last when they addressed you, right?
>
> Actually, this is not true, but to three year old Sam, this was what he understood the term "last name" to mean. Last Tuesday, September 29th, we had a class discussion about the meaning of words and discussed how meanings were in the minds of people. Sam's confusion is a perfect example of what we discussed. *In class, we talked about how words have different meanings for different people. Since Sam is only three, the word "last" to him means something different than it would to an older person. Based on Sam's experience as a little boy, the term "last" has different meaning.*

**Table 4.** Application essay scores.

| | Frequency | Percent |
|---|---|---|
| 60 | 3 | 2.6 |
| 65 | 2 | 1.7 |
| 68 | 1 | .9 |
| 70 | 7 | 6.1 |
| 73 | 2 | 1.7 |
| 75 | 4 | 3.5 |
| 78 | 1 | .9 |
| 80 | 13 | 11.3 |
| 83 | 1 | .9 |
| 85 | 6 | 5.2 |
| 88 | 2 | 1.7 |
| 90 | 16 | 13.9 |
| 93 | 1 | .9 |
| 95 | 6 | 5.2 |
| 98 | 2 | 1.7 |
| 100 | 25 | 21.7 |
| No score | 29 | 20.0 |
| Total | 115 | 100.0 |

**Table 5.** Reasons for poor application of concepts in application essays.

| | Frequency | Percent |
|---|---|---|
| accurate link | 9 | 18.4 |
| poor link | 24 | 49.0 |
| format/grammar | 5 | 10.2 |
| citation/reference page | 6 | 12.2 |
| instructor preference | 1 | 2.0 |
| missing definition of term | 4 | 8.2 |
| Total | 49 | 100.0 |

Additional low scores (i.e., scores less than 80%) reflected other problems such as: an accurate link but no score to determine this ($n$ = 9, 18.4%), incorrect citations/references ($n$ = 6, 12.2%), formatting/grammar errors ($n$ = 5, 10.2%), inaccurately defining the communication concept ($n$ = 4, 8.2%), and issues related to instructor preference ($n$ = 1, 2%).

## III. Discussion.

Since our initial investigation of student portfolios in the basic communication course (Hunt et al., 2000), we were able to analyze the effectiveness of various course assignments, including application essays (Jones et al., 2005). In the present study, we were able to determine the various communication events that students wrote about in application essays, the concepts that students addressed, and that students typically, but not always, made appropriate connections when they wrote application essays. Although students wrote about a wide range of communication events, the majority of them addressed communication in the mass media. This finding is not surprising given the pervasiveness and accessibility of the mass media. Although

this finding is not inherently problematic, it does raise a level of concern in terms of context richness. In other words, we feel that students should grow to develop an appreciation for a wide range of communication contexts. If they devote too much attention to the mass media, they may not develop a full appreciation for other contexts. On the other hand, we found that students often critically reflect on their mass media experiences. This can be interpreted in a positive light given course goals that ultimately ask students to be more critical consumers of the information they are exposed to. This suggests that instructors should consider monitoring the contexts that students write about in the basic course to ensure that they capitalize on the potential for this assignment to have students critically reflect on multiple contexts of communication.

Our recent assessment efforts (Jones et al., 2005) revealed that approximately 18.2% of the application essays focused on concepts from the message responsiveness and persuasion units. We found a notable improvement in the present study where more students wrote about communication concepts from the message responsiveness (26.1%) and persuasion (17.4%) units (see Table 3 for a detailed list). This finding suggests that our revisions to the application essay assignment—including our suggestions during instructor training—were successful in encouraging students to apply communication concepts from all units in the course to real life experiences.

One limitation of the application essay is that it can be difficult to compare students across multiple sections of a large general education course. For example, the application essays analyzed here were often worth different points and even the details of the assignment differed by instructor. However, this too is valuable assessment information for course directors and has been utilized to build a more standardized curriculum.

The findings from Jones et al. (2005) revealed a troubling trend that was occurring with the students—that is, approximately 25% of the students failed to make a direct link between communication events and course concepts. One possible reason this trend began to develop may be due to the application essay assignment instructions being unclear and inconsistent for the students throughout the semester. To rectify this apparent problem, we made significant changes to the instructions for the application essay (See Appendix A). In the present study, participants were exposed to the following revisions. First, students were provided a detailed description and model examples of an application essay for them to reference while completing their own assignments. Additionally, we trained instructors to describe the assignment in a thorough and consistent manner and to make reference to possible application essays throughout the semester. As a result of our revisions to the assignment based upon data from Jones et al. (2005), 63% ($n$ = 72) of the participants made a direct link between the communication events and communication concepts. While the present study was not conducted on such a large scale as Jones et al. (2005), a majority of the 115 essays we analyzed in the present study ($n$ = 72, 63%) accurately identified and linked the communication concept to the communication event. In fact, the data revealed that only 21% percent of the sample ($n$ = 24), as opposed to 25% in Jones et al. (2005), did not make accurate links between communication events and communication concepts. This finding indicates that the revisions we made to the application essay assignment appear to explain the assignment more effectively for students. In addition, by revising course instructor training efforts we have rectified potential uncertainties that teachers and students may have about the application essay assignment.

Finally, instructors might consider implementing written student application essays in their general education courses. If application essays are effective assessment tools of student learning within a basic communication course that focuses primarily on public speaking, then

application essays can be adapted successfully into any type of general education course. Given that general education courses often provide students with concepts and definitions related to the specific course content area, instructors can ask students to apply those concepts and definitions to life examples through the application essay assignment. For example, the instructor of a general education psychology course might ask students to apply concepts related to different personality types to real life examples. These experiences can provide students with opportunities to synthesize and apply course content to experiences outside of the classroom.

The present study is not without limitations. We have been able to demonstrate that a majority of the students are making accurate connections between course material and the communication events they are exposed to outside of the classroom. However, our analysis of these application essays stops short of demonstrating actual student learning. It could be that students make these connections without deeply understanding the course material. Also, it may be that students have different "gestation" periods for fully realizing the applicability of the communication concepts they are exposed to in the basic course. In fact, we often have students come back to us, in some cases several semesters after taking the course, to express that they finally understand how useful the application essay assignment was in helping them see the relevance of the communication theories and principles discussed in the basic course. This suggests that we as basic course directors should employ a multi-method, multi-tool approach to the assessment of course goals and student learning. A comprehensive assessment plan should include psychometric measures as well as more authentic indicators of student performance like student portfolios. Indeed, in the present study, a smaller percentage of students (21%) failed to make strong connections between course concepts and life experiences compared to students in the prior study (25%). While this difference may appear marginal, the noted improvement in students' ability to make direct connections between course content and life experiences highlights the importance of longitudinal assessment in multi-section courses. Longitudinal data can garner information from students who—after completing the basic course—begin to more fully realize the applicability of communication concepts. By analyzing all of this data holistically, we can begin to capture a better image of student performance and better assess the short and long term implications of the application essay assignment. In many ways, the application essay assignment discussed here is just one part of the overall assessment picture.

Overall, the present study provided insights into the assessment of our course. The research indicated a need for change in several areas of course delivery as illustrated in Jones et al. (2005). Initially, the fact that so many students had difficulty making accurate links in the application essays indicated a need to clarify the assignment. As a result, we developed an assignment sheet clearly outlining the details of the assignment with model examples for students to follow (see Appendix A) as analyzed in the present study. The findings from Jones et al. (2005) indicated a need for us to reinforce the importance of the assignment to all instructors to ensure that students have the opportunity to write about course concepts in each of the major units of the course. In a similar vein, the data indicated a pressing need for us to direct the instructors to require students to write about the content delivered in the message responsiveness and persuasion units. This information has been incorporated into our training of course instructors. Our findings from the present study indicate that our revisions to course delivery were effective in helping students make direct linkages between course content and real life situations. Notably, recent advances in communication technology can provide students with a plethora of opportunities to demonstrate their understanding of course concepts. For instance,

students might link course content to their communication experiences through social media such as Facebook, Twitter, and other social networking venues.

Finally, this study was an exercise in the scholarship of teaching and learning. The data we analyzed were largely formative. In other words, we gleaned insights about what the students were learning in the course, but we also learned about areas where we could improve course instruction. Future research might further explore how the application essays are linked to student achievement and learning outcomes. Do students who make strong connections between course concepts and communication events (in the application essay) perform better on course exams? Obviously, this kind of assignment/assessment is not unique to the basic course. Teachers in other courses and disciplines can adapt both the assignment and the procedures for analyzing the data to their own unique requirements. Procedures such as those outlined in the current study are particularly relevant given the pressure educators are under to conduct assessment. Student application essays have wide applicability as a method not only to encourage student reflection, self-expression, and critical thinking, but also to assess program quality and the achievement of specific course goals.

## Appendix 1. Application Essays

### Description

Application essays are opportunities to show how communication directly affects your life. They may include any phenomenon outside of class that are effective examples of course concepts discussed in class and throughout all units of the course. Application essays might include television shows, movies, newspaper articles, comics, guest speakers, personal conversations, etc. In a brief (one page) paper, you are to describe the application essay, link it to a communication concept, and analyze how the application essay is related to the communication concept.

### Format

Again, your paper should be one page long (typed and double spaced with no more than 1.25 inch margins and 12 point font). Your paper will contain two paragraphs:

(1)the description of the application essay,
(2) a link between the application essay and a communication concept, and an analysis of how the application essay is like the concept.

The first paragraph should discuss and/or describe the application essay in detail (who, what, where, when, how). As mentioned above, communication application essays can be anything from a sitcom episode to an article from a newspaper—exercise your imagination here. Basically, this first paragraph explains what happened, how it happened, and why you thought it happened. Just tell the story. Do not forget to cite where you found the application essay (if appropriate).

The second paragraph should identify ONE communication concept (reference class discussions or text material). The first sentence should link the application essay to ONE communication concept (i.e. the communication process model, ethics, listening, audience analysis, language,

delivery, small group communication, etc.); remember to cite where you found the communication concept (i.e. a particular portion of a class lecture, a concept in the textbook, etc.). Next, in the same paragraph, analyze the communication concept being discussed. This analysis should demonstrate that you have a mastery of
how the communication concept helps shape your understanding of the application essay or vise versa. Finally, illustrate why this interaction is important to your life. Provide strong support for your argument(s).

When appropriate, include the application essay with your paper.

Format
This paper is worth _____ points and is part of your total portfolio grade. The following is my criteria for evaluation: Format (_____ pts), Writing (_____ pts), Description (_____ pts), Link (_____ pts), and Analysis (_____ pts). This is a formal writing assignment and should be treated as such. Please plan what you want to write and then follow through with complete, but concise sentences. In other words, do not "think" or "talk" on paper.

## References

Barrow, D.A. (1994). The use of portfolios to assess student learning. *Journal of College Science Teaching, 22*, 148-153.

Black, L. Daiker, D., Sommers, J., & Stygall, G. (1994). *New Directions in Portfolio Assessment*. Portsmouth, NH: Boynton/Cook Publishers.

Chapman, P. (1996). Journals: Pathways to thinking in second-year algebra. *Mathematics Teacher, 89*, 588-590.

Crouse, G.K. (1994). Meeting new challenges in the foreign language classroom. Paper presented at the Annual Central States Conference on the Teaching of Foreign Languages, Kansas City, MO. (ERIC Document Reproduction Service No. ED 378815)

Fallon, M.A., Hammons, J., Brown, S., & Wann, J.S. (1997). Continuous assessment plans: A model for implementation. *Journal on Excellence in College Teaching, 8*, 21-40.

Farris, P., & Fuhler, C. (1996). Using response journals with preservice teachers. *The Clearing House, 70*, 26-29.

Gill, K. (1993). Process and portfolios in writing instruction. *Classroom Practices in Teaching English, 26*, Urbana, IL: National Council of Teachers of English.

Gipe, J., & Richards, J. (1992). Reflective thinking and growth in novices' teaching abilities. *The Journal of Education Research, 86*, 52-57.

Hunt, S.K., Simonds, C.J., & Hinchliffe, L.J. (2000). Using student portfolios as authentic assessment of the basic communication course. *Journal on Excellence in College Teaching, 11,* 57-77.

Jones, A.C., Simonds, C.J., & Hunt, S.K. (2005). The use of application essays as an effective tool for assessing instruction in the basic communication course. *Communication Education, 54*, 161-169.

Patzer, C.E., & Pettegrew, B.S. (1996). Finding a "voice": Primary students with developmental disabilities express personal meanings through writing. *Teaching Exceptional Children, 29*, 22-27.

Reyes, M.L. (1991). A process approach to literacy using dialogue journals and literature logs with second language learners. *Research in the Teaching of English, 25*, 291-313.

Sforzo, G.A. (2005). Assessing outcomes with computer-based vs. traditional paper-based testing: A preliminary study. *Journal on Excellence in College Teaching, 16*, 97-107.

Sircar, S., Fetzer, R.C., Patterson, J., & McKee, H.A. (2009). An assessment of electronic portfolios across the curriculum. *Journal on Excellence in College Teaching, 20*, 121-148.

Slater, T. (1995). *A qualitative and quantitative comparison of the impact of portfolio assessment procedures versus traditional assessment in the college physics course*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA. (ERIC Document Reproduction Services No. ED 391835).

Stefanou, S.E., Hood, L.F., & Stefanou, C.R. (2001). Feedback and change: Assessment of individual contributions within collaborative activities in the higher education classroom. *Journal on Excellence in College Teaching, 12*, 77-91.

Vizyad, L. (1994). Student portfolios: Building self-reflection in the first-grade classroom. *The Reading Teacher, 48*, 362-364.

Voth, B., & Moore, M.M. (1997). The pedagogical values of video portfolios in basic public speaking courses. *Journal of Excellence in College Teaching, 8*(2), 87-105.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 43 – 57.

# Measuring learning through cross sectional testing

**Steve Lovett[1] and Jennie Johnson[2]**

*Abstract: The measurement of student learning is becoming increasingly important in U.S. higher education. One way to measure learning is through longitudinal testing, but this becomes especially difficult when applied to cumulative learning within programs in situations of low persistence. In particular, many Hispanic Serving Institutions (HSIs) find themselves in such situations. Cross sectional testing is a pragmatic alternative, so long as maturity and selection effects can be estimated. The purpose of this paper is to demonstrate the utility and mechanics of measuring learning through cross sectional testing.*

*Keywords: assessment, measurement, student learning, maturity, selection.*

The measurement of student learning is fundamental to the scholarship of teaching and learning. Pace (2011) recently wrote:

> "...assessment is at the core of the entire SoTL enterprise. It is difficult to imagine a robust scholarship of teaching and learning unless our work is cumulative and built on previous research and unless there is a means to systematically evaluate the validity of claims being made about student learning" (p. 107).

Assessment, which begins with the measurement of learning, is becoming increasingly necessary for pragmatic reasons as well. Many parties, from prospective students to employers to governments, are increasingly demanding simple, quantitative performance measures from colleges and universities (Archibald & Feldman, 2008; Burke, 2002; Martell & Calderon, 2005), and accrediting bodies including the Association to Advance Collegiate Schools of Business (AACSB) and the European Quality Improvement System (EQUIS), are shifting their emphasis from input measures (number of faculty with terminal degrees, etc.) to outcome measures, or the extent to which students have achieved educational goals (Mundhenk, 2005; Rubin & Martell, 2009). The purpose of this paper is to demonstrate a simple, pragmatic procedure by which student learning may be measured through cross sectional testing. The focus is on cumulative learning within programs, not within individual courses. The sample with which we demonstrate the procedure comes from the business school of a Hispanic Serving Institution (HSI).

## I. Measuring Learning.

The two most important student outcome measures are persistence – students should finish their programs – and learning – students should gain knowledge, skills or abilities while in their programs. The importance of these two outcomes has been recognized in recent academic research. In a literature search, Robbins, Lauver, Le, Davis, Langley, and Carlstrom (2004) found 408 studies reporting on at least one of these outcome measures between 1984 and 2003.

[1] School of Business, The University of Texas at Brownsville, 80 Fort Brown, EDBC. 1.434 Brownsville, TX, 78520, steve.lovett@utb.edu
[2] School of Business, The University of Texas at Brownsville, 80 Fort Brown, EDBC 1.436, Brownsville, TX,78520, jennie.johnson@utb.edu

Of the two, however, persistence is easier to measure. A student enrolled in one semester either enrolls the next semester or does not, and either graduates within some time frame or does not. Furthermore, the total reenrollment or graduation rate for a program is a valid measure of overall persistence within the program. The temptation for some institutions, therefore, may be to focus on the easily measured outcome – persistence – and to neglect the hard-to-measure outcome – learning. However such an unbalanced set of priorities is not to our advantage as a society. Archibald and Feldman (2008) wrote:

> "...not every strategy a university might design to increase its graduation rate would be a good educational decision. Universities could achieve a higher graduation rate by lowering curricular standards or by encouraging more grade inflation. And any institution could surely achieve higher graduation rates by restricting access to students who are sure bets to graduate. Raising graduation rates in these last two ways is not socially useful since it would weaken the country's commitment to high quality programs and broad based access" (p. 81).

What is needed to maintain a balance between the two outcome measures is a pragmatic means of measuring learning within programs. In this paper we will not offer the reader an instrument by which to measure learning. The content of such an instrument, while obviously a critical issue for any program, is outside the scope of this paper. Rather, this is a method paper – our purpose is to demonstrate a simple, pragmatic procedure by which learning may be measured through cross sectional testing.

This paper will provide the reader with an example of how to measure learning through cross sectional testing in a case-study format. We continue with seven more sections. The next section compares the indirect and direct methods of measuring learning. For example, one useful direct method is testing. The third section continues with a comparison of longitudinal and cross sectional testing. The fourth section explains why maturity and selection effects must be accounted for to measure net gains in learning. The fifth section describes the sample used for this demonstration – freshmen and seniors in a business program at a Hispanic Serving Institution (HSI) – and the sixth, the procedure used to estimated net learning gains. The final sections are a discussion with suggestions for improving and expanding the procedure, and a conclusion.

## II. Indirect versus Direct Measurement.

A key step in managing any outcome is its measurement, and therefore all institutions have an interest in measuring student learning within their programs. Learning can be measured either directly or indirectly. The direct approach (used in this paper) is based on a *demonstration* of learning. The indirect approach is based on an *opinion* regarding learning, and this opinion may come from employers, alumni, teachers or even the students themselves (Martell & Calderon, 2005). However, the indirect approach has its shortcomings.

The simplest form of the indirect approach is self-assessment: ask the students how much they have learned. Perhaps because it is so simple, this method is commonly used, even in formal research. Sitzmann, Ely, Brown and Bauer (2010) found 166 recent studies in which self-assessment was used as a measure of learning. However, not all self-assessments are accurate. For example, Kruger and Dunning (1999) described a series of experiments that demonstrated

that low-performers systematically overestimated their performances, and Clauss and Geedey (2010) found that students were reasonably good at self-assessing their knowledge or recall of facts, but much less able to self-assess their comprehension or their ability to apply knowledge. Sitzmann et al. (2010) performed a meta-analysis and found only weak correlations between self-assessments and learning. They concluded that "… self-assessed knowledge is generally more useful as an indicator of how learners feel about a course than as an indicator of how much they learned from it" (p. 180).

Another convenient measure of learning is the student's Grade Point Average (GPA). However, a course grade is ultimately an instructor's opinion, and different instructors often grade very differently, to the point that a student's grade in any given course falls far short of being a reliable demonstration of learning. Pace (2011) wrote "… procedures for determining grades are generally shrouded in mystery and rest upon processes that may be perfectly legitimate for classroom teaching but do not provide a firm foundation for a systematic exploration of teaching and learning" (p. 108).

Direct measures of student learning are to be preferred and accrediting bodies are beginning to emphasize these. For example, the Association to Advance Collegiate Schools of Business (AACSB) emphasizes direct measurement of learning in the new accreditation standards adopted in 2003, and this may be the most significant change from the prior standards (Thompson, 2004). AASCB acknowledges on their website that indirect measures may have some value, but states clearly that "Such indirect measures, however, cannot replace direct assessment of student performance" (AACSB, 2012). Martell and Calderon (2005) were even more blunt in their rejection of indirect methods of assessment, writing: "… we advise deans to forget about surveys and other indirect measures when thinking about assessment [because they have] very little evidentiary value for assessment of student learning" (p. 223). Therefore, this paper focuses on the direct approach, based on a *demonstration* of learning.

## III. Longitudinal versus Cross Sectional Testing.

While students may demonstrate learning is a variety of ways, one method that gives the researcher significant control is through testing. If the students graduating from one program score higher on a test than those from another on the same test, it is reasonable to claim that the overall performance of the first program is better. The problem, however, is that the students in the first program might have been better prepared initially than those in the second, and a solution may be "before and after" or "value-added" testing – if students could be tested upon entering their programs and then again upon completion, the "gain" or the average difference in scores could be taken as a measure of learning. Of course, the students could also be tested at the beginning and end of a semester in order to evaluate learning within a course, or even several times during a semester (see, for example, Dellwo, 2010), but the emphasis in this paper is on the longer time period and on measuring learning within programs.

The preceding paragraph describes longitudinal testing – the same students are tested at the beginning and end of their programs. Adams and Schvaneveldt (1985) wrote "Advantages of this approach are that one observes or studies the same issue and the same people or events over a long period of time … [so] one can have greater control and ultimately more precise measurement with a longitudinal design" (pp. 115-116). However, they later added "A problem associated with the longitudinal design is the expense of following a population over time" (p. 116).

Rather than longitudinal testing, it would be simpler and less expensive to compare cohorts. For example, the entering class in a given year could be compared to the graduating class in the same year. This describes cross sectional testing. Adams and Schvaneveldt (1985) wrote "The cross sectional approach is obviously not as useful as the longitudinal design to assess change or development, but a number of inferences about change can be properly assessed within the constraints of this approach ... [and] ... for many research problems, the advantages of less time, fewer resources, larger samples, a large array of variables, and the versatility of the cross sectional approach indicate its central utility in social science research" (p. 115).

Cross sectional research is not without its critics. Seifert, Pascarella, Erkel and Goodman (2010) analyzed a large data set from nineteen U.S. institutions, first using a cross sectional method and then using a longitudinal method, and found significant differences between the results of the two methods. They concluded that that "... longitudinal pretest-posttest designs are the best way to estimate what students are learning … " (p. 13), and they referred to longitudinal testing as the "gold standard" (p. 14) in assessment of student learning. However, they also stated that "We clearly recognize that longitudinal, pretest-posttest designs put a greater burden on institutional researchers in terms of time, effort and resources than do cross sectional studies" (p. 14), and this last point is especially relevant to institutions with low persistence.

When persistence is low only a minority of the beginning students may actually graduate from a program, and many of those graduating may be transfer students who began in another program, making the problem of "tracking" students from entrance to graduation extremely difficult. Unfortunately, the low persistence rate scenario appears to be the more common within U.S. higher education. ACT (2010) reported that first to second year retention rates in four-year public institutions recently averaged only about 68%, with only about 43% of students completing their degrees within six years. Furthermore, it is the institutions with fewer resources that suffer the most from low persistence. Gansemer-Topf and Schuh (2006), in a study of 466 American institutions, found that wealthier and more selective institutions had far better retention than poorer or less selective institutions, and in fact reported that "Institutional expenditures and institutional selectivity explained over 60% of the variance in retention and graduation rates" (p. 622). The authors of this paper are particularly aware of the difficulties of measuring learning in situations of low persistence. We work in the business school of a Hispanic Serving Institution (HSI), and we have been frustrated in the past by the problems involved in tracking large numbers of students from entry to graduation. According to the U.S. Census Bureau (2010), U.S. Department of Education (2011), and U.S. Department of Labor (2011), persistence is a particular problem in HSIs.

Given low retention rates, tracking individual students through a program for longitudinal testing becomes prohibitively expensive for many poorer or less selective institutions. Therefore, while longitudinal testing may be preferred, as a pragmatic matter cross sectional testing is less burdensome because it eliminates the need to track individual students, and is therefore more likely to be realized at most institutions. Furthermore, cross sectional testing produces results quickly, while a researcher using longitudinal testing must wait for students to complete their programs before getting results.

## IV. Maturity and Selection Effects.

But at least two difficulties must be considered in cross sectional testing. The first is also a consideration in longitudinal testing; there may be a "maturity" effect. Graduating students are

older than incoming students and so may have gained knowledge or skills independently of their university experience. For example, students gain vocabulary by watching television, and gain math skills by balancing their checkbooks. Pascarella and Terenzini (2005) recognized this difficulty, writing "It is one thing to conclude that increases in subject matter knowledge and academic skills occur *during college*. It is quite another to conclude that these increases occur *because of college*" (p. 70). We must therefore be aware that knowledge and skills can be gained by students outside of the program whose effectiveness we are trying to measure.

The second difficulty is unique to cross sectional testing. There may be a "selection" effect – poorly performing students tend to drop out of the program, while high performing students persist. Tinto's (1975, 1993) model of student departure begins with pre-entry attributes such as skills, abilities and prior schooling – some students are more ready for college than others, and those who are less ready are less likely to perform well and therefore more likely to drop out. This relationship between readiness, performance and retention has documented by researchers such as Allen, Robbins, Casillas and Oh (2008), and the resulting selection effect will tend to raise the average performance of those remaining in the program, creating a false appearance of learning with the program.

Finally, note that maturity and selection are separate effects – that students learn outside of college is one thing, and that less prepared students are more likely to drop out is another. Therefore, both must be accounted for when attempting to measure learning within a program.

## V. Sample.

The subjects of this study were two groups of students attending the business school of an open enrollment, Hispanic Servicing Institution (HSI). The U.S. Department of Education (2011) defines an HSI as a public or private non-for-profit college or university that has a student population of at least 25% Hispanic and serves a higher portion of low to middle income students than their peer universities. Research shows that Hispanics have fallen behind in educational attainment/program completion compared to white student groups as well as other cultural/ethnic groups (Alon, Domina & Tienda, 2010).

One group was comprised of freshmen in the introductory business course, and the other of graduating seniors in the program's capstone course. Both groups coincidentally numbered sixty-three students. Near the end of the courses, both groups were given identical tests of sixty multiple choice questions which accounted for ten percent of the final course grade. The testing was therefore "course-embedded," or part of a course, rather than an entrance or graduation requirement. However, note that the basic techniques demonstrated in this paper would apply equally well to entrance/graduation testing.

There were four categories of questions, one for each of the three subject matter areas emphasized in the program – management, marketing and accounting/finance – and quantitative questions. Three demographic questions were also asked: the student's age (coded as a continuous variable), whether English was the student's first language (dichotomous variable), and whether the student was male or female (dichotomous variable).

Information as to the students' readiness for college was obtained from the school's admissions center. All students, whether first time freshmen or transfers, are required to take reading, writing and mathematics tests before enrolling in classes. The scores from these three tests were standardized and the average of the standardized scores will be referred to below as a "readiness composite" or simply "readiness." Similar data is available at most institutions.

Caison (2007) noted that "...institutions do routinely collect a broad array of information on their students' backgrounds, socioeconomic status, academic progress, and, in many cases, their academic goals and social involvement" (p. 436), and found that this archival information was more useful in predicting retention than the information from a survey instrument.

Table 1 below shows descriptive statistics from the sample. The bottom of the table shows that the seniors did outperform the freshmen on all four outcome scores. The gains in management, marketing and accounting/finance were 0.64, 0.43, and 0.79 standard deviations. Pascarella and Terenzini's (2005) "best estimate" (p. 66) of typical freshmen to senior gains for subject matter knowledge, taken from a synthesis of 17 previously published studies, is 0.87 standard deviations. These results show gains of somewhat less than that. The gains in the quantitative section were 0.90 standard deviations, greater than Pascarella and Terenzini's (2005) estimate of 0.24 standard deviations.

**Table 1.** Descriptive statistics.

| | Freshman (n=63) | Seniors (n=63) | Overall stddev | Gain in stddev |
|---|---|---|---|---|
| 1. Average age | 19.6 | 27.7 | -- | -- |
| 2. Readiness composite | -.357 | .353 | -- | -- |
| 3. % with English as a first language | 47.5 | 43.5 | -- | -- |
| 4. % female | 44.4 | 49.2 | -- | -- |
| | | | | |
| % correct | | | | |
| 5. Management | 69.0 | 78.6 | 15.0 | 0.64 |
| 6. Marketing | 69.1 | 76.0 | 16.2 | 0.43 |
| 7. Acct/finance | 67.7 | 81.5 | 17.4 | 0.79 |
| 8. Quantitative | 57.7 | 78.4 | 22.9 | 0.90 |

The demographics shown at the top of Table 1 indicate a possible maturity effect because the seniors were older than the freshmen. Also, the seniors' readiness composite was higher than the freshmen's, indicating a possible selection effect. Both groups were relatively similar in the percent of students with English as their first language and percent women.

Table 2 below displays correlations between the variables. Age correlated significantly with three of the four outcome scores, indicating a likely maturity effect. The readiness composite correlated even more highly with all four outcome scores, indicating a likely selection effect. Also, note that age was statistically significantly correlated with the readiness composite. Because the analysis below is based on regression with these two as independent variables, this indicates potential problems with multi-colinearity. None-the-less, because the two represent separate effects (see previous discussion), both will be used below. English as a first language and percent female did not correlate significantly with any of the outcome scores, and as Table 1 shows there was little difference between the freshmen and seniors in these variables. Therefore, these last two are not used below.

**Table 2.** Correlations.

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. |
|---|---|---|---|---|---|---|---|
| 1. Age | | | | | | | |
| 2. Readiness composite | .27* | | | | | | |
| 3. English first language | .09 | .10 | | | | | |
| 4. Female | .03 | .09 | .10 | | | | |
| 5. Management score | .20* | .39* | .00 | -.00 | | | |
| 6. Marketing score | -.02 | .30* | -.04 | -.09 | .47* | | |
| 7. Acct/finance score | .24* | .39* | .04 | .11 | .50* | .58* | |
| 8. Quantitative score | .21* | .43* | .04 | -.17 | .51* | .36* | .38* |

* correlation is statistically significant at the 0.05 level (2-tailed)

## VI. Procedure.

Equation 1 will serve as a starting point for cross sectional testing.

(1) Gain = Average Senior Score – Average Freshman Score

Program gains independent of maturity and selection effects will be referred to as the "net gain" of the program, as summarized in equation 2.

(2) Net Gain = Gain – Maturity Effect – Selection Effect

In this study two evaluation items were sought. The first was an assurance that some positive net gains were being realized in each subject matter area. The second was a point estimate of the net gains in each subject matter area. That positive gains were being realized was verified by using a series of four regressions, each with one of the outcome scores as the dependent variable, and with age, readiness, and a dichotomous variable indicating senior status as independent variables.

(3) $\text{score} = b_0 + b_1(\text{age}) + b_2(\text{readiness}) + b_3(\text{senior})$

The results are shown in Table 3. In the cases of marketing, accounting/finance and quantitative skills, the results were encouraging – the beta value of senior status was positive and statistically significant at $p<0.05$. Therefore, even after age and readiness were considered, the seniors still scored significantly higher than the freshmen. In the case of management, however, the beta value of senior status was positive but not statistically significant.

However, these regressions simply show whether learning was statistically significantly different from zero, and a more appropriate goal is to maximize learning. We must therefore calculate a point estimate of learning or gains for each area. Based on the data from Table 1, we might naively calculate the gains in management, for example, as the simple difference between seniors and freshmen in management: 78.6 – 69.0 = 9.6 percentage points. However, this does not take into account the fact that the seniors were on average older (maturity effect) and more ready (selection effect) than the freshmen.

**Table 3.** Regressions.

| | Dependent Variable | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Management | | Marketing | | Acct/finance | | Quantitative | |
| Variable | beta | sig. | beta | sig. | beta | sig. | beta | sig. |
| 1. Constant | .706 | .000 | .821 | .000 | .694 | .000 | .672 | .000 |
| 2. Age | .000 | .916 | -.006 | .030 | .000 | .957 | -.003 | .325 |
| 3. Readiness | .058 | .001 | .051 | .010 | .056 | .005 | .083 | .001 |
| 4. Senior | .052 | .117 | .078 | .038 | .096 | .012 | .174 | .000 |
| model r-square | | .179 | | .127 | | .209 | | .274 |

We therefore need estimates of the maturity and selection effects. These were obtained from a series of four regressions, each with one of the outcome scores as the dependent variable, and with age and readiness as independent variables, but without senior status.

(4) $\text{score} = b_0 + b_1(\text{age}) + b_2(\text{readiness})$

The resulting $b_1$ and $b_2$ coefficients are unbiased estimators of the effects of age and readiness on students' scores without regard to whether the student was a senior or a freshman. The coefficients for the regression using the management score as the dependent variable, for example, are shown in Step A of Appendix 1.

Recall from Table 1 that the freshmen had an average age of 19.6 years and an average readiness composite of -.357. What would be the predicted score of such a group, without regard to senior status? Step B of Appendix 1 shows that we would predict a score of 69.8%, and that the equivalent predicted score of the seniors, with an average age of 27.7 and a readiness composite of .353, would be 76.2%.

The difference between these two figures – 6.4% – is the gain that can be explained through the maturity and selection effects, and this must be subtracted from the total gain to obtain the net gains of the program. Step C of Appendix 1 shows this calculation. The net gain in management was found to be only 3.2%, or 0.21 standard deviations, not the 9.6% calculated above. Repeating this methodology, Appendices 2 through 4 show that the net gain in marketing was 4.7% or 0.29 standard deviations, in accounting/finance it was 5.3% or 0.30 standard deviations, and in quantitative skills it was 9.3% or 0.41 standard deviations. The figures above are similar to Pascarella and Terenzini's (2005) estimate of 0.26 to 0.32 standard deviations for the net effect of attending college.

## VII. Discussion.

First, note that the above analysis raises some concerns as to the sufficiency of the independent variables used. In particular, readiness (selection) appears to have had more powerful effects than age (maturity). Age was statistically significantly correlated with three of the four outcome variables, but readiness correlated even more strongly with all four outcome variables (see Table 2). Furthermore, when both age and readiness were included as independent variables in a regression equation, readiness was statistically significant while age was not. Now, this may be because selection is in fact the more powerful effect, however it also may be because selection

was well measured by the "readiness composite" (calculated through admissions center data), while maturity was poorly measured by age.

In general, readiness and therefore selection is likely to be relatively easy to measure. For example, ability measures such as SAT or ACT scores, and early academic performance measures such as high school graduation rank or GPA have been shown to be reliable predictors of success in college (Harackiewicz, Barron, Tauer, & Elliot, 2002; Pascarella & Terenzini, 2005; Willingham, Lewis, Morgan, & Ramist, 1990), and most institutions have these or similar measures available for both freshmen and seniors. However, measuring maturity means estimating the value or applicability of life experiences to a subject matter, and this is likely to be quite difficult. The authors, for example, have witnessed some very thought provoking classroom debates between students with different backgrounds as to whose experiences made them better able to correctly understand or interpret a particular business case or dilemma. Age as an estimate of maturity is certainly easy to obtain – most students are quite willing to tell a researcher how old they are – but the effect of a year's passing on one student may be very different from its effect on another. Still, there is no doubt that some measure of maturity is needed because students learn outside of the classroom, so a better measurement of maturity is a useful topic for future research.

In any case, the results presented here would naturally direct a faculty's attention toward examining the management area more closely. This was the area in which the least gains were found, and therefore it might present the greatest opportunity for improvement. But again, note that there are two possible reasons why less gains might be found in management. First, it could be that students are in fact learning less in this area. However, it could also be that the test did not accurately measure what they did learn. The issue here is the content of the measurement instrument, or, in other words, the content validity of the dependent variable, and as noted in the introduction, this was outside the scope of this paper. A good source for guidance in this area, however, is Martell and Calderon (2005), who describe the creation of such an instrument as an on-going five-step process in which an entire faculty reflect on what they want their students to learn, define learning goals, measure student performance relative to these goals, report and discuss the results, and then make appropriate changes in teaching or curriculum.

Ideally, the results above might provide a stimulus for the management faculty to begin this process. If so, an initial benefit might be a clearer consensus on learning goals among the management faculty, which in turn might lead to a better measurement instrument. Of course, this has value in itself. Still, if the faculty do no more than this they will have limited themselves to what Vockell and Asher (1995) describe as level 1 research, or data collection. Level 2 research, which depends on level 1 research, means establishing cause-and-effect relationships. In this case the management faculty might experiment with new teaching techniques or curriculum in an effort to improve learning, which is the fifth step in the Martell and Calderon (2005) process mentioned above.

Another interesting topic is the application of the techniques described in this paper to other learning outcomes and other assessment methods. Kraiger, Ford and Salas (1993) and Kraiger (2002) describe three kinds of learning outcomes – cognitive, skill-based and affective – each of which can be assessed through a variety of methods. The dependent variable in this paper is a score on a multiple-choice test, so a cognitive outcome was measured through recognition testing. However, many institutions also seek to teach skills, such as critical thinking or writing, which may be assessed through observation or by reviewing work samples. If, for example, student writing samples could be reliably scored, these scores could take the place of the

dependent variables used in this paper and net gains in writing skills could be estimated using the same method. Finally, institutions may seek affective outcomes such as positive attitudes or motivation, and, as Sitzmann et al. (2010) concluded, these may be accurately measured through self-reports. The techniques described in this paper, in which maturity and selection effects are estimated and subtracted from gains to find net gains, are applicable to any of these *so long as* reliable numeric scores can be assigned to student performances.

## VIII. Conclusion.

Just as excellence in teaching and learning requires openness to different modes of delivery and instructional approaches, practitioners and researchers need to be open to various approaches to the measurement of learning. This paper previously discussed the pros and cons of longitudinal and cross sectional testing to measure learning. Longitudinal testing is useful, but less practical, especially in institutions with low persistence. Such institutions may benefit more from cross sectional testing for measuring learning. According to the U.S. Census Bureau (2010), U.S. Department of Education (2011), and U.S. Department of Labor (2011), persistence is a particular problem in Hispanic Serving Institutions (HSIs), which have lower graduation rates than non-HSIs. Of course, this will have an important impact on tomorrow's workforce due to the continued growth of the Hispanic population in the United States. Therefore, the cross sectional method of measuring learning demonstrated in this paper via case study has particular implications for programs in HSIs and other higher education institutions with similar challenges**.** Cross sectional testing can be used to estimate students' net gains or learning within a program, and the procedure demonstrated here is simple and well within the means of most institutions.

## Appendices

**Appendix 1. Net Gain in Management.**

Step A - regression on management score.

| Variable | beta | std.err | sig. |
|---|---|---|---|
| Constant | .683 | .047 | .000 |
| Age | .002 | .002 | .230 |
| Readiness | .068 | .016 | .000 |

model r-square .162

Step B - calculate predicted values based on age and readiness.

| Freshmen | Seniors |
|---|---|
| .683 | .683 |
| + (.002* 19.6) | + (.002 * 27.7) |
| + (.068 * -.357) | + (.068 * .353) |
| ---------- | ---------- |
| .698 | .762 |

Step C - calculate net gain.

(.786 - .690) - (.762 - .698) = .096-.064 = .032

= 0.21 std. dev.

**Appendix 2. Net Gain in Marketing.**

Step A - regression on marketing score.

| Variable | beta | std.err. | sig. |
|---|---|---|---|
| Constant | .787 | .053 | .000 |
| Age | -.003 | .002 | .228 |
| Readiness | .065 | .018 | .000 |

model r-square .095

Step B - calculate predicted values based on age and readiness.

| Freshmen | Seniors |
|---|---|
| .787 | .787 |
| + (-.003 * 19.6) | + (-.003 * 27.7) |
| + (.065 * -.357) | + (.065 * .353) |
| ---------- | ---------- |
| .705 | .727 |

Step C - calculate net gain.

(.760- .691) - (.727 - .705) = .069 - .022 = .047

= 0.29 std. dev.

**Appendix 3. Net Gain in Accounting/Finance.**

Step A - regression on accounting/finance score.

| Variable | beta | std.err. | sig. |
|---|---|---|---|
| Constant | .652 | .055 | .000 |
| Age | .004 | .002 | .079 |
| Readiness | .074 | .019 | .000 |

model r-square .166

Step B - calculate predicted values based on age and readiness.

| Freshmen | Seniors |
|---|---|
| .652 | .652 |
| + (.004* 19.6) | + (.004 * 27.7) |
| + (.074 * -.357) | + (.074 * .353) |
| --- | --- |
| .704 | .789 |

Step C - calculate net gain.

(.815 - .677) - (.789 - .704) = .138 - .085 = .053

= 0.30 std. dev.

**Appendix 4. Net gain in quantitative score.**

Step A - regression on quantitative score.

| Variable | beta | std.err. | sig. |
|---|---|---|---|
| Constant | .596 | .071 | .000 |
| Age | .004 | .003 | .219 |
| Readiness | .115 | .024 | .000 |

model r-square .195

Step B - calculate predicted values based on age and readiness.

| Freshmen | Seniors |
|---|---|
| .596 | .596 |
| + (.004 * 19.6) | + (.004 * 27.7) |
| + (.115 * -.357) | + (.115 * .353) |
| --- | --- |
| .633 | .747 |

Step C - calculate net gain.

(.784- .577) - (.747 - .633) = .207 - .114 = .093

= 0.41 std. dev.

## References

AACSB. (2012). www.aacsb.edu/accreditation/business/standards/aol/learning_goals. Retrieved June 4, 2012.

ACT, Inc. (2010). *National collegiate retention and persistence to degree rates.* Iowa City, IA.

Adams, G.R., & Schvaneveldt, J.D. (1985). *Understanding Research Methods.* White Plains, NY: Longman.

Allen, J., Robbins, S., Casillas, A., & Oh, I. (2008). Third-year college retention and transfer: Effects of academic performance, motivation, and social connectedness. *Research in Higher Education, 49*, 647-664.

Alon, S., Domina, T., & Tienda, M. (2010). Stymied mobility or temporal lull? The puzzle of lagging Hispanics college degree attainment. *Social Forces, 88*(4), 1807-1832.

Archibald, R.B., & Feldman, D.H. (2008). Graduation rates and accountability: Regressions versus production frontiers. *Research in Higher Education, 49,* 80-100.

Burke, J.C. (2002). *Funding Public Colleges for Performance.* New York: Rockefeller Institute Press.

Caison, A.L. (2007). Analysis of institutionally specific retention research: A comparison between survey and institutional database methods. *Research in Higher Education, 48,* 435-451.

Clauss, J., & Geedey, K. (2010). Knowledge surveys: Students ability to self-assess. *Journal of the Scholarship of Teaching and Learning, 10*(2), 14-24.

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

Gansemer-Topf, A., & Schuh, J. (2006). Institutional selectivity and institutional expenditures: Examining organizational factors that contribute to retention and graduation. *Research in Higher Education, 47*, 613-642.

Harackiewicz, J.M., Barron, K.E., Tauer, J.A., & Elliot, A.J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*, 562-575.

Kraiger, K. (2002). Decision-based evaluation. In K. Kraiger (Ed.) *Creating, Implementing, and Managing Effective Training and Development: State-of-the-Art Lessons for Practice* (pp. 331-376). San Francisco: Jossey-Bass.

Kraiger, K., Ford, J., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning to new methods of training evaluation. *Journal of Applied Psychology, 78,* 311-328.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134.

Martell, K., & Calderon, T. (2005). Assessment in business schools: What it is, where we are, and where we need to go now. In K. Martell and T. Calderon (Eds.) *Assessment of Student Learning in Business Schools: Best Practices Each Step of the* Way (pp. 1-26). Tallahassee: Association for Institutional Research.

Mundhenk, R. (2005). Assessment in the context of accreditation. In K. Martell and T. Calderon (Eds.) *Assessment of Student Learning in Business Schools: Best Practices Each Step of the Way* (pp. 27-42). Tallahassee: Association for Institutional Research.

Pace, D. (2011). Assessment in history: The case for "decoding" the discipline. *Journal of the Scholarship of Teaching and Learning, 11*(3), 107-119.

Pascarella, E.T., & Terenzini, P.T. (2005). *How College Affects Students: A Third Decade of Research.* San Francisco: Jossey-Bass.

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130,* 261-288.

Rubin, R., & Martell, K. (2009). Assessment and accreditation in business schools. In S. Armstrong and C. Fukami (Eds.) *The SAGE Handbook of Management Learning, Education and* Development (pp. 364-384). Los Angeles: Sage.

Seifert, T.A., Pascarella, E.T., Erkel, S.I., & Goodman, K.M. (2010). The importance of longitudinal pretest-posttest designs in estimating college impact. *New Directions for Institutional Research,* Assessment Supplement, Winter, 5-16.

Sitzmann, T., Ely, K., Brown, K., & Bauer, K. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning and Education, 9*(2), 169-191.

Thompson, K. (2004). A conversation with Milton Blood: The new AACSB standards. *Academy of Management Learning and Education, 3*, 429-439.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45,* 89-125.

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition* (2nd ed.). Chicago: University of Chicago Press.

United States Census Bureau. (2010). *Educational attainment in the United States: 2010 Detailed Tables.* Retrieved April 29, 2012 from https://www.census.gov/hhes/socdemo/education/data/cps/2010/tables.html

United States Department of Education. (2011). *Moving America forward: President Obama's agenda for the Latino community*, pp. 1-5. Retrieved April 29, 2012 from http://www.2.ed.gov/about/inits/list/hispanic-initiative/obama-agenda.pdf.

United States Department of Labor. (2011). *The Hispanic labor force*. Retrieved April 29, 2012 from http://www.dol.gov/_sec/media/reports/hispanicelaborforce.pdf.

Vockell, E. L., & Asher, J. W. (1995). *Educational Research* (2nded.). Englewood Cliffs, NJ: Prentice Hall.

Willingham, W.W., Lewis, C., Morgan, R., & Ramist, L. (1990). *Predicting College Grades: An Analysis of Institutional Trends over Two Decades.* Princeton, NJ: Educational Testing Service.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 58 -75.

# Civic engagement in the field of Psychology

**Tiffany Chenneville[1], Susan Toler[2], and Vicki T. Gaskin-Butler[3]**

*Abstract: The purpose of this article is to describe the importance of, and recommendations for how best to promote, civic engagement among undergraduate psychology majors. In this article, we will describe how the goals of civic engagement are consistent with the specific curricular goals of undergraduate psychology programs. We also will (a) review the empirical support for civic engagement in the field of psychology and describe the implications of this method for teaching students about diversity; (b) discuss some of the challenges associated with incorporating civic engagement in psychology courses as well as provide strategies for overcoming these challenges; (c) discuss some of the unique ethical issues related to civic engagement in the field of psychology; and (d) provide recommendations, using specific examples, for how to incorporate service-learning activities as a means of encouraging civic engagement in psychology courses.*

*Keywords: civic engagement, psychology, service-learning*

## I. Introduction.

Working for the common good is a fundamental concept in the field of psychology and, in fact, in this article, we argue that the curricular goals of undergraduate psychology programs are consistent with the goals of civic engagement. Promoting civic engagement in the field of psychology is supported by the literature, yet challenges exist when attempting to incorporate civic engagement as a central component of psychology courses.

The concepts of service-learning and civic engagement are not new to the field of psychology. William James, considered by many to be the "father" of psychology in America, argued a century ago that youth should be encouraged to participate in community service as a means of mobilizing people by giving them a sense of common purpose. James argued that our society addressed the need for solidarity and a sense of common purpose by mobilizing military action rather than social action (James, 1910). John Dewey, president of the American Psychological Association in 1899, was the founder of *progressive education*, pedagogy that became the foundation for what we know today as service-learning or civic engagement (Rocheleau, 2004). Dewey proposed that students need to struggle with social problems through the application of concepts and principles learned in the classroom. This emphasized the pedagogical method of using community service to enhance students' understanding of facts and theories. Both Dewey and James emphasized the importance of theoretical relevance. That is, viable theories must be capable of addressing every day social and ethical problems. Progressive education must encourage students to actively participate in the democratic process and tackle difficult issues related to social justice.

[1] Assistant Professor of Psychology, University of South Florida St. Petersburg, chennevi@stpt.usf.edu
[2] smtoler@usfsp.edu
[3] vgaskinb@mail.usf.edu

As the progressive education movement evolved, dialogues returned to the responsibility of educators to produce citizens who had both the knowledge and the sense of civic responsibility to engage in community activism and social problem solving. This model emphasized nurturing a set of values and social ideals that produced citizens who participated in the democratic process and were civically engaged. The educational goal went beyond facilitating the acquisition of facts and theories to nurturing a set of values and social ideals that produced citizens who actively participated in the democratic process and were civically engaged. Civic engagement, then, emphasized the educational goal of instilling democratic and social values. In contrast the service-learning model emphasized understanding facts and knowledge from the classroom through community service.

Near the end of Dewey's life, the pedagogical foundations of service-learning were criticized for lack of academic rigor and breadth. There was a growing sense that progressive education became technical education, preparing students for jobs rather than careers (Rocheleau, 2004). Additional criticisms of this movement focused on indoctrination of values that promoted a brand of liberal citizenship and community activism. It was argued that traditional education focused on imparting facts, knowledge and theories in a value neutral setting (Fish, 2003). Educational trends in the mid to late twentieth century shifted to a focus on careers and self-interest concurrent with a sense of social and political alienation (Bellah, Madsen, Sullivan, Swindler, & Tipton, 1985).

Rocheleau (2004) has suggested that this social alienation and the call for relevance in the college classroom set the stage for a resurgence of service-learning and civic engagement near the end of the twentieth century. Social crises such as the terrorist acts of September 2001, the economic meltdown of 2008 and the ethical and moral bankruptcy of our political and business leaders have given educators cause to reconsider their responsibilities to educate ethical and responsible citizens as well as scholars. It is no coincidence that President Obama's call for a new type of civically engaged American citizen emerged from his own experience as a community organizer in the south side of Chicago and as a lecturer at the University of Chicago. The University of Chicago was the birthplace of John Dewey's philosophy of pragmatism and progressive education and home to the first psychology laboratory.

*A. Service-learning versus Civic Engagement.*

"Service-learning" and "civic engagement" often are used interchangeably and, while they are related, there are important distinctions between the two terms. The American Psychological Association (n.d.) defines civic engagement as "individual and collective actions designed to identify and address issues of public concern". Bringle and Hatcher (1996) describe service-learning as "a credit-bearing educational experience in which students participate in an organized service activity that meets identified community needs and reflects on the service activity in such a way as to gain further understanding of course content, broader appreciation of the discipline, and an enhanced sense of civic responsibility" (p. 222). While interrelated, these terms clearly signify different constructs.

In this article, we conceptualize service-learning as a means by which citizen scholars are created. Service-learning typically refers to a pedagogical approach that allows for students to better understand course material by providing a means by which to apply what they are learning in real life settings. In and of itself, in the absence of clearly specified goals, service-learning can be self-serving for the student. While community members may benefit in the short-term from

the students' participation in community settings, service-learning that is designed as a student-centered approach may lack the components necessary to achieve significant changes within the student, the university, or the community. Civic engagement extends beyond mere service-learning. Civic engagement reflects a reciprocal relationship between the student, the university, and the community. Therefore, intention is important when designing service-learning activities. In this article, we discuss service-learning as a means by which to promote civic engagement, which should be the ultimate goal of all service-learning activities.

*B. Learning Outcomes in Psychology.*

The American Psychological Association supports the idea that, as educators, we have a responsibility to inculcate our students with certain values. Indeed, an overarching goal of a liberal arts education is to assist students in "recognizing, understanding, and respecting the complexity of sociocultural and international diversity" (American Psychological Association [APA], 2007, p. 210). Suggested student learning outcomes include sensitivity to individuals from diverse backgrounds and abilities, heightened consciousness of prejudicial attitudes and discriminatory behaviors, and an understanding of how "power, privilege and oppression may affect prejudice, discrimination, and inequity" (APA, 2007, p. 20). What better classroom is there for understanding these social forces than in a homeless shelter, an inner city elementary school, or a battered women's shelter? Service-learning opportunities provide exactly these types of pseudo classrooms in which students not only extend their learning of concepts taught at the university, but also give back to the community in some meaningful way.

The American Psychological Association identifies even more specific learning outcomes to educate psychology majors in the value of "empirical evidence and tolerance for ambiguity" as well as the importance of "assessing and justifying their engagement with respect to civic, social, and global responsibilities" (APA, 2007, p. 17). Allowing students to engage in community service while at the same time giving our theories and facts the pragmatic test of realism offer powerful and sometimes challenging learning opportunities. In the sections to follow, we will discuss some of these learning opportunities in the field of psychology as well as some of the important considerations to be made when thinking about implementing civic engagement activities.

## II. Civic Engagement in Psychology.

In this section, we will discuss the empirical support for civic engagement via service-learning in psychology courses and the implications of this method for teaching students about diversity.

*A. Empirical Support.*

While the benefits of service-learning have been well documented (e.g., Prentice, 2007), there is a dearth of information about best practices for designing and implementing service-learning activities in higher education courses. And, most of what is available falls outside the field of psychology. However, there is some research that specifically addresses the benefits of, and recommendations for, implementing service-learning and civic engagement into the undergraduate psychology curriculum (e.g., Kretchmar, 2001; Fenzel, 2001; Grimes, 1998; and McDonald, Caso, & Dee, 2005).

Kretchmar (2001) reported that, when given a choice between participating in a service-learning activity and completing a research project, over 88% of students enrolled in an introductory psychology course opted for service-learning. In this course, students were given the opportunity to secure their own placements from a list of possible community agencies (e.g., Habitat for Humanity, Juvenile Detention Center). Of those who participated in service-learning, the large majority (80%) reported positive results related to learning and service commitment.

Similarly, Fenzel (2001) reported favorable course evaluations from the large majority (between 89 and 100%) of undergraduate students enrolled in one of three courses with a significant service-learning component (i.e., Introductory Psychology, Child Psychology, Substance Abuse and its Effects). The service-learning activities varied by course. For example, students enrolled in Introductory and Child Psychology courses participated in tutoring programs at various sites while students enrolled in the Substance Abuse course were required to teach a 6-hour unit on drug use prevention to middle school students. Most students (77%) indicated an increase in personal and social responsibility and many (59 to 88%) reported an increased interest in community service as a result of their service-learning experience.

Grimes (1998) also reported the positive results of incorporating service-learning into an Applied Psychology course. In this case, the service-learning component involved mentoring at-risk, African American middle school students. Of the 14 students enrolled in the course, over 80% described the course as one of the best courses they had ever taken and believed the course should become part of the required curriculum. Over 70% reported that the course increased their understanding of psychological concepts and their interest in public service. And, in fact, over 90% indicated an interest in continuing their work as a mentor.

Citing not only the benefits for students, but also for instructors and the community, McDonald, Caso, and Dee (2005) describe the value of teaching and learning operant principles in animal shelters. As part of a service-learning requirement, students enrolled in a Psychology of Learning course volunteered as dog trainers in an animal shelter. In addition to enhanced self-esteem and an increased interest in social causes among students, both instructors and community members benefited from this requirement.

*B. Teaching Diversity.*

The importance of incorporating a focus on multicultural issues into psychology courses cannot be overstated. Indeed, many undergraduate psychology curricula include specific courses designed to address issues related to diversity (e.g., Multicultural Psychology, Psychology of Women); and in addition, or at the very least, imbed issues of diversity by incorporating multicultural issues within all courses, regardless of the subject matter (e.g., discussing how theories of aging may differ across cultures in a Developmental Psychology course). The intersection between civic engagement and teaching diversity can be viewed from two different perspectives. First, promoting civic engagement among students via service-learning activities may serve as a vehicle for teaching students about diversity by exposing them to other cultures and sub-cultures within the community, regardless of the intended content of the course. For example, students may be exposed to the conditions of inner city schools while completing the service-learning component of a Behavior Modification course. Second, civic engagement can be incorporated into courses specifically designed to teach issues related to diversity (e.g., Psychology of Women, Multicultural Psychology).

Mio, Barker, and Tumambing (2009) define multicultural psychology as "the systematic study of behavior, cognition, and affect in settings where people of different backgrounds interact" (p. 3). Therefore, the pedagogical task of psychology courses containing multicultural content is to enable students to cross boundaries of race, gender, ability, class, and other differences, in order to better prepare them to make meaningful connections with others in our increasingly diverse society.

Multicultural psychology aids students in understanding the impact of culture on the development of worldviews, communication patterns, acculturation, and identity development. In addition, students in multicultural psychology classes are acquainted with issues related to stereotyping, prejudice, discrimination, and racism. Cultural factors related to health and mental health also are concerns of multicultural psychology. The ultimate goal of multicultural coursework is to move students toward the development of multicultural competence (Mio et al., 2009).

The American Psychological Association states that civic engagement includes a range of civic acts and "[a]n engaged citizen should have the ability, agency, and opportunity to move comfortably among these various types of civic acts" (APA Education Directorate, 2009, p. 1). Therefore, civic engagement is particularly important in multicultural classes. First, it affords students the opportunity to "address issues of public concern" (APA Education Directorate, 2009, p. 1). Second, civic engagement activities enable students to expand their knowledge of and appreciation for the experiences of those from diverse communities (Goodman & West-Olatunji, 2008; Barrow, 2008).

Beyond incorporating civic engagement into multicultural psychology courses, multicultural content can be incorporated into all psychology courses through service-learning activities designed to enhance communication, collaboration, and critical thinking across race, ethnicity, and culture (Grier-Reed, Detzner, Poch, & Staats, 2010). Hurtado (2007) argues for linking diversity, education, and civic engagement as a means by which we can "advance students' awareness of the origins of complex social problems and employ new forms of pedagogy involving dialogue, discussion, experiential learning, reflection, social critique, and commitment to change" (p. 186). In essence, incorporating civic engagement activities in psychology courses promotes the development of multicultural competence. Laird, Engberg, and Hurtado (2005) noted that colleges and universities' creation of programs that foster students' "engagement with social diversity" (p. 448) prepare students to deal with a diverse society. Furthermore, teaching courses that discuss multicultural content also increases the likelihood that students will work and live in diverse communities after graduating from college (Laird, Engberg, & Hurtado, 2002).

## III. Challenges.

Despite the fact that research has demonstrated the value of civic engagement for undergraduate psychology majors (Fenzel, 2001; Grimes, 1998; Kretchmar, 2001; McDonald, Caso, & Fugit, 2005), barriers exist that can deter faculty from incorporating service-learning as a means of promoting civic engagement into their courses. These barriers, or challenges, can be categorized into three groups: administrative challenges, faculty challenges, and student challenges. While these challenges are discussed as distinct categories in the sections to follow, it is important to recognize that a significant amount of overlap exists between these areas. For example,

administrative challenges cannot be separated entirely from faculty challenges as there clearly is a reciprocal relationship between administrative demands and faculty commitments.

*A. Administrative Challenges.*

It goes without saying that administrative support is imperative if efforts to promote civic engagement are to be successful. Without a stated commitment, higher education institutions are likely to get sidetracked with competing initiatives. However, even at universities that include civic engagement as part of their mission statement, funding is likely to be an issue. Especially in difficult economic times where budget cuts_result in universities that are understaffed, the support necessary to sustain initiatives often is lacking. In addition to funding issues, a lack of enthusiasm for civic engagement may underlie resistance at the administrative level. Further, institutions without a process for centralized decision-making and those that lack a strong sense of shared governance are likely to find it difficult to implement service-learning into the curriculum, much less promote civic engagement university wide (Ward, 1996). Similarly, lack of presidential support and poor faculty participation are likely to be barriers to institutionalizing service-learning initiatives that promote civic engagement.

*B. Faculty Challenges.*

Faculty interested in incorporating service-learning components into their courses or otherwise promoting civic engagement face many challenges, all of which ultimately relate to time constraints. Developing a curriculum that promotes civic engagement is time consuming. Many university faculty juggle multiple roles and responsibilities (e.g., teacher, researcher, mentor, and/or faculty governance) making it difficult to devote the time and energy necessary to successfully launch a civic engagement initiative, much less sustain one, which requires ongoing supervision and documentation. Bulot and Johnson (2006) estimate that courses that contain service-learning require an extra two to ten hours per week (4.5 hours on average) compared to courses without a service-learning component. Even at institutions where administrative support is adequate, which may reduce the demands placed on faculty with regard to managing civic engagement initiatives, faculty must prioritize competing demands on their time. This is especially true of faculty interested in tenure or promotion where the value of promoting civic engagement is not clearly articulated.

A lack of faculty development related to service-learning and civic engagement also may present a challenge. Even faculty who are intrigued by the concept of civic engagement may not believe they possess the knowledge or skills to successfully implement civic engagement activities. Lacking self-efficacy with regard to civic engagement, faculty likely will avoid pursuing opportunities to incorporate service-learning in their courses. Indeed, some faculty may not even realize they already promote civic engagement in their courses. In addition to adequately defining service-learning and civic engagement, faculty development efforts should emphasize the need for creating a balance between the needs of the course, the program, and the department with the needs of the students and the community agencies involved (Osborne & Renick, 2006).

Forging community partnerships is likely to be one of the greatest challenges faced by faculty members interested in civic engagement. Identifying and contracting with community agencies can be both time consuming and confusing. Legalities must be addressed, which can

serve as a significant barrier. The faculty-community agency partnership must be crafted carefully to ensure that student involvement extends beyond mere volunteerism (Bolut and Johnson, 2006). Without clear expectations, both parties may be less than satisfied with the experience.

In a study of faculty implementing service-learning in undergraduate and graduate gerontology-related courses, Bulot and Johnson (2006) described the need to overcome stereotyped thinking among students as a significant barrier. Specifically, the issue becomes how faculty can confront discriminatory attitudes among students without losing focus on other important course-related content. Bulot and Johnson (2006) also describe the challenge of balancing flexibility and rigidity, noting the importance of clear expectations but also acknowledging the need to take into consideration various student variables when designing service-learning activities.

*C. Student Challenges.*

Various student variables may be perceived as challenges to the successful implementation and maintenance of service-learning and civic engagement initiatives. For example, the age, maturity level, and life experience of individual students may affect the extent to which certain service-learning activities are useful. Bulot and Johnson (2006) suggest that nontraditional students (24 years or older) may pose unique challenges given the likelihood of competing responsibilities such as family obligations and work demands. On many campuses, especially in urban settings, even traditional students (under the age of 24) are working full- or part-time to put themselves through college. Thus, instructors must be creative when incorporating service-learning into their courses. In fact, Long, Larsen, Hussey, and Travis (2001) stress the importance of allowing students choices when designing service-learning activities to avoid some of the difficulty associated with students' outside commitments, which Bulot and Johnson (2006) refer to as the *social situation*. Another challenge among students may be confusion about what is expected of them, which speaks again to the need for clearly articulated goals and expected outcomes as well as functional relationships between faculty/universities and participating community agencies. Finally, students may simply lack interest in service-learning, which represents a challenge.

*D. Distance Learning.*

Despite controversy about its effectiveness, distance learning is a current trend that must be acknowledged. While a discussion about the implications of distance learning on civic engagement are beyond the scope of this article, we would be remiss if we did not mention the inherent difficulties associated with incorporating service-learning into online courses. Kenworthy-U'Ren (2008) provides an excellent summary of the key implications associated with incorporating service-learning into online courses: (1) reflection activities must be adapted to an online environment; (2) service-learning projects will have to be re-structured and maybe even re-operationalized, taking into account logistical issues; (3) there may be a shift among key stakeholders (e.g., faculty, students, community partners) with some more comfortable than others with the inevitable changes that will be required; (4) service e-learning may require more or different resources from community partners, thus limiting the involvement of agencies incapable of meeting the technological needs of the course; (5) partnership opportunities may be strengthened by Internet access given that students can locate potential placements; and (6)

students may benefit from the convenience of electronic journaling. Readers interested in learning more about opportunities to combine distance learning with civic engagement initiatives are referred to the section on distance learning in Appendix A.

*E. Overcoming Challenges.*

Clearly, incorporating service-learning activities into the undergraduate curriculum and, thus, promoting civic engagement among college students can be challenging for all of the reasons described above. However, many of these barriers can be overcome. Creating a university climate that promotes citizen scholars will require action on the part of all stakeholders to include not only administrators, faculty, and students, but also community collaborators. Both top down (i.e., administrative) and bottom up (i.e., students, faculty) initiatives are needed for civic engagement to be integrated successfully into higher education settings.

Civic engagement must be incorporated into the university's mission before we can expect individual departments to embrace the idea of imbedding service-learning into the curriculum. This will require thoughtful discussions and collaboration between administrators, faculty, students, and community agencies. Having a clear mission will set the stage for pursuing specific initiatives to create an environment where civic engagement is valued and practiced. Creating such an environment may require the reorganization of university goals and priorities. Because of the time demands associated with implementing service-learning activities, planning is needed to ensure that necessary resources are available and that stakeholders are not unduly burdened.

Organizational restructuring notwithstanding, there are some strategies that may ease the burden for individual faculty attempting to implement service-learning activities into their courses. For example, teaching assistants (TA) can help individual instructors manage activities related to service-learning/civic engagement. Even when funds are not available to pay TAs, many students are willing to volunteer as a TA for the experience it provides.

## IV. Ethical Issues.

While morality and ethics are important across disciplines, professional ethics become a particular issue when incorporating service-learning/civic engagement activities into psychology courses. Because service-learning activities in psychology courses are likely to expose students to vulnerable populations, ethical issues must be addressed. Although a standard part of the curriculum in many graduate psychology programs, ethics is not a standard part of the curriculum in most undergraduate psychology programs. While ethics may be embedded into courses across the undergraduate curriculum, a course dedicated to ethics is not among the most frequently listed courses in undergraduate psychology curriculum (Perlman & McCann, 1999). In this section, we will discuss some of the major ethical issues of concern when designing and implementing service-learning activities into psychology courses. While the topic areas correspond to some of the ethical standards described in the Ethical Principles and Codes of Conduct of the American Psychological Association (APA) (2002), it is important to understand that the APA does not accredit undergraduate psychology programs. Clearly, undergraduate students majoring in psychology, even those who are student members of the APA, are not held to the same principles and standards expected of doctoral level psychologists. However, the principles and standards outlined in the APA Ethical Principles and Code of Conduct (2002)

provide a framework for thinking about ethical issues that may arise when incorporating service-learning in undergraduate psychology courses.

*A. Competence.*

Competence is difficult to define given that it is variable across time, task, and situation. However, if simplified, professional competence generally refers to one's skills and abilities that affect performance and can be compared to a measurable standard. Within the field of psychology, it is clearly understood that discipline-specific competencies vary across education and training level as well as specialty area. The expected competence of an undergraduate psychology major differs dramatically from the expected competence of a licensed psychologist or even psychology graduate student, with the standards much higher for the latter. However, as Rosenthal, McKnight, and Price (2001) demonstrate, confusion persists about the standard level of training for psychologists, even among undergraduate psychology students. Therefore, it is the responsibility of faculty to educate both students and community partners about the appropriate role of undergraduate psychology students. For example, undergraduate students should never be sent out into the community to provide psychological services given that they lack the necessary training and experience. These limitations need to be clearly delineated to avoid potential confusion about what the student can and cannot do in a given placement. Failure to do so could result in significant harm to the general public.

*B. Privacy and Confidentiality.*

Confidentiality is important and privacy rights must be considered when designing service-learning activities. For example, consider a child psychology course that places students in a local elementary school to volunteer (e.g., read to children) and requires students to keep track of their observations in a journal to be submitted to the instructor. Students should be required to ensure the confidentiality not only of the children they observe and with whom they work, but also the teachers and other school personnel. While information shared between students and the individuals they serve through community placement is not considered privileged, respect for privacy should guide decisions about what to disclose and to whom.

*C. Informed Consent and Assent.*

Respect for individual autonomy also is an important issue when developing courses designed to promote civic engagement via service-learning activities. In the case of university-community partnerships, the obvious and primary concern is for the recipient of services provided through various community agencies (e.g., students in an elementary school, battered women in a domestic violence shelter). Questions can arise with regard to informed consent or, in the case of children, informed assent. For example, should informed consent be obtained from parents whose children will be receiving classroom tutoring from a local university student? Further, is assent from the children themselves necessary? Assuming consensus can be reached about what situations warrant consent, which is a lofty goal in and of itself, then one must consider what information should be disclosed as part of the consent process.

Also relevant to this discussion is respect for the individual autonomy of students. Specifically, what information should be provided to students enrolling for a course with a

significant service-learning component or, further, to students majoring in psychology in a program whose mission dictates that civic engagement be fostered? Handelsman and Rosen (1987) describe the limitations associated with implying student consent based on course registration. That is, given variations in teaching philosophies and styles that, to a large degree, dictate assignments and grading, not to mention content, it may be unfair to imply consent based on course registration. Handelsman and Rosen (1987) remind educators that students' decisions should be informed and, thus, adequate information should be provided prior to course registration. This is extremely important in the case of service-learning. Service-learning activities often require a substantial amount of time outside of the classroom, which may be difficult or even impossible for some students. It, therefore, may be prudent to somehow notify students in advance about the requirements of the course so that an informed decision can be made.

*D. Supervision.*

Site supervision is critical for service-learning placement. Ensuring that appropriate site supervision occurs will require that sites, themselves, be monitored as well as students' progress at those sites. To this end, it is important to establish formal connections with site supervisors, when possible, to ensure that they understand the nature of the students' assignment and the importance of their evaluations in determining students' grades.

*E. Research.*

Research can be an integral part of service-learning initiatives designed to promote civic engagement. In fact, Nigro and Wortham (2006) encourage including action research as a part of service-learning activities. Student research designed and conducted in collaboration with community partners can be extremely valuable, especially from the perspective of promoting civic engagement. However, the potential gains of research must be balanced against the risks and, even when the gains clearly outweigh the risks, precautions must be taken to ensure that risks that do exist are minimized. All of the ethical issues previously described apply to research activities. Student researchers must be *competent* in the areas of research design and methodology and may require *supervision* to carry out research projects. Appropriate actions must be taken to ensure the *privacy rights and confidentiality* of research participants. Of particular concern when thinking about conducting research in community agencies via service-learning placements is *informed consent and assent*. This issue becomes less problematic with "formal" research given that permission must be obtained from appropriate Institutional Review Boards (IRB) (e.g., university IRB and hospital IRB in the case of a service-learning placement in a hospital). More problematic is "informal" research. For example, consider the Behavior Modification course described in an earlier section. As described, students are placed in a local elementary school where they are required to apply learning principles when tutoring children and are required to track the children's progress as part of a course assignment. Essentially, the primary course assignment involves a single subject design. Assuming only the instructor has access to the data collected, should informed consent/assent be obtained from children and their parents? What if others (e.g., classmates) have access to the data collection? Or, what if a student wants to publish his/her findings? These are important questions that should be considered proactively.

*F. Recommendations for Addressing Ethical Issues.*

As demonstrated, many ethical issues arise within the context of promoting civic engagement among undergraduate psychology majors. And, although a discussion of liability is beyond the scope of this article, it is important to recognize that overlap exists between ethical standards and legal standards. Policies and procedures should be established for responding to ethical issues as they arise. However, it is impossible to foresee what issues might arise in a given class or setting or for a particular student; and therefore, reactive approaches are limited. Faculty, administration, and community partners are encouraged to be proactive in their approach to instilling the importance of professional ethics. One option is to incorporate a professional ethics course, specific to the field of psychology, into the undergraduate psychology curriculum and, further, to designate this course as a prerequisite to courses with a service-learning component. This option has significant limitations, though. Among these limitations is the well documented fact that knowledge does not always translate into behavior. This means that even students who perform well in an ethics course may not demonstrate appropriate behavior outside the classroom.

In his work on spirituality and psychotherapy, Plante (2007, 2009) recommends a model of ethics that focuses on five central values: (1) respect, (2) responsibility, (3) integrity, (4) competence, and (5) concern. It seems this model fits well for all aspects of fostering civic engagement among undergraduate psychology majors from administrative policy to faculty course development to the establishment of community partnerships to the promotion of professional behavior among students.

## V. Psychology Course Examples.

Service-learning, with the goal of promoting civic engagement among undergraduate students, can be incorporated into psychology courses. In this section, we will provide some examples of how civic engagement has been incorporated into psychology courses at a relatively small southeastern university with approximately 3,800 undergraduate students and 400 graduate students. The student to faculty ratio at this institution is 19:1. In the courses described, student enrollment averages anywhere between 30-50 students per semester. Following the class examples is a discussion of how some of the ethical issues described above were addressed.

*A. Child Psychology.*

Child Psychology is an upper level course offered by psychology departments, either as part of the major or as a general education course available to all undergraduates. The course content generally includes an overview of cognitive, physical, emotional and psychological growth and development from birth through adolescence. Typically, enrolled students have very little experience with children. Most do not have children of their own. Many do not have younger siblings, and the majority of students have limited exposure to children through employment or leisure activities. Civic engagement has been incorporated into this course by requiring students to volunteer a minimum of two hours a week over the course of a semester in an elementary school or daycare setting. Students are provided the opportunity to become classroom volunteers and observers of children across a wide range of developmental stages. Throughout the semester, students are required to maintain a journal in which they reflect on specific developmental theories. This journaling exercise requires students to reconcile their personal observations and

interactions with children with the theories introduced in the classroom. For example, journal topics focus on observations of the temperamental differences in children and reflections on the "goodness of fit" with their teacher to observations and reflections on Vygotsky's view of children's use of private speech. Due dates for journal entries are scheduled to coincide with lectures on specific topics. Students are asked to discuss their experiences in the community, which often results in very lively and sometimes emotionally charged discussions.

*B. Psychology of Women.*

Psychology of Women is an upper level general education course often offered as an elective to undergraduate psychology majors. The course content includes providing a "challenging and thought provoking" (Hyde, 2007, p. ix) stimulus for students to engage research related to the psychology of women. This course also enables students to understand women's experiences resulting from biological and social/cultural factors. In addition, this course allows students to understand the psychology of women from a multicultural perspective.

Civic engagement has been incorporated into this course by requiring students to volunteer a minimum of ten hours over the course of the semester in a local community agency that provides services for women and/or girls. It is noteworthy that many sites do not provide services exclusively to this population; rather, many sites also provide services to men/boys (e.g., elementary school and middle schools, homeless shelters, women's shelters, and women's residences). Students' volunteer activities range in depth and breadth; some students merely observe the target population (i.e., the clients who receive services from a particular agency). However, others complete tasks that ultimately benefit the target population or, in some cases, provide direct service to the target population. Students are required to have a site supervisor who verifies the completion of their volunteer hours and who evaluates their performance. During their volunteer experience, students are required to complete a journal in which they reflect on their experience. Journal entries are completed after each site visit. At the conclusion of the volunteer experience, students are required to complete a research paper and conduct a presentation designed to integrate their experiences with recent psychological research on the psychology of women. Students whose volunteer activities do not directly relate to the psychology of women are instructed to choose a research topic that relates to the agency's role in providing services for women or girls. The ultimate goal of the assignment is to strengthen students' knowledge of research in a particular area of interest and to critically engage the research in light of their civic engagement experience. Students are assessed on how well they are able to integrate their knowledge of course material with their civic engagement activity. Although students often report that completion of the volunteer experience, journal, and research paper are challenging, the majority of students recommend continuing to include civic engagement in future courses.

Students typically are challenged when their civic engagement experiences do not confirm stereotypes related to race, class, and gender. Students often enter their settings with preconceived notions about the women or children served by particular agencies. However, by the end of their experience, most are able to understand that many of the circumstances in which women find themselves transcend race, class, and gender. As a result, students are more knowledgeable and more culturally competent at the conclusion of their civic engagement activities.

*C. Behavior Modification.*

Behavior Modification is an upper level course offered by psychology departments with prerequisites such as Statistics and/or Research Design and Methodology. This course provides an in depth introduction to behavioral analysis. It is designed to help students learn how to target and measure behavior, develop a behavioral intervention program, and implement a research design using the basic principles of behavior modification. Students are responsible for completing a semester long behavior modification project as they volunteer in an underprivileged public school. Each student is required to become certified as a "Mentor" in the public school system because of the intensive involvement with elementary, middle or high school students. Students enrolled in the course also are required to obtain parental consent for their work with one or more children. Behavior Modification students identify a target behavior that will enhance a child's functioning in the school setting. They conduct a functional assessment of the behavior to understand environmental factors (antecedents and consequences) impacting the target behavior. A behavioral intervention is designed and implemented in the classroom, baseline and intervention data is collected and discussions focus on lessons learned over the course of developing and implementing the intervention. The final week of the semester culminates in written papers and oral presentations to fellow classmates regarding their behavioral interventions in the school setting. Students require a great deal of oversight and encouragement as they attempt to apply behavioral principles to "real life" circumstances. Faculty who attempt this level of service-learning need to be prepared to meet with student fear and resistance to what initially appears to be a daunting task. The leap from discussion of principles in the classroom to application in the schools requires a level of problem solving that most undergraduate students have not been required to face. As the students settle into their projects, active class discussion returns to understanding basic concepts, such as "reinforcement", or "discriminative stimuli" that only an applied assignment can demand. As the semester nears an end, and projects are due, students generally experience the ultimate joy of accomplishing a truly challenging project. The best part of the semester is listening to students' oral presentations, as they describe the successes of their own students, as well as challenges faced while working in the classroom setting.

*D. Our Approach to Dealing with Ethical Issues.*

Due to the ethical concerns associated with incorporating service-learning into undergraduate psychology courses, faculty within our department agreed to implement a requirement that students complete a course in professional ethics in psychology prior to enrolling in a course involving a service-learning component. The ethics course originally was designed for this purpose and specifically addresses issues related to competence, privacy and confidentiality, informed consent/assent, and research involving human participants. This approach, however, is not without problems. As is a problem with any prerequisite requirement, there is the issue of course availability and sequencing. The ethics course is taught only once a year, which can be problematic. We have dealt with this issue in a number of ways. In some cases, we have allowed students to enroll in the ethics course as a co-requisite (simultaneous with the service-learning course) as opposed to a pre-requisite. In other cases, the faculty member teaching the service-learning course has incorporated ethics into the course curriculum/syllabus. For example, one semester, the faculty member who teaches the ethics course guest lectured in one of the service-learning courses. While this approach addresses the need to expose students to professional

ethics, we know that knowledge of ethical standards and principles does not always translate into ethical behavior.

In addition to formal instruction on ethics, faculty members incorporating service-learning into their courses do a number of things in an attempt to be proactive with regard to ethical behavior. First, the importance of working within one's competence is stressed to students. For example, when volunteering at a community agency, students are instructed to behave strictly within the roles defined as part of the service-learning requirement and not to act like they know more than they do. Second, we routinely address issues related to privacy and confidentiality by instructing students not to disclose identifying information about community members in class or in their coursework. For example, if there is a journal or other written course requirement, students are told to use first names only, initials, or pseudonyms to protect the identity of community members. Third, issues of consent/assent are dealt with individually depending on the setting and the requirement. For example, in the Child Psychology course described above where students are required to volunteer in an elementary school setting, the school shares responsibility for consent/assent. That said, students are instructed to disclose their role to all concerned and to respect individual autonomy, addressing issues related to both competence and consent/assent. Finally, students required to conduct research as part of a service-learning requirement must obtain approval through the university's institutional review board (IRB) and, where necessary, external IRBs (e.g., K-12 school system IRB, hospital IRB).

In closing, readers should recognize that the examples described in this section provide but a few of the possibilities that exist for incorporating service-learning into undergraduate psychology courses and, thus, promoting civic engagement among undergraduate psychology majors. For more examples, readers are referred to Bringle and Duffy (2006) who provide additional illustrations of how to integrate service-learning into psychology courses.

## VI. Conclusion.

In this article, we have attempted to demonstrate the utility of promoting civic engagement among undergraduate psychology majors via the inclusion of service-learning activities into relevant courses. As emphasized in this article, significant barriers exist, which include lack of administrative support, time constraints among faculty, and student resistance or complete lack of interest among students to engage in service-learning. However, with thoughtful planning, dedication and a commitment to the ideals of civic engagement, these challenges can be overcome. And, while unique ethical issues must be considered when incorporating service-learning activities into psychology courses, professional and ethical behavior among students, faculty, and community collaborators can be encouraged if stakeholders are willing to be proactive. Readers interested in learning more about this topic are encouraged to pursue the additional resources contained in Appendix A.

## Appendix A. Civic Engagement and Service Learning Resources

*Civic Engagement and Service Learning Journals*:

*Journal for Civic Commitment
*Journal for Civic Engagement
*Michigan Journal of Community Service-Learning

*Civic Engagement and Service Learning Websites:*

*International Partnership for Service Learning: http://www.ipsl.org/
*National Service-Learning Clearinghouse: http://www.servicelearning.org/

*Distance Learning and Civic Engagement/Service-Learning Resources:*

*Bennett, G., & Green, F.P. (2001). Promoting service learning via online instruction. *College Student Journal, 35*(4), 491-497.
*Dailey-Hebert, A., Donnelli-Sallee, E., & DiPadvoa-Stocks, L. (2008). Service-elearning: Educating for citizenship. Information Age Publishing: Charlotte, NC.

*Diversity and Civic Engagement/Service-Learning Resources:*

*Everett, K.D. (1998). Understanding social inequality through service learning. *Teaching Sociology, 26*(4), 299-309.
*Marullo, S. (1998). Brining home diversity: A service-learning approach to teaching race and ethnic relations. *Teaching Sociology, 26*(4), 259-275.

*Other Resources:*

*101 Ideas for Combining Service & Learning (separated by subject area). Available at: http://www.ahead.org/uploads/2008/handouts/block8/8.2%20-%20Handout%202%20-%20101%20ideas%20for%20SL%20by%20subject%20area.pdf
*Howard, J. (2001). Service-learning course design workbook. University of Michigan: Ann Arbor.
*Service-learning in psychology: A resource sheet. Available at: http://www.denison.edu/campuslife/servicelearning/psychology%20resource%20sheet%202009.pdf

## References

American Psychological Association. (n.d.). Civic engagement. Retrieved from http://www.apa.org/education/undergrad/civic-engagement.aspx.

American Psychological Association Education Directorate (2009). *Civic engagement and service-learning*. Washington, DC: American Psychological Association. Retrieved August 10, 2009, from http://www.apa.org/ed/slce/civicengagement.html

American Psychological Association. (2007). *APA guidelines for the Undergraduate psychology major.* Washington, DC: APA board of educational affairs task force on psychology major competencies. Retrieved from www.apa.org/ed/resources.html

American Psychological Association (2002). Ethical principles and code of conduct. Retrieved April 30, 2009 from http://www.apa.org/ethics/code2002.html.

Barrow, F.H. (2008). Who is in charge of this: Service-learning as a context for youth empowerment through policy participation in after-school program. *The Journal for Civic Commitment, 11*. Retrieved February 18, 2009, from http://www.mc.maricopa.edu/other/engagement/Journal/Issue11/Barrow.s html.

Bella, R.N., Madsen, R., Sullivan, W., Swidler, A, & Tipton, S.M. (1985). *Habits of the Heart: Individualism and Commitment in American Life.* Berkeley: University of California Press.

Bringle, R.G., & Duffy, D.K. (2006). *With Service in Mind: Concepts and Models for Service-Learning in Psychology.* Sterling, VA: Stylus Publishing.

Bringle, R.G., & Hatcher, J.A. (1996). Implementing service learning in higher education. *The Journal of Higher Education, 67*(2), 221-239.

Bulot, J.J., & Johnson, C.J. (2006). Rewards and costs of faculty involvement in intergenerational service-learning. *Educational Gerontology, 32,* 633-645.

Fenzel, L.M. (2001, April). Enhancing the benefits of service-learning in undergraduate psychology courses. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Fish, S. (2003, May 16). Aim Low. *Chronicles of Higher Education,* p. C5.

Goodman, R.D., & West-Olatunji, C. A. (2009). Engaging students in outreach as an extension of advocacy to achieve praxis. *The Journal for Civic Commitment, 12*. Retrieved February 18, 2009, from http://www.mc.maricopa.edu/other/engagement/Journal/Issue12/Goodman.html.

Grier-Reed, T., Detzner, D., Poch, R.K., & Staats, S. (2010). Toward multicultural community engagement. *Academic Exchange Quarterly, 14*(3). Retrieved from http://rapidintellect.com/AEQweb/cho4671.htm

Grimes, T.R. (1998, August). Creating opportunities for service-learning through the applied psychology course. Paper presented at the Annual Convention of the American Psychological Association, San Francisco, CA.

Handelsman, M.M., & Rosen, J.A.A. (1987). Informed consent of students: How much information is enough? *Teaching of Psychology, 14* (2), 107-109.

Hurtado, S. (2007). Linking diversity with the educational and civic missions of higher education. *The Review of Higher Education, 30*(2), 185-196.

Hyde, J.S. (2007). *Half the human experience: The psychology of women* (7th ed.). Boston: Houghton Mifflin.

James, W. (1910, August). The moral equivalent of war. *McClure's Magazine,* pp. 463-468.

Kenworthy-U'Ren, A. (2008). Creating paths forward for service elearning. In A. Dailey-Hebert, E. Donnelli-Sallee, & L.N. DiPadvoa-Stocks (Eds), *Service-eLearning: Educating for Citizenship*, (pp. 9-22). Charlotte, NC: Information Age Publishing.

Kretchmar, M.D. (2001). Service-learning in a general psychology class: Description, preliminary evaluation, and recommendations. *Teaching of Psychology, 28*(1), 5-10.

Laird, T. F. N., Engberg, M. E., Hurtado, S. (2005). Modeling accentuation effects: Enrolling in a diversity course and the importance of social action engagement. *The Journal of Higher Education, 76*(4), 448-476.

Laird, T. F. N., Engberg, M. E., & Hurtado, S. (2002). Modeling the effects of a diversity course on students' preparation for a diverse democracy. *Paper presented at the Annual Meeting for the Study of Higher Education*. Sacramento, CA.

Long, A.B., Larsen, P., Hussey, L., & Travis, S.S. (2001). Organizing, managing, and evaluating service-learning projects. *Educational Gerontology, 27,* 3-21.

McDonald, T.W., Caso, R., & Fugit, D. (2005). Teaching and learning operant principles in animal shelters: Perspectives from faculty, students, and shelter staff. *Journal of Instructional Psychology, 33*(4), 310-321.

Mio, J.S., Barker, L.A., & Tumambing, J. (2009). *Multicultural Psychology: Understanding our Diverse Communities* (2nd ed.). New York: McGraw-Hill.

Nigro, G., & Wortham, S. (2006). Service-learning through action-research partnerships. In R.G. Bringle & D.K. Duffy (Eds), *With Service in Mind: Concepts and Models for Service-Learning in Psychology*, (pp. 161-170). Sterling, VA: Stylus Publishing.

Osborne, R.E., & Renick, O. (2006). Service-learning. In W. Buskist & S.F. Davis (Eds), *Handbook of the Teaching of Psychology,* (pp. 137-141). Malden, MA: Blackwell Publishing.

Plante, T.G. (2009). Five ethical values to guide professional behavior. In T.G. Plante, *Spiritual practices and psychotherapy: Thirteen tools for enhancing psychological health*, (pp. 103-114). Washington, DC: American Psychological Association.

Prentice, M. (2007). Service-learning and civic engagement. *Academic Questions, 20* (2), 135-145.

Rocheleau, J. (2004). Theoretical Roots of Service-Learning: Progressive Education and the Development of Citizenship. In B.W. Speck and S.L. Hoppe (Eds.), *Service-learning: History, Theory, and Issues* (pp. 3-21). Westport, CT: Praeger Publishers.

Rosenthal, G.T., McKnight, R.R., & Price, A.W. (2001). Who, what, how, and where the typical psychologist is…the Profession of Psychology scale. *Journal of Instructional Psychology, 28* (4), 220-224.

Ward, K. (1996). Service-learning and student volunteerism: Reflections on institutional commitment. *Michigan Journal of Community Service-learning, 3,* 55-65.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 76 – 85.

# From herb garden to wiki: Responding to change in naturopathic education through scholarly reflection

**Airdre Grant[1] and Meg O'Reilly[2]**

*Abstract: The study of naturopathy in Australia has undergone a number of significant changes that have impacted dramatically on scholarly and academic practice. Naturopathic education has shifted from apprenticeship style with a charismatic teacher, to several private colleges offering advanced diplomas/degrees with Government accreditation, and into the university system. In the university context, reflective practice and scholarship have been key to effective responses to change. These changes include the introduction of new models of technology and delivery, access to greater resources and collegial networks and adoption of innovative teaching approaches; all of which have had a significant impact on a developing pedagogy in naturopathy. Utilising a ethnographic methodology, interviews were conducted with ten naturopathy lecturers to investigate their scholarly reflective approaches to decision making and pedagogy. Some of the changes to academic practice within this context and the innovations arising are discussed. These innovations, whilst increasingly common in higher education, are new pedagogical approaches within the contexts of naturopathy, a discipline that traditionally has sat outside formal learning structures.*

*Keywords: Reflective decision making, integrative teaching practices, technology innovations*

## I. Introduction.

This paper reports on a Learning and Teaching Fellowship project at Southern Cross University that investigated how a shift from the private education sector into the higher education sector changed academic practice in the delivery of education in naturopathy Academic lecturers who had moved from the private sector to teach in the university program, reflected on their teaching and discussed the innovations they had integrated into their pedagogy as a result of the changing educational context for the discipline. This project was undertaken with approval of the University's Human Research and Ethics Committee.

### *A. Context.*

Complementary medicine is an umbrella phrase that covers a number of different modalities including naturopathy, osteopathy, aromatherapy and different forms of bodywork. During the last decade of the twentieth century and into the first decade of the twenty-first century, complementary medicine usage in Australia and worldwide expanded dramatically (Cohen, 2002). In Australia in particular, one in four members of the public now make use of some form of complementary health care product or service on a regular basis (MacLennan, Myers, & Taylor, 2006; MacLennan, Wilson, & Taylor, 1996, 2002). A surge in the popularity of

[1] School of Health and Human Services, Southern Cross University, airdre.grant@scu.edu.au
[2] Teaching and Learning Centre, Southern Cross University

complementary medicine meant that courses of study were increasingly sought after in the decade 1990–2000, and enrolment numbers increased sharply across a range of programs offered by the private colleges (McCabe, 2008). Since 2000, the climate has changed again – courses have closed and student numbers are declining, as the long-awaited professional registration has still not eventuated.

Historically naturopathic education existed outside formal education pathways. There is currently one university degree in naturopathy (as a major in a Bachelor of Clinical Science) and several Government accredited private colleges that offer an undergraduate degree in Australia. This movement into more formal structures has created a shift in academic practice (Baer, 2008). There is however, a paucity of educational research in naturopathy and this paper thus aims to contribute to knowledge of academic practice in the field.

*B. Current context for educational delivery.*

In general, naturopathy education in Australia is currently delivered via different awards. Students can gain either an Advanced Diploma of Naturopathy or a degree (e.g. Bachelor of Complementary Medicine, Bachelor of Naturopathy, and Bachelor of Health Sciences–Complementary Medicine). In 1995 Southern Cross University (SCU) in Lismore, New South Wales (NSW) launched a new four-year undergraduate Bachelor of Naturopathy (BNat) degree. At that time, this was an Australian first. For some time it was the only stand-alone degree of its kind in a university context. It is now subsumed into a Bachelor of Clinical Sciences with a major in naturopathy. Naturopathic curriculum typically has three primary pillars – (1) herbalism (sometimes known as phytotherapy), (2) nutrition, and (3) tactile therapies. It may also include (4) homoeopathy. Curriculum also covers (5) medical sciences – anatomy and physiology, chemistry and biochemistry, symptomatology and pathology, and electives such as reflexology, aromatherapy and other health related topics.

## II. Scholarly reflection and changing technologies.

Reflective practice, as defined by Schön (1983), is about a practitioner considering deeply their experience as a lecturer, and also being coached or having peer review. The opportunity to develop reflective thought is seen as critical with Chickering and Dalton (2006) noting this step as part of the process of integrating in one's thinking processes and the formation of beliefs. As Brookfield (2006) writes, the best teaching incorporates a critically reflective stance and the lecturers in this study were forced into reflecting on their teaching practice by changes in the teaching environment. They reflected on the increase of technology in their programs and impact of more formal and resourced teaching structures.

Technology-based learning and teaching in higher education is becoming almost a taken-for-granted proposition in many undergraduate courses (Adams, 2008; Pajo & Wallace, 2001). The pace of uptake across the sector is both unrelenting and brisk, with studies of all aspects – student performance, satisfaction, anxiety, motivation, participation, equity and lecturer approaches also proliferating (Alavi & Leidner, 2001; Allen & Seaman, 2004; Gray et al., 2009). At the same time, students are becoming more engaged with technology in both their personal and study lives (Chang, Kennedy, & Petrovic, 2008; Lefoe, O'Reilly, & Parrish, 2007; O'Reilly, Lefoe, Philip, & Parrish, 2008). In response to this drive from both students and the higher education sector, lecturers in the naturopathy course sought to integrate relevant, efficient, scalable and reusable technologies.

## III. Methodology.

The scholarly reflections of academic lecturers who taught in the naturopathy program were gathered using interviews and reflective prompts. The approach to the collection and interpretation of data for this investigation was *constructivist* in epistemology and *ethnographic* in methodology. Consistent with a constructivist epistemology, its focus is upon the "world of lived reality and situation-specific meanings" (Schwandt, 1994, p.118), or, in other words, it seeks to provide an understanding of the how naturopathic lecturers responded to the changes in their teaching environment. Consistent with an ethnographic methodology, it "involves first-hand intensive study of the features of a given culture and the patterns in those features" (Gall, Borg, & Gall, 1996, p.607).

This lens was chosen as a methodology of naturalistic inquiry, interviews were arranged and open-ended stimulus questions were designed. These stimulus questions about the BNat were:

1. Within the course in which you were involved, can you identify any changes/innovations you have introduced? For example: field trips, laboratories, observations, new delivery styles, assessment techniques, technological innovations.
2. What do you feel was innovative in the delivery of the BNat?
3. Can you also comment on the extent to which you have engaged in or been stimulated to engage in research in your field of interest?

Ten individual interviews with key academic lecturers from the disciplinary grouping of Natural and Complementary Medicine (NCM) were undertaken in 2009. Interviews were arranged by email, the Learning and Teaching Fellowship project was explained and semi-structured interviews, based on the three questions above, were conducted. Notes were taken and later transcribed for analysis and reporting purposes. A thematic analysis was undertaken and themes emerging from the transcripts were synthesised in order to determine the key experiences, concerns and achievements. What follows is a report on the experiences, concerns and achievements of the lecturers.

## IV. Findings – Case studies of changing academic practices.

A number of key themes emerged pertaining to reflective decision-making and pedagogically driven changes to a range of academic practices. We have chosen to present these themes as a series of modality based case studies, to illustrate changes in academic practice.

### *A. Herbal Medicine to wiki.*

There is large herb garden at SCU Lismore campus with a covered outdoor learning area (COLA). Classes are conducted where students see, smell and gather herbs for analysis in the chemistry laboratories. Wishing to capture students' insights gained from this hands-on garden work, the lecturer (also known as the unit assessor or UA) introduced the use of a wiki with one page on each herb as an easy to access *Materia Medica*. Using this wiki, students were required to contribute information from the literature and also from their practical observation. To encourage participation the lecturer structured sections for introductory information, additional resources and a glossary. Students contributed according to their nominated area of interest. For example one student, who developed an interest in the herb *senna* and its use as 'ballerina tea' for weight loss, researched the topic and was the prime contributor to the wiki page. Other

students were able to comment on these entries and all contributions – the original students' and any follow-up comments were visible and easily tracked by the UA. A social bookmarking tool was also introduced to assist students in sharing and collecting web resources. The UA's comments illustrate her reflections on the innovative use of the wiki:

> *It appeared that the students were intimidated if the wiki was too pretty and were reluctant to get involved in it as they thought their work may not be good enough. I reworked the assessment so they had to begin at the beginning. The students had to construct the whole thing and work on creating their own pages the way they wanted to without a construct from me (Herbal Medicine UA, 2009).*

Another significant change to practice highlighted by the UA is that when herbal medicine entered the tertiary sector, naturopathy was located in the health sciences. However, herbal medicine has a long history that is deeper and stronger than folklore and pharmaceuticals. The UA thus designed the unit to align the curriculum with its traditions by attaching herbal medicine to the discipline of history. This allowed students to develop knowledge along lines that are more complex than solely adopting a science base. The connection with historical traditions proved useful when the UA was able to make collegial connections with the Smithsonian Institute in Washington DC in the USA. As a consequence of this connection students benefited from exposure to guest lecturers from the Smithsonian Institute. These lectures were recorded as podcasts. The lecturer said of this development:

> *We have a herb garden we can conduct classes in the outdoor learning area and that's been great; it's like being immersed in the herbs. But when colleagues from the Smithsonian Institute (in Washington DC) came to the Uni and gave talks on Socrates as part of their work as historians, suddenly the students were able to see how the growth of herbal medicine has been one long continuum. Up until then I think they had just been thinking about the herbs, now they saw them as part of the history of medicine… and here were two historians talking to them about ancient texts and healing practices. It was a wonderful moment for me as a teacher (Herbal Medicine UA, SCU, 2009)*

*C. Quizzes and podcasts for Nutrition.*

Classes in nutrition are shared with other disciplinary groupings in the university and so naturopathy students naturally mix with students from exercise science, osteopathy and education. Furthermore, the introductory unit in nutrition is a unit available as an elective across all undergraduate courses, and therefore caters for a large cohort of internal and external students. The UA sought assistance from production staff to develop a unit that was consistent across all cohorts and campuses, and which utilised a number of technologies for a broad range of participants, as well as being easy to manage. This has meant changing integrating a range of technologies for flexible learning. The unit now includes wikis for student collaborative work and podcasts of recorded lectures, as well as synchronous audio based interaction online supplemented with the use of an in-built platform for sharing applications (including PowerPoint™, websites, whiteboard and specialist software for dietary nutrient calculations). The UA also identified a need to stimulate reflective practice as part of the student learning, noting:

> *I wanted to design an assessment item that stimulated deeper thought through reflective practice, [and] that encouraged students to look into the literature and see how it informed their study of nutrition, so they did more than simply learn how to calculate food values. I saw it as important to stimulate sustained engagement with the materials. Online quizzes combined with a food diary promoted that engagement (Nutrition UA, 2009).*

Audiographic software was used as a teaching tool to reach both external and internal students that, combined with use of podcasts, enabled them to access recorded lectures and go over topics they may have missed or wanted to study further. The technology reinforced student learning, supported successful completion of assessment tasks and extended the reach of the unit. The UA critically evaluated her innovations and concluded that pedagogically, the changes to requirements for student reflection have added depth to the learning process as well as enabling access to a geographically widespread cohort.

*D. Vodcast, blogs, peer review for Tactile Therapies.*

The teaching of tactile therapies primarily requires a hands-on approach. Students learn massage routines through observation and usually practice with a partner in class in a purpose-built classroom that has massage tables and is screened for privacy. In response to student requests for more practice time and thus to facilitate learning outside practical classes, the UA created vodcasts whereby techniques were demonstrated and uploaded to the online unit site. These could be downloaded to an mp3 player for self-paced skill development, revision and reinforcement. An online workbook was also developed containing interactive exercises to promote group work. Reflective practice was developed through use of a weekly blog and participation in the online discussion board. Other assessments such as written case studies, practical examinations and demonstrating an understanding of theory required students to engage with critical thinking and best practice application. Students also participated in supervised practicum work in the university student clinic. There were opportunities for students to work as trainee tactile therapists in the community, not only in places such as hospices and rehabilitation units but also during University-organised community events.

The UA has completed the Graduate Certificate in Higher Education (L&T) and in so doing found the practice of peer review of teaching particularly beneficial to trigger reflective practice and to stimulate a scholarly approach to teaching:

> *A colleague from Teaching and Learning came to my classes and observed and offered feedback. I found it tremendously useful. It validated my practice and the constructive feedback informed me on how to enhance the student learning experience. Up until then I had just taught as best I knew how. This really helped me improve my teaching practice (Tactile Therapies UA, 2009).*

*E. Using software in Homoeopathy.*

The subject of homoeopathy can sit uneasily within a university curriculum, as there exists lively debate regarding its validity and efficacy. Homoeopathy has a tradition of being taught along lines that have not shifted a great deal since its conception in 1799 by the German physician Samuel Hahnemann. A major push within homoeopathy teaching was the use of assessment tools

that promote philosophical and reflective practice – key skills required of practitioners. A specialist homoeopathy software program was installed at the university teaching clinic so students could learn to use a tool that made repertorising the *Materia Medica* both swift and efficient (repertorising refers to the calculation of an appropriate remedy through selective symptom matching, which is often very complex and time consuming). The software package was used both in the classroom and student clinic.

*F. Online cases studies and scenario based learning in Medical Sciences.*

The medical sciences stream included subjects such as anatomy, physiology, chemistry and biochemistry, symptomatology and pathology. Lecturers were active in seeking innovations to enhance engagement with the often dense material of medical sciences. Classes were delivered in fully set-up laboratories designed for use across five undergraduate programs in Health and Human Sciences. Anatomy classes involved not only the use of models as is typical, but were also supplemented by 3D graphic computer based activities and optional activities within the cadaver lab (wet lab). Students worked through activities in rotation with the computer based activities reinforcing the learning from anatomical models and wet lab specimens. Lecturers met together to discuss strategies to make student wet lab engagement more meaningful, to cater effectively to the high number of students rotating through the facility. A meeting like this may not necessarily have happened in private colleges where lecturers attend for their classes only and are usually not employed on a full time basis or located on site.

Physiology classes were delivered using an audiographic interface in order to provide an equivalent experience to students on each of the university's three main campuses. These media rich lectures were also recorded for self-paced learning and revision. Assessment was conducted through a series of online exams and laboratory tests. Similarly, chemistry and biochemistry classes incorporated hands-on experiments and used a series of online quizzes to reinforce key foundation knowledge.

Symptomatology and pathology were taught through two inter-related units. The design of these has changed significantly over time and a detailed report can be found in a previously published paper (O'Reilly & Wojcikowski, 2008). In brief, the units use case based scenarios in a lock step fashion requiring students to submit responses in a series of related steps. They have in the past been provided with feedback on each stepwise submission using a feature of Blackboard™ called Adaptive Release™. From the student feedback and UA's ongoing reflections on how to improve teaching, changes were made for less in-class tutorial hours and more on-line delivery through interactive self-paced case scenarios. The latest version of the unit was designed with Scenario Based software. As a Scenario Based Learning (SBL) design, this is as close to an authentic practitioner caseload as possible and is intensively engaging for students. Design of these units now also facilitates a more automated marking process. As a result of a coherent and scholarly approach to enhancing the student learning via technology-supported engagement, this academic was awarded a VC's citation for excellence in teaching.

*G. Community-based, case-based learning and computer-based clinical decision making in Clinic and Practicums.*

As part of the medical sciences curriculum, students embark upon case based learning from their 3rd year of study and are also introduced to specialties through guest lecturers such as pediatricians, pharmacists and mental health specialists. The pedagogical design aims to enhance

diagnostic knowledge that is further developed through liaison with local hospitals. Naturopaths are typically excluded from the public health system however the negotiation by the Head of Nursing increased interaction between the university and hospitals in the area, and naturopathy students were enabled to go into palliative care units, rehabilitation units and aged care facilities to give massages. They were also encouraged to work in Aboriginal Health Care units. This increased their opportunity to see other health professional interacting with patients as well as strengthening collegial networks. As part of a commitment to promote collegial relationships and enhance learning, SCU also established an agreement with Bastyr University in Washington State (USA) that allows for exchange for 4th year students.

These opportunities exposed students to the training structures of other institutions. It also brought them into contact with the array of technologies used in the health sector and introduced them to the possibility of incorporation into their own practice. As one clinic supervisor noted:

> *Students are able to make the most of the programs available to augment decision making. For example there is a program called HyperHealth ™ that helps with differential diagnosis. The Natural Standard™ is another program which, when students type in a condition, offers suggestions for vitamin and mineral supplements that they can use to inform their case studies. We have CDs of herbal and nutritional compendiums which they can access… These programs do not replace their decision making. We teach them how to use them wisely and also the importance of keeping knowledge current. They get to see that good quality up to date data is easily available (Clinical Supervisor, 2009).*

The UA for clinical education chose to enroll in the Graduate Certificate in Higher Education (L&T) in order to explore and improve teaching practice. This opportunity to improve pedagogical approaches was seen as significant:

> *I undertook the study in the Graduate Certificate in Higher Education (L&T), and used my study as an opportunity to do a literature review on assessment strategies in clinical assessment. This really informed my practice as Clinical Director and I changed the assessment protocols at the clinic so that testing was more thorough and tested knowledge on all levels. I introduced Viva Voce examinations and problem based learning so that the learning for students was real, immediate and relevant early on (Clinic Director, 2009).*

*H. Enhanced PowerPoint presentations for Grand Rounds.*

Clinic, practicums and Grand Rounds together form the core practical components of the medical sciences curriculum. Students in their 4th year present assessable case studies to peers, lecturers and invited guests. In this way, students practised presentation skills as well as sharing reflections and insights about practice strategies with peers. This teaching model of Grand Rounds was borrowed from medicine and initiated by the Professor of Natural and Complementary Medicine who holds both medical and naturopathic qualifications. Students also performed role plays and used a number of technologies including PowerPoint™, and IT tools such as hyperlinks within their presentation to provide snapshot videos to enhance their

presentations. Peer review was incorporated into the formal assessment as a strategy to promote student engagement, critical thinking skills and scholarly reflective practice.

## V. Discussion.

The Learning and Teaching Fellowship project sought to inquire about the innovations in teaching practices, technology led and otherwise, and to investigate the scholarly reflective decision making activities of lecturers. Whist some of the innovations are standard practice in higher education, integration with the formal higher education/training sector brought change to pedagogy in naturopathic education. These changes were impacted by greater access to resources, exposure to collegial exchange across a range of disciplines and engagement with self-regulatory systems.

At SCU lecturers from the private sector now had access to resources that had previously been largely unavailable to them, such as a large university library with reference librarians, chemistry and computer laboratories. Employment was now primarily on continuing basis thus affording greater continuity, support and availability to students (employment on a sessional basis is generally the case in the private colleges). As academics, naturopathic lecturers were expected and supported to upgrade their qualifications and to embark on research and scholarship and to participate in course and program reviews routinely. Opportunities for collegial exchange, professional development and support for conference participation now existed. The affordances of these resources meant that lecturers were now drawn into a model of supported academic practice and encouraged to develop their scholarly outputs.

Lecturers had to respond to issues such as the resourceful use of educational technology, a diverse student cohort, cross campus teaching, governance issues such as course and program reviews, enforced deadlines consistent with institutional prerogatives, incorporating university Graduate Attributes and the ongoing requirements to upgrade qualifications. This mix of requirement and challenges had impact on the way naturopathy was taught. On the one had it introduced a range of resources and a degree of standardisation, on the other hand the intimacy of the 'cottage industry' was felt by some to be a loss.

The pluses were that the opportunities within the university to advance qualification through engagement with further study, through submissions for grants, awards and/or scholarly publications were all seen as positives by lecturing staff.

## VI. Conclusion.

The findings from this study point to the reflective practices that academics undertook about the introduction of new technologies, innovations and pedagogies in naturopathy. All the naturopathy lecturers interviewed expressed that they had gone through significant changes in their teaching practice as a result of the changes in delivery for the subjects and their exposure to a more involved educational system. This reflective process impacted upon their academic practice as they underwent a process of professional upheaval and reshaping of professional practice. Further investigation into academics' scholarly reflective approaches to decision making, and teaching in the discipline indicated that some chose to pursue postgraduate educational studies and others also deepened their inquiry into disciplinary pedagogies. These changes and shifts in pedagogy have proved positive in bringing the educational culture of naturopathy more into mainstream. The benefits have been more access to resources and collegial inquiry and support, along with upgrading of the skills of the lecturers and deepening of

their professional commitment to excellence in education for neophyte naturopathic practitioners. This can only have a positive impact on the profession, which being unregistered, is vulnerable to criticism from society and other allied professions who may view complementary medicine with caution and reserve.

## Acknowledgements

The authors wish to acknowledge the opportunity provided by Southern Cross University Vice-Chancellor's Learning and Teaching Fellowship and thank colleagues for evidence and reflective data provided. Associate Professor Gail Wilson provided valuable feedback.

## References

Adams, J. (2008). Understanding the factors limiting the acceptability of online course and degrees. *International Journal on E-Learning, 7*(4), 573–587.

Alavi, M., & Leidner, D. (2001). Research commentary: Technology-mediated learning – A call for greater depth and breadth of research. *Information Systems Research, 12*(1), 1–10.

Allen, I., & Seaman, J. (2004). Entering the mainstream: the quality and extent of online education in the United States 2003 and 2004. *Available at* http://www.sloan-c.org/publications/survey/pdf/entering_mainstream.pdf,

Baer, H. (2008). The growing legitimation of complementary medicine in Australia: success and dilemmas. *Australian Journal of Medical Herbalism, 20*(1), 5–11.

Brookfield, S. (2006). *The Skillful Teacher: On Trust, Technique and Responsiveness in the Classroom* . (2nd ed.). San Francisco: Jossey-Bass.

Chang, R., Kennedy, G., & Petrovic, T. (2008). *Web 2.0 and user-created content: Students negotiating shifts in academic authority*. Paper presented at the Hello! Where are you in the landscape of educational technology? Proceedings ascilite.

Chickering, A.W., & Dalton, J.C. (2006). *Encouraging Authenticity and Spirituality in Higher Education*. San Francisco: Josey-Bass.

Cohen, M. (2002). New professor, new department, new journal, new era in medicine *Journal of Complementary Medicine, 1*(2).

Gall, M.D., Borg, W.R., & Gall, J.P. (1996). *Educational research: An introduction* ( 6 ed.). White Plains, NY: Longman.

Gray, K., Kennedy, G., Waycott, J., Dalgarno, B., Bennett, S., Chang, R., et al. (2009). *Educating the Net Generation*. Strawberry Hills: Australian Learning and Teaching Council.

Lefoe, G., O'Reilly, M., & Parrish, D. (2007). *Community, exchange and diversity: The Carrick Exchange.* Paper presented at the In Proceedings of the 24th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education, Singapore.

MacLennan, A.H., Myers, S.P., & Taylor, A.W. (2006). The continuing use of complementary and alternative medicine in South Australia: Costs and beliefs in 2004. *Medical Journal of Australia, 184*(1), 27-31.

MacLennan, A.H., Wilson, D.H., & Taylor, A.W. (1996). Prevalence and cost of alternative medicine in Australia. *Lancet, 347*(9001), 569-573.

MacLennan, A.H., Wilson, D.H., & Taylor, A.W. (2002). The escalating cost and prevalence of alternative medicine. *Preventive Medicine, 35*(2), 166-173.

McCabe, P. (2008). Education in naturopathy and western herbal medicine in Australia: Results of a survey of education providers. *Complementary Therapies in Clinical Practice, 14*(3), 168-175.

O'Reilly, M., Lefoe, G., Philip, R., & Parrish, D. (2008). The Carrick Exchange for Higher Education: Design factors for user engagement' In J. Luca & E. Weippl (Eds.), *Proceedings of EdMedia 2008 World Conference on Educational Multimedia, Hypermedia and Telecommunications* (pp. 221–230). Vienna: Association for Advancement of Computers in Education.

O'Reilly, M., & Wojcikowski, K. (2008). *Clinical Diagnosis Online.* Paper presented at the "Hello! Where are you in the landscape of educational technology?" Proceedings ascilite Melbourne.

Pajo, K., & Wallace, C. (2001). Barriers to the uptake of web-based technology by university teachers. *The Journal of Distance Education, 16* (1), 70-84.

Schön, D. (1983). *The Reflective Practitioner*. London: Arena, Ashgate.

Schwandt, T.A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. L. Denzin, Y.S. (Ed.), *Handbook of Qualitative Research* (pp. 118). Thousand Oaks, CA: Sage

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 86 – 102.

# Enhancing students' scientific and quantitative literacies through an inquiry-based learning project on climate change

**Aaron M. McCright**[1]

*Abstract: Promoting sustainability and dealing with complex environmental problems like climate change demand a citizenry with considerable scientific and quantitative literacy. In particular, students in the STEM disciplines of (biophysical) science, technology, engineering, and mathematics need to develop interdisciplinary skills that help them understand the social dynamics of environmental problems and solutions. To this end, this study examines how participation in a semester-long inquiry-based learning project that involves sociological research on climate change beliefs, attitudes, and behaviors enhances the scientific and quantitative literacies of STEM students. The results suggest that participation in a sociological inquiry-based learning project helps STEM students to (a) improve their knowledge of scientific and statistical principles and processes, (b) hone their scientific research skills, and (c) gain respect for sociology specifically and social science more generally. While the inquiry-based learning project described here deals with climate change, educators can adapt it to deal with other environmental social science research topics (e.g., water use, energy conservation, food security, sustainability).*

*Keywords: scientific literacy, quantitative literacy, climate change, interdisciplinary learning, inquiry-based learning, survey research*

## I. Introduction.

Understanding the complexity of global environmental problems, such as climate change, and proposed solutions, such as sustainability, usually requires collaboration across disciplinary boundaries by a range of scholars and stakeholders who are scientifically and quantitatively literate. Within the general public, promoting sustainability and dealing with pressing environmental problems is likely to be more effective with a citizenry that is scientifically and quantitatively literate and supportive of the interdisciplinary work necessary to address and understand complex problems as well as support their solutions. All of this suggests that we need continued improvement in the education of university students, especially those in the STEM disciplines of (biophysical) science, technology, engineering, and mathematics who may apply their knowledge and skills to deal with environmental challenges in the future. Specifically, we need STEM majors who are scientifically and quantitatively literate and who are sufficiently literate in social science to collaborate with social scientists and better understand the social implications of their own technical work.

One way to achieve this is to increase the quantity and quality of experiences for university-level STEM students to conduct interdisciplinary, hands-on research on pressing environmental problems of our day. Such a strategy has an additional benefit of helping to retain strong students in the STEM disciplines—especially those from historically underrepresented groups (Alper, 1993; Uriarte et al., 2007)—since such hands-on projects increase students' interest in science and illustrate the application of science in their everyday lives (U.S. National Research Council, 2003, 2007).

[1] Lyman Briggs College, E-35 Holmes Hall, 919 E. Shaw Lane, Michigan State University, East Lansing, MI 48825, mccright@msu.edu.

To this end, this study addresses the following question. *What is the role of social science in the education of students in the STEM disciplines? More specifically, can participation in a semester-long inquiry-based learning project that involves sociological research on climate change enhance the scientific and quantitative literacy of STEM students?* Answering these questions is increasingly important not only for dealing with the complex environmental problems and solutions mentioned above, but also in the short term for promoting interdisciplinary teaching and learning on university campuses (e.g., Chandramohan & Fallows, 2009; Klein, 2010) and amid mounting pressure to provide evidence of the value-added of a broad liberal arts curriculum (e.g., Arum & Rokso, 2011; U.S. Department of Education, 2006).

Drawing upon existing social science education research and specifically employing insights on inquiry-based learning, this study empirically examines how participating in authentic, inquiry-based sociological research about climate change beliefs, attitudes, and behaviors may improve STEM students' scientific and quantitative literacy by helping them see how basic research processes (e.g., research design principles, sampling, measurement, and statistical analyses) span different disciplines. This project was administered in a junior-level course on environmental social science in Michigan State University's Lyman Briggs College (LBC), a residential learning community for the study of science and society. Begun in 1967, LBC offers STEM students an interdisciplinary curriculum that integrates the natural sciences with the social sciences and humanities to deepen their understanding of science.

After briefly reviewing the relevant social science education research in the next section, the third section details the research design, measurement instruments, student samples, procedures, and analytical techniques employed. After describing the course and the semester-long inquiry-based learning project the students completed, the fourth section then presents the results of the quantitative analyses of pretest and posttest data. The final section discusses this study's limitations and implications, as well as the promise of similar inquiry-based learning projects on other environmental topics for preparing STEM majors to effectively work on such complex environmental problems and solutions in the future.

## II. Relevant Scholarship.

With a general shift in higher education from focusing on teaching (i.e., what instructors do) to focusing on learning (i.e., what students do) (e.g., McKinney, 2008; Svinicki, 1999), faculty and administrators increasingly promote active learning in which students learn by doing something (e.g., Mohamed 2008; Pedersen, 2010) and collaborative learning in which students learn by working with their peers (e.g., Johansen, Scaff, & Hargis, 2009; Parrott & Cherry, 2011).

Social scientists have long utilized active learning strategies and techniques in their university courses. At the most general level, this includes different ways of facilitating class discussion (often of controversial topics) (e.g., Yamane, 2006). Social scientists also employ other active learning techniques such as role-playing (e.g., Teixeira-Poit, Cameron, & Schulman, 2011), task exercises (e.g., Nguyen & Trimarchi, 2010), and simulations (e.g., Steck et al., 2011). Further, social scientists often promote collaborative learning via group research projects (e.g., Longmore, Dunn, & Jarboe, 1996) and problem-based learning activities (e.g., Ross & Hurlbert, 2004).

Given the call for more education research in social science (Albers, 2008; Whitman et al., 2007), the call for more inquiry-based learning in social science (e.g., Atkinson & Hunt, 2008), and the increasing importance of social science for understanding our complex environmental problems and their solutions, my interest here is to examine the role of sociological inquiry-based learning in promoting scientific and quantitative literacy (see Rusche & Jason, 2011). Most broadly, inquiry-based learning follows an apprenticeship model whereby students perform manageable versions of authentic research with the guidance of the course instructor (Vajoczki et al., 2011). Inquiry-based learning in the STEM disciplines—especially in laboratory courses—increases students' scientific literacy (e.g., Brickman et al., 2009; Derting & Ebert-May, 2010; Handelsman et al., 2004). Social scientists utilize inquiry-based learning

projects that last for a few class meetings or span the entire semester (e.g., McKinney & Busher, 2011; Raddon, Nault, & Scott, 2008).

Parallel to, and often integrated with, the promotion of inquiry-based learning is the promotion of quantitative or statistical literacy within the social science curriculum (e.g., Grawe, 2011; Lindner, 2012; Wilder, 2009). Indeed, several studies examine how the analysis of available data may increase students' quantitative literacy (e.g., Burdette & McLoughlin, 2010; Himes & Caffrey, 2003). Further, demanding that students formally present posters describing their research may help them hone their scientific communication skills (e.g., Levine-Rasky, 2009).

Since effectively dealing with the complex environmental problems and solutions of our day calls for STEM majors who are scientifically and quantitatively literate and have at least basic social science skills, this study empirically examines the effectiveness of participation in a sociological inquiry-based learning project for increasing the scientific and quantitative literacy of STEM students. Given the importance and salience of global environmental problems, the STEM students participating in this inquiry-based learning project investigated the climate change beliefs, attitudes, and behaviors of other college students.

## III. Methods.

### *A. Research Design.*

The participants in this study are juniors and seniors in Michigan State University's Lyman Briggs College who have completed similar introductory STEM courses in mathematics, chemistry, and biology as well as a freshman-level course introducing them to basic concepts, theories, and methods in the history, philosophy, and sociology of science (HPS). Some students take a year of physics in their first two years, though most of these students—especially those in the popular majors of human biology and physiology—take their physics courses as juniors or seniors. Their STEM courses have smaller enrollments, allow for greater faculty-student interaction, and involve more active and collaborative learning than do most STEM courses in MSU's College of Natural Science.

This study utilized a quasi-experimental design, which is illustrated in Figure 1. In essence, the design is a compromise between a one-group pretest-posttest design and an untreated control group design with pretest and posttest (e.g., Cook & Campbell, 1979). It is essentially a one-group pretest-posttest design with a posttest measure from an untreated control group. Briefly, pretest and posttest data were gathered from the students in the experimental group: the 27 students in my "The Natural Environment: Perceptions and Practices" course who participated in the sociological inquiry-based learning project, which is described in section 3.3). Posttest data also were gathered from a control group of 130 students in five other similar HPS courses: "Topics in History, Philosophy, and Sociology of Science," "Technology and Culture," two sections of "Topics in History of Science," and "Science Technology, and Public Policy." Limited course time and logistical reasons precluded gathering pretest measures from the control group prior to the beginning of the semester.

There is no reason to believe there is a systematic difference in the scientific and quantitative literacy of students in the experimental group (enrolled in my course) and of those in the control group (enrolled in one of the five similar HPS courses). All of these 157 students signed up for their courses approximately nine months earlier, well before this education research project was designed. Prior to the first day of classes, none of the students in the experimental group knew they would take part in the inquiry-based learning project. Also, because the LBC curriculum is quite constrained, most LBC students have taken similar STEM classes at similar times over their academic career. Enrollment in all six courses varied between 27 and 31; all six courses had similar workloads in terms of weekly readings, classroom activities, graded essay exams, etc. An examination of the pretest measures for the experimental group and the posttest measures for the control group reveals that students in both groups had the same level of sophistication in their perceptions of science (not reported here). Though, the

| | Pre-Test | Intervention | Post-Test |
|---|---|---|---|
| **Experimental Group**<br>27 students in "The Natural Environment: Perceptions and Practices" who participated in the sociological inquiry-based learning project | $O_1$ | X | $O_2$ |
| **Control Group**<br>130 students in five courses: "Topics in History, Philosophy, and Sociology of Science," "Technology and Culture," two sections of "Topics in History of Science," and "Science Technology, and Public Policy" | | | $O_2$ |

**Figure 1. The Quasi-Experimental Design of the Study.**

students in the experimental group seemed to have a slightly stronger belief that statistics will help them understand science than did those in the control group. Finally, the students in the experimental group began the semester with a relatively high level of confidence in their ability to understand science and statistics.

The pretest and posttest surveys contained groups of items measuring perceptions of scientific principles, attitudes toward the social sciences and statistics, self-assessment of scientific and statistical skills, and assessed knowledge of scientific and statistical processes. The surveys were identical with one major exception. The wording or numbers in the items used to measure scientific and quantitative literacy were changed slightly so the students in the experimental group had to make new calculations or judgments to select the correct answer on the posttest survey (rather than merely responding with their answer from the pretest survey).

The pretest survey was administered online to the experimental group prior to the first day of class. The students were asked not to use any resources (e.g., books, the internet, friends, etc.) when completing the pretest survey. Examination of the responses on the pretest survey (and comparison of these with their responses on the posttest survey) strongly suggests that experimental group students answered the questions on this survey with no assistance. At any rate, had some students used resources to answer the questions (while being prohibited from using such resources on the in-class posttest survey), this would have made it more difficult to find that participation in the inquiry-based learning project influenced their scientific or quantitative literacy. The posttest survey was then administered to both the experimental group and the control group via written questionnaire during class time in the 12th week of the semester. Timing and logistical pressures prevented administration of the posttest survey at the end of the semester. This earlier administration of the posttest survey likely leads to an underestimation of the effects of participation in the inquiry-based learning project on the knowledge, skills, and attitudes indicators described below.

*B. Variables and Analyses.*

Table 1 lists the names, coding, means, and standard deviations of the variables used in this study. Three types of outcome variables were examined: knowledge, skills, and attitudes. Four of these variables are composite indexes, and three are single item measures. Nine items were used to create the "variables and statistical knowledge index" (Cronbach's Alpha=.50), which measures how much knowledge students have regarding basic scientific and statistical principles and processes. Briefly, the items ask students to distinguish independent from dependent variables, evaluate sampling procedures, distinguish inductive from deductive logic, distinguish quantitative measures from qualitative measures, identify univariate statistics, interpret a

correlation coefficient, and predict and interpret an R-squared value. Correct answers to the items were coded as 1 and all other responses as 0, so the resulting index ranges from 0 to 9.

The three skills variables asked students to either assess or rate their confidence in various scientific and statistical skills. The "interdisciplinary research skills self-assessment" variable is a single item that measures how much students agree or disagree with the following statement: "I have the skills to participate in a research project that integrates social science and natural science." The responses on this item range from 1 ("strongly disagree") to 5 ("strongly agree"). Six items were used to create the "scientific and statistical skills self-assessment index" (Alpha=.84), which measures how much weaker (1="mine is much weaker") or stronger (5="mine is much stronger") students believe their following skills are compared to other LBC students: understanding scientific methods; designing a scientific study; gathering scientific data; analyzing scientific data; performing statistical analyses; and interpreting the results of statistical analyses. This index ranges from 6 ("mine is much weaker" for 6 items) to 30 ("mine is much stronger" for 6 items). Also, Two items were used to create the "confidence in scientific/statistical skills index" (Alpha=.40), which measures how much students disagree (1="strongly disagree") or agree (5="strongly agree") with the following statements: "I feel confident in my ability to understand science" and "I feel confident in my ability to understand statistics."

The three attitudinal variables asked students about their thoughts regarding sociology specifically, social science generally, and statistics. The "perception of how scientific sociology is" variable is from a group of items that asked students to assess how scientific a range of ten endeavors are. These included core natural science fields such as physics and chemistry, core social science fields such as sociology, quasi-scientific fields such as medicine, and non-scientific fields such as astrology. The responses on this item range from 1 ("not at all scientific") to 5 ("completely scientific"). Three items were used to create the "favorable attitudes toward social science index" (Alpha=.64), which measures how much students disagree (1="strongly disagree") or agree (5="strongly agree") with the following statements: "Understanding the social sciences is crucial to my education," "I understand how the social sciences are related to the natural sciences," and "LBC students should receive more education in the social sciences." The resulting index ranges from 3 to 15. Finally, the "knowing statistics helps to understand science" variable is a single item that measures how much students disagree (1="strongly disagree") or agree (5="strongly agree") with the following statement: "I believe that learning statistics will help me better understand science."

In some of the statistical analyses described below, a key predictor variable and five control variables were also utilized. The key predictor variable is membership in the experimental group, coded 1 for "yes" and 0 for "no." Gender is coded 0 for male and 1 for female. Class standing ranges from 1 ("freshman) to 4 ("senior"), and cumulative GPA ranges from 1 ("less than 2.50") to 7 ("3.76 to 4.00"). A dummy variable named "completed college statistics course" measures whether students have completed (1) or have not completed (0) a junior-level statistics course for STEM majors. Finally, an Honors College membership dummy variable distinguishes those students who are members of the MSU's Honors College (1) from those students who are not members (0).

The analyses were conducted in two stages. In the first stage, paired-samples t-tests were used to analyze change on the seven knowledge, skills, and attitudes indicators across the experimental group's pretest and posttest responses. Finding statistically significant changes on these variables is reasonable evidence that participation in the sociological inquiry-based learning project had an effect. But, as an additional check in the second stage, several multivariate OLS regression models were then run to analyze variation in posttest values for students in the experimental group or in the control group. The results of analyses in this second stage provide greater confidence that whatever changes identified in the first stage were due to participation in the inquiry-based learning project and not simply a result of maturation or general education outside of the classroom over the course of the semester.

**Table 1.** Descriptive Statistics for Variables in the Study.

| Variable | Coding | Mean | SD |
|---|---|---|---|
| Knowledge | | | |
| variables and statistics knowledge index | 0 (9 incorrect answers) to 9 (9 correct answers) | 4.50 | 1.94 |
| Skills | | | |
| interdisciplinary research skills self-assessment | 1 (strongly disagree) to 5 (strongly agree: "I have skills to participate in a research project that integrates social and natural science") | 3.77 | .88 |
| scientific/statistical skills self-assessment index | 6 ("mine is much weaker" for 6 questions) to 30 ("mine is much stronger" for 6 questions) | 20.04 | 3.75 |
| confidence in scientific/ statistical skills index | 2 ("strongly disagree" for 2 questions) to 10 ("strong agree" for 2 questions) | 8.22 | 1.34 |
| Attitudes | | | |
| perception of how scientific sociology is | 1 (not at all scientific) to 5 (completely scientific) | 3.12 | .90 |
| favorable attitudes toward social science index | 3 ("strongly disagree" for 3 questions) to 15 ("strongly agree" for 3 questions) | 11.24 | 2.38 |
| knowing statistics helps to understand science | 1 (strongly disagree) to 5 (strongly agree: "I believe that learning statistics will help me better understand science") | 4.01 | .93 |
| Key Predictor | | | |
| experimental group | 0 (not in my course) to 1 (in my course) | .17 | .38 |
| Controls | | | |
| gender (female) | 0 (male) to 1 (female) | .48 | .50 |
| class standing | 1 (freshman) to 4 (senior) | 3.37 | .67 |
| cumulative GPA | 1 (less than 2.50) to 7 (3.76 to 4.00) | 5.38 | 1.41 |
| completed college statistics course | 0 (not yet taken/completed) to 1 (have completed) | .46 | .50 |
| Honors College membership | 0 (no) to 1 (yes) | .27 | .44 |

The means and standard deviations for these variables are for the posttest survey data (N=157).

*C. The Inquiry-Based Learning Project.*

The 27 experimental group students were enrolled in my environmental social science course titled "The Natural Environment: Perceptions and Practices." This is a discussion-based HPS course on Americans' values, beliefs, attitudes, and behaviors regarding the biophysical environment. Like the other HPS courses in Lyman Briggs College, this popular course is offered each semester and is regularly fully enrolled. During the semester in question, this HPS course focused specifically on the topic of climate change.

Throughout the semester, the students in this course read and discussed works on the social causes and effects of climate change and on the dynamics of climate change public opinion, media coverage, and politics. They also participated in a sociological inquiry-based learning project that asked them to investigate climate change public opinion at MSU. As a part of a collaborative team consisting of four or five students, they worked to answer an intriguing research question about the climate change attitudes, beliefs, and behaviors of students in MSU's three residential colleges. On the first day of class, the students were given the handout displayed in Table 2, which laid out in approximate chronological order the tasks that students would complete individually, in their teams, and as a class throughout the semester as part of the inquiry-based learning project.

In the first two weeks of the semester, the students explored their personal interests in climate change public opinion and quickly formed collaborative teams with like-minded peers to create a unique research question. The research questions for the six collaborative teams are displayed in Table 3. Between weeks three and five, the teams wrote and revised survey questions and made decisions about the overall structure and design of the survey. Between the sixth week and tenth week of the semester, we administered the survey online. Giving the respondents approximately four full weeks to respond allowed the students to achieve a sample size of 428 completed surveys. In the eleventh week of the semester, the students were led through a brief tutorial in basic univariate, bivariate, and multivariate statistics and in the SPSS software the students were using on their laptops in the classroom. During weeks twelve and thirteen, the teams analyzed their data and interpreted the results of their analyses. During weeks fourteen and fifteen, the teams prepared and revised their own research posters and reviewed the research posters of other teams. Finally, in the sixteenth week of the semester, the teams presented their research posters at the LBC Research Symposium.

At the beginning of the semester, I performed a more managerial role vis-à-vis the project; as the semester went on, I shifted to perform a more advisory role. Since these STEM students were going to do sociological research on a topic for which they had little expertise (i.e., climate change public opinion), I began the semester with an analogy that I was like the head of a climate change public opinion research group and they were its members who would carry out the actual work. I guided them toward manageable research questions and helped them as they struggled with the first draft of their survey items. Once the students became more proficient with the measurement of their concepts though, I transitioned to more of a consultant, giving advice and suggestions about the wording and ordering of survey questions as I do in any paid or pro bono consultancy. Thus, the students came to exercise a great deal of responsibility—tempered by my guidance and advice—for the creation of the entire survey, its administration, and the eventual analysis of their data.

## IV. Results.

Table 4 presents the results of the paired-sample difference-of-means t-tests within the experimental group to assess change in knowledge, skills, and attitudes over time between the pretest survey and the posttest survey. Compared to just prior to the start of the semester, students in the twelfth week of the inquiry-based learning project had greater knowledge of scientific principles and basic statistics—as seen in the values in the first row of the table. In other words, students' knowledge of scientific and statistical principles and processes significantly improved. Of the nine knowledge items, they went from answering approximately

**Table 2.** (Roughly) Chronological List of Tasks for Our Inquiry-Based Learning Project.

Keep the following conditions in mind as you read the list of tasks below:
- Each team's research question must deal with individuals' values, perceptions, beliefs, policy preferences, and/or behaviors with regard to climate change.
- We will create and administer an online survey via surveymonkey.com to students in Michigan State University's three residential colleges.
- We will analyze our survey data with SPSS statistical software.
- Each team will create a formal research poster to display the results of its analyses.

Those tasks you will do individually are identified with an "I," those you will do in your research team are marked with a "T," and those we will do together as a class are marked with a "C."

Project Preparation and Research Proposal Tasks
Explore your personal interest in one or more climate change topics (I)
Join a research team with a shared interest in an important climate change topic (I)
Create a compelling research question that your team will answer with data from our online survey (T)
Write up a brief research proposal as a team (T)

Survey Construction Tasks
Write the necessary survey questions you think should be included in the survey for you to have data to answer your team's research question (I, T)
Examine our pooled survey questions (C)
- Look for redundancies or contradictions
- Look for logical groups of similar survey questions
- Think about the logical ordering of the survey questions within a section
- Think about the logical ordering of the sections of survey questions

Create multiple revisions of our lists of pooled survey questions until we finalize the list of selected survey questions, the wording of these survey questions, and the ordering of these questions (C)
Write all the necessary directions for individual questions and sections of questions (C)
Write a brief introductory paragraph introducing the survey (C)
Pre-test the final version of the survey after it is posted online (I, C)
Create our numerical coding scheme (C)

Survey Administration Tasks
Gain permission from each residential college to use its e-mail distribution list (C)
Write all e-mail correspondences with our potential respondents: pre-notice e-mail; cover letter e-mail; first follow-up e-mail; second follow-up e-mail (C)
Identify ways to advertise our survey to increase our response rate: mailbox flyers; word of mouth; Facebook; classroom visits; student groups; etc. (C)

Data Analysis Tasks
Become familiar with some basis statistical analyses and our SPSS software (I, C)
Identify what survey questions your team will be working with for your research question (I, T)
Determine the expected relationships between your team's variables (I, T)
Run the appropriate statistical analyses to examine these relationships (I, T)
Work on interpreting the results of these statistical analyses (I, T)
Write a draft of your team's research poster (I, T)
Peer review another team's research poster (I, T)
Finalize your team's research poster (I, T)
Present your team's research poster at the Lyman Briggs College Research Symposium (T)

**Table 3. The Research Questions of the Six Collaborative Teams.**

| |
|---|
| Are men or women more likely to take climate change actions? Are there gender differences in the type of action, public or private, taken regarding climate change? |
| Does the amount of university-level science training influence knowledge of climate change health effects? |
| How does one's residential college influence one's view on climate change and health-related behaviors? |
| How does academic major influence perception of climate change? |
| Does one's preferred media source for climate change information influence one's understanding of and concern for climate change? |
| How do students' political views, both party affiliation and ideology, affect their perceptions of the economic and environmental risks associated with climate change? |

four items correctly before the semester to answering slightly more than six items correctly about three-quarters of the way through the semester. The students seemed to be on the verge of becoming proficient in their statistical analyses at that point of the semester, so later administration of the posttest survey likely would have revealed an even greater increase in their knowledge scores.

The middle section of Table 4 shows that students in the experimental group assessed their interdisciplinary research skills and their general scientific and statistical skills as significantly stronger in the twelfth week of the semester than just prior to the start of the semester. Yet, their self-reported confidence in their scientific and statistical skills did not change over time. The lack of change in the latter is at least partially due to a ceiling effect whereby experimental group students rated their confidence in their scientific and statistical skills as quite high on the pretest survey. The bottom section of Table 4 shows that students in the experimental group developed more favorable attitudes toward sociology and social science more generally, but participating in the sociological inquiry-based learning project did not affect their belief about the importance of knowing statistics for understanding science. Again, the average pretest response on the latter item suggests that there is a ceiling effect in play.

On their own, the results in Table 4 suggest that participation in the sociological inquiry-based learning project had a positive effect on students' scientific and statistical knowledge, skills, and attitudes (see, e.g., Burdette & McLoughlin, 2010; Himes & Caffrey, 2003). Yet, it might be that students' scores on these knowledge, skills, and attitudes measures generally increase over the course of the semester regardless of what class they are in and/or whether or not they participated in the sociological inquiry-based learning project. After all, this could occur as a result of maturation or general education outside of the classroom. Comparing the posttest survey responses of the students in the experimental group to those of similar students in similar courses (the control group) will allow us to determine with greater confidence if the significant changes in knowledge, skills, and attitudes in Table 4 are due to participation in the sociological inquiry-based learning project.

Table 5 reports the results of seven multivariate OLS regression models that examine variation in posttest values for students in either the experimental group (N=27) or in the control group (N=130). The entries are standardized coefficients to more easily compare the effects of the predictor and control variables. The key predictor variable is the dummy variable for experimental group membership. Each of the seven models displayed in Table 5 examine the effect of participating in the sociological inquiry-based learning project while controlling for the effects of gender, class standing, cumulative GPA, completion of a college-level statistics course, and Honors College membership.

According to the first model in Table 5, students in the experimental group had higher scores on the science and statistical knowledge index than did students in the control group. Combining this with the results in Table 4, it is reasonable to conclude that participation in the sociological inquiry-based learning project improved students' knowledge of scientific and

**Table 4.** Results of Paired-Sample T-Tests for Pretest/Posttest Differences on Knowledge, Skills, and Attitudes Indicators (N=27).

| Indicators | Pretest | | Posttest | | T Statistic |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | |
| **Knowledge** | | | | | |
| Variables and Statistical Knowledge Index | 3.96 | 1.48 | 6.37 | 1.82 | 6.24*** |
| **Skills** | | | | | |
| Interdisciplinary Research Skills Self-Assessment | 3.89 | .75 | 4.37 | .63 | 3.32** |
| Scientific/Statistical Skills Self-Assessment Index | 20.67 | 3.58 | 22.30 | 3.58 | 2.36* |
| Confidence in Scientific/Statistical Skills Index | 8.81 | 1.08 | 8.70 | 1.20 | -.41 |
| **Attitudes** | | | | | |
| Perception of How Scientific Sociology Is | 2.93 | .96 | 3.59 | .69 | 3.34** |
| Favorable Attitudes Toward Social Science Index | 10.59 | 2.14 | 12.30 | 1.56 | 3.82** |
| Knowing Statistics Helps to Understand Science | 4.15 | .77 | 4.26 | .66 | .68 |

* p<.05 ** p<.01 *** p<.001

statistical principles and processes. That is, the increase in this knowledge within the experimental group between the pretest before the semester started and the posttest in the twelfth week of the semester led these students' posttest knowledge scores to be significantly greater than those for the control group students. As one might expect, students in the Honors College also scored higher on scientific and statistical knowledge than did their non-Honors College counterparts, but there was no correlation between this knowledge and students' class standing, cumulative GPA, or experience with a college statistics course. Also, male students had slightly higher knowledge scores than did female students.

The next three columns in Table 5 deal with the three skills indicators. Students in the experimental group had higher posttest values on these three indicators than did students in the control group. Recall that the results in Table 4 showed that participation in the inquiry-based learning project seemed to have a positive effect on the experimental group students' assessment of their interdisciplinary research skills and of their scientific and statistical skills but no effect on their confidence in their scientific and statistical skills. The results in Table 5 mean that it is reasonable to conclude that participation in the sociological inquiry-based learning project improved the experimental group students' assessment of their interdisciplinary research, scientific, and statistical skills. The statistically significant positive coefficient on experimental group membership in the fourth model in Table 5 merely reflects the fact that experimental group students were more confident in their scientific and statistical skills in general than were control group students.

Class standing, cumulative GPA, and Honors College membership had no effect on any of the skills indicators. Those students who had completed a college-level statistics course

assessed their scientific and statistical skills as weaker than those of other LBC students and expressed lower confidence in these same skills as compared to those students who had not completed a college-level statistics course. Also, male students assessed their scientific and statistical skills as stronger than those of other LBC students and expressed greater confidence in these same skills as compared to female students. This latter finding is consistent with much sociology of science scholarship on the gendered pattern of skills self-assessment and confidence among university students (e.g., Committee on Science, Engineering, and Public Policy, 2007; Etkowitz, Kemelgor, & Uzzi, 2000).

Finally, the last three columns in Table 5 deal with the three attitudes indicators. Students in the experimental group had higher posttest values on these three indicators than did students in the control group. The results in Table 4 suggest that participation in the inquiry-based learning project had a positive effect on experimental group students' perception of how scientific sociology is and their favorable attitudes toward social science but no effect on their belief that knowing statistics helps them to understand science. The results in Table 5 mean that it is reasonable to conclude that participation in the sociological inquiry-based learning project improved students' attitudes toward sociology specifically and social science in general. The statistically significant positive coefficient on experimental group membership in the fourth model in Table 5 merely reflects the fact that experimental group students were slightly more likely to believe that knowing statistics helps them to understand science than were control group students.

As with the skills indicators, class standing, cumulative GPA, and Honors College membership had no effect on any of the attitudes indicators. Those students who had completed a college-level statistics course had less favorable attitudes toward social science and were less likely to believe that knowing statistics helps them to understand science compared to those students who had not completed a college-level statistics course. Also, male students perceived sociology to be more scientific than did female students, and the former were more likely to believe that knowing statistics helps them to understand science than were the latter.

**Table 5.** Standardized Coefficients from OLS Regression Models Predicting Knowledge, Skills, and Attitudes Indicators (N=157).

| | Knowledge | Skills | | | Attitudes | | |
|---|---|---|---|---|---|---|---|
| **Independent Variables** | **Variables and Statistical Knowledge Index** | **Interdisciplinary Research Skills Self-Assessment** | **Scientific and Statistical Skills Self-Assessment Index** | **Confidence in Scientific and Statistical Skills Index** | **Perception of How Scientific Sociology Is** | **Favorable Attitudes Toward Social Science Index** | **Knowing Statistics Helps to Understand Science** |
| experimental group | .51*** | .34*** | .33*** | .22** | .34*** | .22** | .17* |
| gender (female) | -.16* | -.10 | -.35*** | -.28*** | -.22** | .01 | -.18* |
| class standing | .02 | -.03 | .01 | -.09 | .06 | .05 | .01 |
| cumulative GPA | .06 | .04 | .15 | .17 | -.13 | .02 | -.01 |
| completed college statistics course | -.13 | -.07 | -.15* | -.17* | .13 | -.16* | -.17* |
| Honors College membership | .21*** | .06 | -.01 | -.05 | -.01 | .07 | .12 |
| adjusted $R^2$ | .26 | .08 | .19 | .11 | .18 | .04 | .04 |

* p<.05 ** p<.01 *** p<.001

## V. Discussion.

To briefly recap, the results presented above suggest that participation in a sociological inquiry-based learning project helped STEM students to (a) improve their knowledge of scientific and statistical principles and processes, (b) hone their scientific research skills (if we assume that their self-assessment of these skills correlates with an objective assessment of these skills), and (c) gain respect for sociology specifically and social science more generally. More studies of this nature should be conducted to determine the robustness of these findings. The results of such studies may find that participation in such a sociological inquiry-based learning project may help students improve their scientific knowledge and quantitative skills and shift their attitudes to (a) get more out of interdisciplinary undergraduate majors and minors; (b) secure a position in an interdisciplinary internship or research team; or (c) get into an interdisciplinary graduate or professional program.

Several limitations of this current study demand attention, suggest that these results should be treated with caution, and provide avenues for future research. First, this study employed a quasi-experimental design largely because of the structural impediments to randomly sorting students into different courses and the logistical difficulties of gathering data from students in other instructors' courses at multiple points in a semester. Yet, significant planning in certain curricular settings may allow for a truly randomized experimental design to more effectively determine the influence of participation in such an inquiry-based learning project. Second, this study utilized imperfect measures of scientific process skills and quantitative literacy. Better measures do exist in the literature and, with revision and shortening, may be employed in future studies of this type. Third, this study only examined the immediate impacts of participating in an inquiry-based learning project. While significant short-term impacts on scientific and quantitative literacy were found, such impacts may not endure over a longer time frame. Thus, future research should aim to follow up with participants in the weeks and months after the semester.

Despite these limitations, this study does suggest that such an inquiry-based learning project may be fruitful for cultivating the scientific and quantitative literacies and social science skills necessary for STEM students to effectively deal with complex environmental problems and solutions. Thus, other educators may utilize projects similar to one described here to help their students gain hands-on experience doing research on a wide range of environmental topics. Further, this study provides an effective model for integrating sociology (and social science more generally) into a STEM curriculum. This is important for two reasons: the increasing workforce demands for university graduates with STEM degrees and the shifting priorities within the country's federal scientific funding bodies. This paper concludes by elaborating on these two reasons.

University officials, prospective employers, and public policy-makers expect that university graduates with STEM degrees have a wide range of personal and social skills and interdisciplinary and international experiences upon graduation to compete effectively in the increasingly global 21st century workforce (U.S. National Research Council, 2009). Interdisciplinary undergraduate programs are emerging across the nation, but far too many undergraduates seeking interdisciplinary experiences still struggle to find them. Many students seek out such experiences on an individualized, ad hoc basis—through a specific internship or a specific research opportunity, such as an NSF-funded Research Experience for Undergraduates. Systematizing or formalizing our interdisciplinary research opportunities—with one example being the inquiry-based learning project described here—may allow groups of students to share common experiences and thus further benefit through collaborative learning.

Within the U.S. scientific community, there is an increasing premium on interdisciplinary scholarship in recent years. For instance, the NSF includes several interdisciplinary programs (e.g., Cyber-enabled Discovery and Innovation; Dynamics of Coupled Natural Human Systems), competitions for centers that formalize interdisciplinary research teams (e.g., Science of Learning Centers; Science and Technology Centers), and programs that support students (e.g., Integrative Graduate Education and Research Traineeship Program). Many of these initiatives

and opportunities at the NSF and beyond (e.g., National Institutes of Health; National Oceanic and Atmospheric Administration; National Aeronautics and Space Administration) not only support scholarship that spans multiple disciplines in the natural sciences but also scholarship that spans the much wider cultural divide between the natural sciences and the social sciences. Indeed, compared to decades past, the future of scientific opportunities for interdisciplinary scholarship looks quite bright, with increased funding opportunities, publication outlets, conferences, and graduate programs. Thus, we also should be improving the quality and quantity of our interdisciplinary educational opportunities for undergraduate STEM and social science majors. The type of inquiry-based learning project described in this study may be one fruitful avenue in this regard.

## Acknowledgements

Special thanks to Deb DeZure, Assistant Provost for Faculty and Organizational Development and Director of the Lilly Teaching Fellows Program at Michigan State University. Thanks also to Robert Shelton, Jan Bokemeier, Tom Dietz, Chris Ganchoff, Mark Waddell, Catherine Westfall, and Cindi Young. I extend my gratitude to my students for their enthusiastic participation in my research project. Support for this project was made available through the Lilly Teaching Fellows Program at Michigan State University.

## References

Albers, C. (2008). Improving pedagogy through action learning and scholarship of teaching and learning. *Teaching Sociology*, *36*, 79-86.

Alper, J. (1993). The pipeline is leaking women all the way along. *Science*, *260*, 409-411.

American Association of University Women. (1992). *How Schools Shortchange Girls*. Washington, DC: AAUW.

Arum, R., & Roksa, J. (2011). *Academically Adrift*. Chicago: University of Chicago Press.

Atkinson, M.P., & Hunt, A.N. (2008). Inquiry-guided learning in sociology. *Teaching Sociology*, *36,* 1-7.

Brickman, P., Gormally, C., Armstrong, N., & Hallar, B. (2011). Effects of inquiry-based learning on students' science literacy skills and confidence. *International Journal for the Scholarship of Teaching and Learning*, *3*(2), 1-22.

Burdette, A.M., & McLoughlin, K. (2010). Using census data in the classroom to increase quantitative literacy and promote critical sociological thinking. *Teaching Sociology*, *38,* 247-257.

Chandramohan, B., & Fallows, S.J. (Eds.). (2009). *Interdisciplinary Learning and Teaching in Higher Education*. London: Routledge.

Committee on Science, Engineering, and Public Policy. (2007). *Beyond Bias and Barriers*. Washington, DC: National Academy Press.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation*. Chicago: Rand McNally.

Derting, T.L., & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology. *CBE Life Science Education*, *9,* 462-472.

Etkowitz, H., Kemelgor, C., & Uzzi, B. (2000). *Athena Unbound*. New York: Cambridge University Press.

Garkovich, L. (1982). A proposal for building interdisciplinary bridges. *Teaching Sociology*, 9, 151-168.

Grawe, N.D. (2011). The potential for teaching quantitative reasoning across the curriculum. *International Journal for the Scholarship of Teaching and Learning*, *5*(1), 1-12.

Handelsman, J., Ebert-May, D., Beichner, R., Burns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S.M., & Wood, W.B. (2004). Scientific teaching. *Science*, *304*, 521-522.

Himes, C.L., & Caffrey, C. (2003). Linking social gerontology with quantitative skills. *Teaching Sociology*, *31*, 85-94.

Johansen, D., Scaff, C., & Hargis, J. (2009). Interdisciplinary project-based model for enhanced instruction of courses. *International Journal for the Scholarship of Teaching and Learning*, *3*(1), 1-14.

Klein, J.T. (2010). *Creating Interdisciplinary Campus Cultures: A Model for Strength and Sustainability*. San Francisco: Jossey-Bass.

Levine-Rasky, C. (2009). The public poster session. *Teaching Sociology*, *37*, 309-316.

Lindner, A.M. (2012). Teaching quantitative literacy through a regression analysis of exam performance. *Teaching Sociology*, *40,* 50-59.

Longmore, M.A., Dunn, D., & Jarboe, G.R. (1996). Learning by doing. *Teaching Sociology*, *24,* 84-91.

McKinney, K. (2008). Correlates of success in the sociology major. *International Journal for the Scholarship of Teaching and Learning*, *2*(1), 1-14.

McKinney, K., & Busher, M. (2011). The sociology research experience capstone course at three institutions. *Teaching Sociology*, *39*, 290-302.

Mohamed, A.R. (2008). Effects of active learning variants on student performance and learning perceptions. *International Journal for the Scholarship of Teaching and Learning*, *2*(2), 1-15.

Nguyen, T., & Trimarchi, A. (2010). Active learning in introductory economics: do MyEconLab and Aplia make any difference? *International Journal for the Scholarship of Teaching and Learning*, 4.1, 1-18.

Parrott, H.M., & Cherry, E. (2011). Using structured reading groups to facilitate deep learning. *Teaching Sociology*, *39*, 354-370.

Pedersen, D.E. (2010). Active and collaborative learning in an undergraduate sociological theory course. *Teaching Sociology*, *38*, 197-206.

Raddon, M., Nault, C., & Scott, A. (2008). Integrating the complete research project into a large qualitative methods course. *Teaching Sociology*, *36*, 141-149.

Ross, S.M., & Hurlbert, J.M. (2004). Problem-based learning. *Teaching Sociology*, *32*, 79-93.

Rusche, S.N., & Jason, K. (2011). You have to absorb yourself in it. *Teaching Sociology*, *39,* 338-353.

Steck, L.W., Engler, J.N., Ligon, M., Druen, P.B., & Cosgrove, E. (2011). Doing poverty. *Teaching Sociology*, *39,* 259-273.

Svinicki, M.D. (1999). New directions in learning and motivation. *New Directions for Teaching and Learning*, *80*, 5-27.

Teixeira-Poit, S.M., Cameron, A.E., & Schulman, M.D. (2011). Experiential learning and research ethics. *Teaching Sociology*, *39,* 244-258.

U.S. Department of Education. (2006). *A test of leadership: A report of the Commission appointed by Secretary of Education Margaret Spellings*. Washington, DC: U.S. Department of Education.

U.S. National Research Council. (2003). *BIO2010*. Washington, DC: National Academies Press.

U.S. National Research Council. (2007). *Rising above the gathering storm*. Washington, DC: National Academies Press.

U.S. National Research Council. (2009). *Transforming agricultural education for a changing world*. Committee on a Leadership Summit to Effect Change in Teaching and Learning; Board on Agriculture and Natural Resources. Washington, DC: National Academies Press.

Uriarte, M., Ewing, H.A., Eviner, V.T., & Weathers, K.C. (2007). Constructing a broader and more inclusive value system in science. *BioScience*, *57,* 71-78.

Vajoczki, S., Watt, S., Vine, M.M., & Liao, X. (2011). Inquiry learning: Level, discipline, class size, what matters? *International Journal for the Scholarship of Teaching and Learning*, *5*(1), 1-11.

Whitman, P.D., Richlin, L., Arboleda, A., Garrett, N.D., Robinson, E., Viale, D., Seward, E., Biseda, M., Alexander, J., Andrade-Lopez, E., Brooks, A., Carter, L., Ha, A.O., Owens, R., Thompson-Eggersten, K.A., & Vercoe, M.J.H. (2007). The status of the scholarship of teaching and learning in the disciplines. *International Journal for the Scholarship of Teaching and Learning*, 1.1, 1-17.

Wilder, E.I. (2009). Responding to the quantitative literacy gap among students in sociology courses. *Teaching Sociology*, *37,* 151-170.

Yamane, D. (2006). Course preparation assignments. *Teaching Sociology*, *34,* 236-248.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 102 – 123.

# Science, technology, engineering, and mathematics graduate teaching assistants teaching self-efficacy

**Sue Ellen DeChenne[1], Larry G. Enochs[2], and Mark Needham[3]**

*Abstract: The graduate experience is a critical time for development of academic faculty, but often there is little preparation for teaching during the graduate career. Teaching self-efficacy, an instructor's belief in his or her ability to teach students in a specific context, can help to predict teaching behavior and student achievement, and can be used as a measure of graduate students' development as instructors. An instrument measuring teaching self-efficacy of science, technology, engineering, and mathematics (STEM) graduate teaching assistants (GTAs) was developed from a general university faculty teaching instrument to the specific teaching context of STEM GTAs. Construct and face validity, measurement reliability, and factor structure of the instrument were determined from survey data of 253 STEM GTAs at six universities. STEM GTA teaching self-efficacy correlated to various measures of GTA professional development and teaching experience. Implications and applications for faculty involved in GTA professional development, supervision, and research are discussed.*

*Keywords: Teaching self-efficacy; STEM GTA Professional Development; Faculty development; Teaching experience*

Graduate teaching assistants (GTAs) in science, technology, engineering, and mathematics (STEM) disciplines have and are going to continue to have a large influence on the teaching of undergraduate students. Many of the first instructional experiences that undergraduates have in college are closely associated with their GTAs. STEM GTAs teach both major and non-major students, potentially impacting the scientific literacy of the college educated population and the knowledge and retention of STEM majors (Fencle & Scheel, 2005; Miller, Pfund, Pribbenow, & Handelsman, 2008; O'Neal, Wright, Cook, Perorazio, & Purkiss, 2007). STEM GTAs often have more contact hours with students than the professors do, especially in large introductory undergraduate courses where GTAs are usually responsible for teaching laboratory or recitation sections (Fagen & Wells, 2004; Golde & Dore, 2001). For example, GTAs provide 91% of biology and 88% of chemistry laboratory instruction at research universities (Abraham et al. 1997; Sundberg, Armstrong, & Wischusen, 2005). In addition, many STEM doctoral students are interested in an academic career where they will be the future professors teaching the next generation of undergraduate students (Golde & Dore, 2001).

Although the graduate experience is a critical time for development of future academic professionals (Austin, 2002), many receive no formal professional development in teaching (Abraham et al., 1997; DeChenne et al., 2009; Golde & Dore, 2001; Meyers, Lansu, Hundal, Lekkos, & Prieto, 2007; Piccinin & Fairweather, 1996-97; Prieto & Scheel, 2008; Rushin et al., 1997). While there are individual programs showing promising results (e.g., Burton, Bamberry, & Harris-Boundy, 2005; Davis & Kring, 2001; Webber, Gabbert, Kropp, & Pynes, 2007), many

[1] Postdoctoral Fellow, University of Nebraska – Lincoln, National Center for Research on Rural Education, sdechenne2@unl.edu
[2] Oregon State University, enochsl@onid.orst.edu
[3] Oregon State University, mark.needham@oregonstate.edu

studies indicate that GTA professional development may be ineffectual (Commander, Hart, & Singer, 2000; Fagen & Wells, 2004; Jones, 1993; Luft, Kurdziel, Roehrig, & Turner, 2004; Prieto & Scheel, 2008; Shannon, Twale, & Moore, 1998). Most of the research about GTA professional development, however, is descriptive with few measures of the effectiveness of professional development (DeChenne, 2012). Reliable and valid measures of constructs related to improvement in teaching are critical for advancing the scientific knowledge base and teaching practice of STEM GTAs.

Teaching self-efficacy is a construct which is related to teaching effectiveness (Bandura, 1997) and is a domain specific construct focusing on teacher perceptions of their own ability to "organize and execute courses of action required to successfully accomplish a specific teaching task in a particular context" (Tschannen-Moran, Hoy, & Hoy, 1998, p. 233). Teaching self-efficacy has been shown to be a valuable predictor for student achievement, teacher retention, and persistence in the face of teaching difficulties (for a review, see Tschannen-Moran, et al., 1998). Given the empirically established relationship to student and teaching outcomes, teaching self-efficacy can contribute to our understanding of STEM GTA teaching. Before this can be accomplished however, a valid and reliable measure of STEM GTA teaching self-efficacy is needed.

## I. Literature Review.

### *A. Social Cognitive Theory.*

Self-efficacy beliefs are a central component of skill development in social cognitive theory (Bandura, 1986, 1997), which provides a mechanism for development of skills through interactions of behavior, personal attributes, and environmental circumstances. According to Bandura (1997):

> Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments…[self-efficacy] beliefs influence the courses of action people choose to pursue, how much effort they put forth in given endeavors, how long they will persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-hindering or self-aiding, how much stress and depression they experience in coping with taxing environmental demands, and the level of accomplishments they realize (p. 3).

Research has demonstrated that with professional development for a specific skill, self-efficacy is positively correlated with performance (Bandura, 1997; Gist, Schwoerer, & Rosen, 1989; Pajares, 1996a). Self-efficacy beliefs can predict performance and have been used in the literature as a measure of such performance, especially when the performance is difficult to measure quantitatively, such as in teaching (e.g., Burton et al., 2005; Prieto & Altmaier, 1994; Young & Bippus, 2008).

According to social cognitive theory, there are four sources of self-efficacy: mastery experiences, vicarious experiences, verbal persuasions, and psychological states (Bandura, 1997). Mastery experiences are related to evaluation of performance in the task; successful performance increases self-efficacy, whereas failure decreases self-efficacy. The best mastery experiences are those that challenge the individual, but in which they are ultimately successful. Vicarious experience is provided through modeling. Being able to observe someone perform the

skill successfully improves the observer's self-efficacy, especially when the model is someone similar to the observer. Verbal persuasions include feedback on performance; receiving timely feedback about how well a person performed a skill can influence self-efficacy. Psychological states such as how someone feels when performing the skill (e.g., anxious, calm, excited) can influence their perceived self-efficacy about that skill.

*B. Teaching Self-Efficacy.*

Research has demonstrated that teaching self-efficacy impacts many student outcomes and teacher behaviors. A teacher's self-efficacy beliefs positively impact student learning and the actual success or failure of a teacher's behavior (Henson, 2002). These beliefs are also related to teacher instructional practices (Borko & Putnam, 1996; Haney, Czerniak, & Lumpe, 1996) and to student achievement and psychological wellbeing; teachers with high teaching self-efficacy tend to perform better and their students benefit (Ashton & Webb, 1986; Tschannen-Moran et al., 1998).

Teaching self-efficacy usually develops early in a teacher's career and becomes relatively stable over time (Morris & Usher, 2011; Tschannen-Moran et al., 1998), which makes the graduate experience especially critical for developing teaching self-efficacy in future professors (Hoy, 2003-2004). Being a GTA is usually the first instructional experience for most university faculty, and GTA professional development is often the only education in instruction that faculty receive (Tanner & Allen, 2006). Morris and Usher (2011) found that early successful instructional experiences, which are a combination of mastery experiences and verbal persuasions, are important for developing high teaching self-efficacy of teaching award winning professors, and that their teaching self-efficacy solidified within the first few years as a faculty member.

*C. Teaching Self-Efficacy Instrument Design.*

Bandura (1997) proposed that because self-efficacy beliefs are explicitly self-referent in nature and directed toward perceived abilities about given specific tasks, they are powerful predictors of behavior. Teaching self-efficacy refers to organizing and executing courses of action required to successfully accomplishing *a specific teaching task in a particular context*. Many measures of self-efficacy address specific tasks, yet fall short of providing the particular context (Dellinger, Bobbett, Olivier, & Ellett, 2008; Henson, 2002; Pajares, 1996a).

In development of teaching self-efficacy instruments, it is important to address the situational specificity of different teaching contexts and tasks and to balance the general and specific tasks (Bandura, 1997; Henson, 2002; Pajares, 1996a). If too specific, the instrument is not likely to measure overall teaching self-efficacy, just a specific teaching task in a specific context. For example, a measure of a genetics GTA's belief in their ability to teach loading DNA samples into an agarose gel for electophoresis to freshman level biology students is too specific. If the context is too broad, however, the instrument may simply be measuring general personality traits instead of self-efficacy specific to the task (Pajares, 1996b). Given that self-efficacy refers to an individual's beliefs about his or her abilities to accomplish a behavior or task in a *specific context*, it is important to design the instrument to the context in which the person is performing the task (Bandura, 1997).

Recently, instruments have been designed with subscales for teaching self-efficacy (Chang, Lin, & Song, 2011; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Hoy, 2001). These instruments measure teaching self-efficacy and specific types of tasks within teaching, providing subscales of efficacy and giving greater range in measuring teaching self-efficacy. These three instruments have between three and six correlated subscales including: student engagement, instruction, classroom management, motivation of students, course design, technology usage, learning assessment, etc. These instruments can be used as an overall measure of teaching self-efficacy or as a more specific subscale score.

*D. GTA Teaching Self-Efficacy.*

Teaching self-efficacy has been studied in GTAs both in individual programs and across multiple programs. Studies indicate that teaching experience seems to increase teaching self-efficacy, although it depends on the level of teaching responsibility. In an early study of teaching self-efficacy with new counseling psychology professors, Tollerud (1990) demonstrated that GTA teaching experience positively impacted teaching self-efficacy. Using a wider population of GTAs, teaching experience was correlated to teaching self-efficacy in several studies (Liaw, 2004; Prieto & Altmaier, 1994; Prieto, Yamokoski, & Meyers, 2007). However, in a study of business GTAs, Burton et al. (2005) found that teaching experience was not correlated to teaching self-efficacy. In a study of International Teaching Assistants (ITAs), Kim (2009) found that teaching experience was correlated with teaching self-efficacy related to instructional strategies and classroom management, but not those related to student engagement. Theoretically, experience should provide ample mastery experiences and verbal persuasions to impact teaching self-efficacy.

Level of responsibility for teaching can also have an impact on GTA teaching self-efficacy. In two studies, GTAs in non-instructive roles (e.g., graders) had significantly lower teaching self-efficacy scores than GTAs who had teaching roles (e.g., assistant, primary instructors; Prieto & Meyers, 1999) and there were significant differences between GTAs who were graders or assistants and those who had primary responsibility for the classroom (Prieto et al., 2007). Prieto and Altmaier (1994), however, found no significant differences in teaching self-efficacy based on amount of teaching responsibility (e.g., primary instructor, assistant instructor, general instructor). According to social cognitive theory (Bandura, 1997), GTAs in courses with more responsibility should have far more experiences that impact their teaching self-efficacy.

The impact of GTA professional development on teaching self-efficacy is not as well established. In four studies of GTAs from multiple programs, two studies found no impact of GTA professional development on teaching self-efficacy (Liaw, 2004; Tollerud, 1990). Two other studies, however, did find that GTA professional development increased teaching self-efficacy (Prieto & Altmaier, 1994; Prieto & Meyers, 1999). In a qualitative study of foreign language GTAs, Mills and Allen (2007) found that GTA professional development was highly influential. Studies of individual GTA professional development programs indicated that teaching self-efficacy increased after the GTA professional development program was completed (Burton et al., 2005; Hadre, 2003; Komarraju, 2008; Meyers et al., 2007; Sargent, Allen, Frahm, & Morris, 2009; Young & Bippus, 2008). These mixed results are likely due to the variable quality of GTA professional development available to study participants in the quantitative

studies of multiple programs. If the quality of the GTA professional development was poor, then there should be little impact on GTA teaching self-efficacy.

Teaching self-efficacy of GTAs has been measured with an instrument originally designed for psychology students (Prieto & Altmaier, 1994) or one taken from the K-12 teaching context (often the Teacher Sense of Efficacy Scale; Tschannen-Moran & Hoy, 2001), but it has been recognized that teaching in STEM is fundamentally different from other disciplines and this difference should be recognized in roles of GTAs (Golde & Dore, 2004; Lindblom-Ylanne, Trigwell, Nevgi, & Ashwin, 2006; Torvi, 1994; Verleger & Velasquez, 2007). STEM GTAs are rarely responsible for a course (Abraham et al., 1997; DeChenne et al., 2009; Sundberg et al., 2005; Torvi, 1994), but instead teach laboratory and recitation sections, so usually act as a conduit between the students and course professor. STEM GTAs need to understand complex grading rubrics and have skills allowing them to facilitate questions without giving students answers. STEM students often work independently or in small groups on complex projects that can span a term or more of coursework (Moore & Diefes-Dux, 2004; Pomalaza-Raez & Groff, 2003; Taylor, Heer, & Fiez, 2003). GTAs must understand these long-term projects, how to facilitate learning, and help students at different points of scholarship and with often frustrating problems. All of these activities require STEM GTAs to have excellent interpersonal skills.

Given that the STEM GTA context is quite different from many other teaching contexts, measuring teaching self-efficacy in STEM GTAs arguably requires a context specific instrument. Published studies using teaching self-efficacy as a measure, however, are limited to inclusion of STEM GTAs within a study of multiple programs usually using a generic instrument originally designed for the psychology context (Prieto & Altmaier, 1994). A teaching self-efficacy instrument is needed for STEM GTA teaching context. To validly measure teaching self-efficacy, this instrument should correlate to measures of teaching experience and measures of GTA professional development. Like other newer teaching self-efficacy instruments, an instrument measuring subscales within STEM GTA teaching self-efficacy would provide the possibility of a more sophisticated understanding of STEM GTAs teaching self-efficacy. Such an instrument would be useful in further research on STEM GTA teaching self-efficacy and for faculty responsible for GTA supervision and professional development. The purpose of this study, therefore, is to develop and validate an instrument measuring STEM GTA self-efficacy in teaching, and explore relationships between STEM GTA teaching self-efficacy, GTA professional development, and teaching experience.

## II. Methods.

### *A. Participants.*

Data were collected from GTAs in various STEM departments at six USA universities; three in the Pacific Northwest, two in the Southwest, and one in the Midwest. Five universities (including the originating university) had a Carnegie basic classification of RU/VH (Research Universities with Very High research activity) and one was a DRU (Doctoral/Research University). Engineering and technology GTAs taught across various engineering disciplines (e.g., aerospace, biological, biomedical, chemical, civil, computer, construction, electrical, environmental, industrial, manufacturing, mechanical, and petroleum). Science GTAs taught biochemistry, biology, chemistry, geosciences, microbiology, molecular biology, and physics. Also included in the sample were GTAs who taught mathematics.

*B. Instrument Development and Modification.*

The instrument used in many GTA teaching self-efficacy studies was the Self-Efficacy Toward Teaching Inventory – Adapted (SETI-A) (Prieto & Altmaier, 1994), which had been adapted for general GTA use from a teaching self-efficacy instrument that was specific for counseling psychology educators (Tollerud, 1990). Another post-secondary level instrument, the College Teaching Self-Efficacy Scale (CTSES), had been recently developed (Prieto Navarro, 2005). A team, including two science educators and two engineering education faculty, discussed what types of items should be included in a STEM GTA teaching self-efficacy scale and reviewed the items on the CTSES and SETI-A. The CTSES was chosen since it required less extensive modification and the team collaborated to adapt the CTSES to the STEM GTA context.

As part of a larger study of STEM GTA teaching self-efficacy, the CTSES needed to be streamlined; items specific to STEM GTA teaching were added or modified from the general college instructor context and items not usually part of a STEM GTA duties were removed. The CTSES was long (44 items) and contained two six-point scales, one for self-efficacy and one measuring actual instructor action for each item. Only the self-efficacy scale was retained, but changed to a five-point scale because of limitations of data collection, as the instrument was distributed in print and the data were collected on a scantron bubble form with five response options per question.

Seven items related to overall course design and planning were removed because STEM GTAs were rarely involved in course design or were the primary instructors responsible for a course. The CTSES also contained five items on reflective practice, many of which required teaching the same course repeatedly. Many GTAs, especially in engineering, did not teach in the same course repeatedly, so these items were removed. Three items that were unclear to the researchers or included technical pedagogical language were removed. There were also two pairs of redundant items, so one item from each pair was removed. Four items were rewritten to be more specific to the STEM GTA context; and given the large amount of group work in STEM laboratory classes; one item related to student interaction was added. Face validity of the items was reviewed by two additional social science faculty members with knowledge of both social cognitive theory and instrument design; they were asked to evaluate whether each item represented an aspect of GTA teaching self-efficacy, comment on clarity, and suggest revisions or additions. The final STEM GTA-Teaching Self-Efficacy Scale (STEM GTA-TSES) as administered contained 28 items measured on a five point scale anchored with A (no confidence) and E (complete confidence).

At the five institutions outside the originating university, three extra questions were asked in addition to the items in the STEM GTA-TSES – two demographic questions (university and department affiliations) and a question indicating the GTA's primary teaching role (laboratory, recitation, lecturer, course instructor or grader). At the originating university, the instrument contained measures of STEM GTA-TSES, GTA professional development, GTA teaching experience, and additional demographic questions such as gender, nationality, department, and career interest. GTA professional development was measured at the originating university: (a) as the number of total hours spent in teaching professional development at the university, department, and through university coursework in teaching; and (b) through an instrument measuring the GTA's perception of their GTA professional development (DeChenne et al., 2012). Teaching experience was measured by totaling the number of quarters taught (semesters converted to quarters) and by two items asking GTAs to rate their own experience. One item

asked GTAs to compare themselves to other GTAs in their department (less experience to more experience) and the other asked them to rate their own experience (beginner to expert).

*C. Administration.*

GTAs were administered the STEM GTA-TSES once near the end of the semester or quarter. Data were collected from the various sites from fall 2008 through fall 2010, with one of two administration techniques used depending on location. Questionnaires were distributed to the GTAs through the department mail system, collected in a sealed container in the departmental office, and returned to the researchers through the mail (or collected directly by a researcher). Alternatively, questionnaires were administered during a GTA professional development class, collected by a faculty or staff member not involved in the GTA professional development and returned through the mail (or collected by one of the researchers at that time).

*D. Analysis.*

STEM GTAs from all universities were used to run the reliability and factor analysis while a subset of the STEM GTAs from the originating university was used for correlational analysis. The 28 STEM GTA-TSES items were analyzed with all data using principle axis exploratory factor analysis (EFA) with Varimax rotation, Kaiser Criterion, and Scree test. Confirmatory factor analysis (CFA) was then used with all data to examine if a second-order factor structure provided good fit and demonstrated construct validity. A CFA of the GTA professional development items was also used to examine whether the variables measuring this latent factor provided good fit and demonstrated construct validity. EQS 6.1 software and Satorra-Bentler Robust estimation to correct for multivariate non-normality were used for the CFA analysis (Byrne, 1994). Robust corrected comparative fit index (CFI), non-normed fit index (NNFI), and root mean square error of approximation (RMSEA) were used to assess model fit. CFI and NNFI values $\geq 0.90$ and RMSEA values $\leq 0.08$ imply acceptable fit (Browne & Cudeck, 1993).

Internal consistency of multiple-item indices measuring these concepts was examined with Cronbach's alpha reliability coefficients. An alpha coefficient of approximately $\geq 0.65$ indicated that items measure the same concept and justified combining items into a single index (Cortina, 1993).

Using the GTAs from the originating university, Pearson correlations ($r$) between the STEM GTA-TSES and GTA professional development and teaching experience measures were determined. According to Cohen (1988), correlations less than .10 are considered small or weak, those around .30 are moderate or medium and those greater than .50 are large or strong. Using GTAs from the originating university, differences in teaching self-efficacy by gender, career goals, and nationality were determined with $t$-tests and effect sizes were examined using point-biserial correlations ($r_{pb}$). Instructional role and college of instruction were similarly compared using data from all GTAs across institutions.

## III. Results.

In total, there were 253 participants: 177 from the originating university and 76 across the other five universities. Engineering GTAs comprised 68% of participants with 32% in science or mathematics. Twenty-five percent of GTAs described their primary role as grading and the

remaining 75% indicated classroom instruction as their main role, with laboratory instructor (42%) being the most common and course instructor the least common (5%). Twenty-seven percent of GTAs were female, 47% were ITAs, and 64% were interested in college or university teaching as a career. Seventeen percent of the sample had no GTA professional development of any kind. Sixty-seven percent had less than two years teaching experience and the GTAs had taught an average of 3.2 different courses.

*A. GTA-Teaching Self-Efficacy Instrument.*

The EFA of the 28 teaching self-efficacy items revealed four factors explaining 51% of the variance. Two of the factors had a Kaiser criterion (Guttman, 1954) greater than one and a Scree test (Cattell, 1966) also suggested that two factors could be found in the data. Since both of these indicated that there were two factors, the exploratory factor analysis was rerun forcing two factors. To strengthen the factors, all of the items that cross-loaded between the factors were removed, leaving 18 items (Costello & Osborne, 2005). An EFA with those 18 items revealed two clean factors explaining 46% of the variance (Table 1). The factors were labeled self-efficacy for learning environment (*learning* = 11 items) and instructional strategies (*instructional* = 7 items). All factor loadings were between .49 to .77 for learning environment and .51 to .71 for instructional strategies. Both factors were also highly reliable (*learning* $\alpha$ = .90, *instructional* $\alpha$ = .85). All variables met the criterion of item total item correlations being greater than .40, and deletion of any item did not improve reliability. Means for each factor were high (*learning* = 4.07, *instructional* = 4.20), indicating that for each factor the GTAs were confident in their ability to carry out these teaching duties and responsibilities.

Given the strong correlation between learning environment and instructional strategies with all GTAs ($r$ = .66), a higher order structure was possible, which was not uncommon in teaching self-efficacy scales (Chang, Lin, & Song, 2011; Skaalvik & Skaalvik, 2007; Tschannen-Moran & Hoy, 2001) and was advocated in their development (Dellinger, et al., 2008). A second-order CFA was performed on the items and there was a good fit in the second order structure (NNFI = .92, CFI = .93, RMSEA = .04; Figure 1). All variables loaded between .62 to .74 for learning environment and .58 to .70 for instructional strategies, and each factor loaded highly on the second-order GTA teaching self-efficacy construct (*learning* = .87, *instructional* = .85). All factor loadings were significant at $p < .05$. Reliability of the single factor structure was .92 with a mean of 4.10, all variables met the criterion of item total item correlations being greater than .40, and deletion of any item did not improve reliability. These results indicated that this instrument could be used to measure the underlying concept of STEM GTA teaching self-efficacy; measuring total teaching self-efficacy as well as the *learning* and *instructional* subscale self-efficacies.

*B. Correlational and Comparative Analysis.*

The overall STEM GTA-TSES and the learning environment and instructional strategies subscales showed significant positive correlations with several measures of teaching professional development and teaching experience (Table 2). These measures (originating university GTA sample) indicated significant moderate positive correlations of the STEM GTA-TSES and both

**Table 1.** Factor Loadings for Exploratory Factor Analysis of STEM GTA-TSES Subscales with All GTAs.

| How confident am I in my ability too[2]… | V# | Factor Loadings[1] | |
|---|---|---|---|
| | | Learning[3] | Instructional[3] |
| Promote student participation in my classes? | 1 | **.77** | .13 |
| Make students aware that I have a personal investment in them and in their learning? | 2 | **.70** | .22 |
| Create a positive classroom climate for learning? | 3 | **.68** | .22 |
| Think of my students as active learners, which is to say knowledge builders rather than information receivers? | 4 | **.65** | .22 |
| Encourage my students to ask questions during class? | 5 | **.62** | .32 |
| Actively engage my students in the learning activities that are included the teaching plan/syllabus? | 6 | **.61** | .31 |
| Promote a positive attitude towards learning in my students? | 7 | **.59** | .33 |
| Provide support/encouragement to students who are having difficulty learning? | 8 | **.57** | .34 |
| Encourage the students to interact with each other? | 9 | **.54** | .34 |
| Show my students respect through my actions? | 10 | **.50** | .34 |
| Let students take initiative for their own learning? | 11 | **.49** | .35 |
| Appropriately grade my students' exams/assignments? | 12 | .11 | **.71** |
| Evaluate accurately my students' academic capabilities? | 13 | .23 | **.69** |
| Prepare the teaching materials I will use? | 14 | .30 | **.62** |
| Spend the time necessary to plan my classes? | 15 | .25 | **.60** |
| Clearly identify the course objectives? | 16 | .32 | **.58** |
| Provide my students with detailed feedback about their academic progress? | 17 | .32 | **.55** |
| Stay current in my knowledge of the subject I am teaching? | 18 | .34 | **.51** |
| Eigenvalue | | 7.67 | 1.63 |
| Percent (%) of total variance explained | | 39.59 | 6.27 |
| Cumulative percent (%) of variance | | 39.59 | 45.86 |
| Factor Mean[2] | | 4.07 | 4.20 |
| Cronbach $\alpha$ | | .90 | .85 |

[1] Principal axis factor analysis with Varimax rotation.
[2]Items coded on a 5 point scale of 1 = not at all confident to 5 = very confident.
[3]Factor loadings > .40 are in boldface.

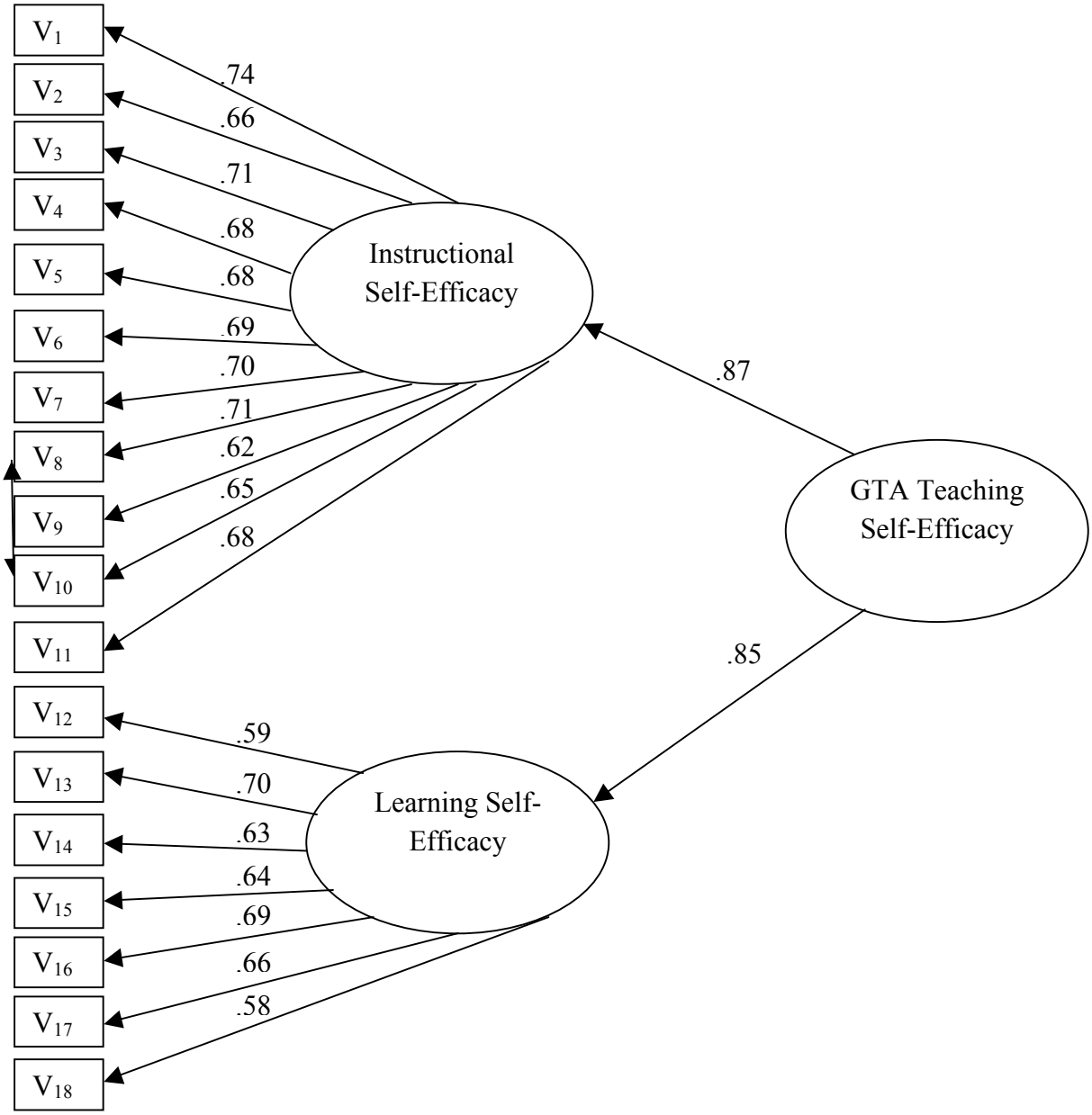


**Figure 1**. Second Order Confirmatory Factor Analysis of STEM GTA Teaching Self-Efficacy with All GTAs.
See Table 1 for variables corresponding to codes (e.g., $V_1$). All factor loadings indicated are significant at $p < .05$. Model fit indices are NNFI = .924, CFI = .934, & RMSEA = .043. To achieve these fit indices the errors between $V_8$ and $V_{10}$ were allowed to co-vary**.**

subscales with GTAs perception of professional development (GTA-TSES $r = .30$). There was also a small significant correlation of the number of hours reported in GTA professional development with teaching self-efficacy and the two subscale measures (GTA-TSES $r = .18$). There were significant moderate correlations of the STEM GTA-TSES and both subscales with

measures of teaching experience asking GTAs to rate themselves compared to their peers (GTA-TSES $r$ = .34) and on a scale from beginner to expert (GTA-TSES $r$ = .45). There was a small significant correlation with the instructional strategies subscale and the number of quarters that GTAs taught (*instructional* $r$ = .16).

**Table 2.** Correlation Analysis with Originating University GTAs.

| | Measures[1] | Mean | GTA[2] TSES | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|
| GTA-TSES[2] | | 4.1 | | | | | | | |
| A | Learning Instructiona | 4.1 | .95** | | | | | | |
| B | l | 4.2 | .86 ** | .66** | | | | | |
| Professional Development | | | | | | | | | |
| C | Perception[3] | 3.2 | .30** | .27** | .29** | | | | |
| D | Hours[4] | 20.8 | .18* | .16* | .17* | .22** | | | |
| Teaching Experience | | | | | | | | | |
| E | Compare[5] | 3.3 | .34** | .27** | .38** | .04 | .12 | | |
| F | Rating[6] | 2.9 | .45** | .38** | .44** | .22** | .22** | .67** | |
| G | Quarters[7] | 5.2 | .11 | .07 | .16* | <.01 | .10 | .46** | .46** |

*$p$ < .05 (2 tailed), **$p$ < .01 (2-tailed)
[1]All scales were rated on a scale of 1 to 5, with 5 being the best in each scale.
[2]GTA-Teaching Self-Efficacy Scale
[3]GTA ratings of how well they learned teaching skills
[4]Total hours of professional development including university, department, and credit coursework
[5]Item asking: Compared to other GTAs how much teaching experience do you have?
[6]Item asking: Rate your own teaching experience?
[7]Numbers of quarters as a GTA

Data collected with GTAs at the originating university indicated only one group difference in mean scores on the STEM GTA-TSES or the *learning* and *instructional* subscales. Consistent with Prieto and Altmaier (1994), teaching self-efficacy in STEM GTAs does not vary in this study by gender (STEM GTA-TSES males = 4.13, females = 4.09; $t$ = .46 $p$ = .646, $r_{pb}$ = .04), career plans (STEM GTA-TSES academic = 4.18, other = 4.02; $t$ = 1.65, $p$ = .101, $r_{pb}$ = .13), college of instruction (STEM GTA-TSES science/math = 4.06, engineering = 4.11; $t$ = .711, $p$ = .478, $r_{pb}$ = .05), or instructional role (STEM GTA-TSES graders = 4.15, classroom instruction = 4.10; $t$ = .56, $p$ = .578, $r_{pb}$ = .04). In two other populations, however, the teaching self-efficacy of GTAs varied by instructional role (Prieto & Meyers, 1999; Prieto et al., 2007). Unlike these studies, STEM GTAs were rarely the course instructor (<5% with this population) which was the instructional role reported to have different teaching self-efficacy (Prieto et al., 2007) and makes up almost half of the combined assistant/full responsibility group in Prieto and Meyers (1999).

The one group difference occurred with ITAs, who had significant higher instructional strategies self-efficacy than United States GTAs (USTAs) (*instructional* ITA = 4.34, USTA = 4.09; $t$ = 2.73, $p$ = .007, $r_{pb}$ = .20). However, the overall teaching self-efficacy of ITAs and USTAs was similar (STEM GTA-TSE ITA = 4.17, USTA = 4.08; $t$ = 1.01, $p$ = .316, $r_{pb}$ = .08). The effect size for the instructional strategies was close to moderate (Cohen, 1988), but as both

means were greater than 4 (out of 5), ITAs and USTAs both felt confident in their abilities to plan, prepare, execute, and evaluate their classes, which is similar to other GTA results (Prieto & Altmaier, 1994; Prieto et al., 2007).

## IV. Discussion.

The purpose of this study was to develop an instrument for measuring teaching self-efficacy of STEM GTAs and explore relationships between STEM GTA teaching self-efficacy, GTA professional development, and teaching experience. Essential to this process was to work toward establishing measurement reliability and both face and construct validity of the STEM GTA teaching self-efficacy measure. Assertions related to instrument validity, reliability, and correlations must be viewed, however, as sample dependent.

### *A. Instrument Modification and Development.*

The STEM GTA teaching self-efficacy instrument as developed has two subscales, instructional strategies and learning environment, which may be used individually or combined as a single measure of teaching self-efficacy. This structure is not unlike the Teacher Sense of Efficacy Scale (Tschannen-Moran & Hoy, 2001); which has three factors – student engagement, instructional strategies, and classroom management – that could be used to measure overall teaching self-efficacy. The STEM GTA-TSES, however, did not include items relating to classroom management since STEM GTAs are teaching adults, not children. The two factor structure provides more flexibility in using the instrument. It not only provides a global score of teaching self-efficacy, but if self-efficacy of STEM GTAs relating to classroom instruction or ability to create an active and positive learning environment is needed, then this instrument also offers that option. When this instrument is used to evaluate GTA professional development or individual GTA development, the subscales can be useful in determining where changes are occurring in GTA teaching self-efficacy.

Possible limitations include sample size, scale sensitivity, and face validity with STEM GTAs. Costello and Osborne (2005) argue that in exploratory factor analysis, a ratio of at least 10 participants to each item in the instrument provides an average of less than one (0.70) item misclassified on the wrong factor. This study is close to achieving the desired sample size, with a ratio of nine participants per item. Also by Costello and Osborne's categorization, the individual factors in the STEM GTA-TSES (Table 1) are good; "a factor with…5 or more strongly loaded items (.50) are desirable and indicate a solid factor" (p. 5). Additionally, the CFA with all GTAs reveals a solid second-order factor structure (Figure 1), which indicates that the two specific self-efficacy subscales can collapse into one broader teaching self-efficacy factor. This could be predicted from social cognitive theory, which indicates that teacher efficacy should be task specific (Bandura, 1997; Pajares, 1996a). Given the complex nature of the teaching task, a multiple factor structure should be expected (Dellinger, et al., 2008; Henson, 2002), but additional research with this instrument should use CFA to confirm the two-factor structure. Another possible limitation is the sensitivity in a five-point scale, which is required by the data scanning software used. The measure might be more sensitive with a larger scale. Bandura (2006), for example, advocated for a 0 -100 scale in 10 point increments and there is evidence in a middle school environment that this scale may be more predictive in a regression equation on achievement than a 6 point scale (Pajares, Hartley, & Valiante, 2001). Although

face validity was measured with research faculty, it was not determined for the respondents. Further research with this instrument should include a debriefing interview with STEM GTAs after they have completed the items in the measure. This interview can be used to determine how the STEM GTAs interpret the items and if they feel that these items represent the concept of teaching self-efficacy.

*B. Related Relationships.*

A correlation analysis with the originating university GTAs suggested that the STEM GTA-TSES and learning environment and instructional strategies self-efficacy sub-scales were related to both theoretical and empirical constructs. There was a high correlation between the two subscale factors in this instrument and the between each subscale and the overall teaching self-efficacy. The two subscales measure related activities in the classroom - learning environment and instructional strategies – and since overall teaching self-efficacy is determined from both subscales, they should be correlated.

Prior research (Burton et al., 2005; Liaw, 2004; Prieto & Altmaier, 1994; Prieto et al., 2007; Tollerud, 1990) has generally shown a positive effect of GTA teaching experience on self-efficacy. With this sample of STEM GTAs the *instructional* subscale correlated with all measures of teaching experience, but the *learning* subscale and STEM GTA-TSES did not correlate with quarters of teaching experience. With half of this sample ITAs, this is consistent with Kim (2009); who also found that the student engagement subscale of the Teacher Sense of Efficacy Scale (Tshannen-Moran & Hoy, 2001) did not correlate to teaching experience but the classroom management and instructional subscales did. Additionally, both sub-scales and the STEM GTA-TSES were highly correlated with measures of teaching experience that were GTA self-reports (Table 2). These self-reports asked the GTAs to rate their experience from beginner to expert and to rate how they compared themselves to others. Each of these should be correlated; all three are self-assessments that should be reinforcing each other. Award winning professors also use referential comparisons as a source for teaching self-efficacy, probably because there are few objective measures of teaching quality available in college instruction (Morris & Usher, 2011). It is not surprising that this is also appears true of STEM GTAs.

Examining the SETI-A (Prieto & Altmaier, 1994; Prieto et al., 2007; Tollerud, 1990) showed that most of these items are similar to the *instructional* factor rather than the *learning* factor. This may also account for the different results of *learning* factor and STEM GTA-TSES with teaching experience. The learning factor measured items that would be ideal in the learning environment of a student-centered classroom. There is evidence that student centered teaching is less common in hard scientific disciplines (Lindblom-Ylanne et al., 2006; Luo, Grady, & Bellows, 2001). If there is less student-centered teaching by these STEM GTAs then it wouldn't be surprising for teaching experience to be uncorrelated to the learning sub-scale.

The populations used in the prior studies on GTA teaching self-efficacy were very different in composition than the STEM GTAs in this study. Two-thirds of these GTAs had less than two years teaching experience compared to about 40% in other studies (Liaw, 2004; Prieto & Altmaier, 1994). Being completely responsible for a course was rare in this sample (5%) whereas over 40% of GTAs were completely responsible for a course in other studies (Prieto & Altmaier, 1994; Prieto et al., 2007). Additionally, this sample was completely STEM GTAs, whereas STEM GTAs comprised less than 50% of two of the studies (Prieto & Altmaier, 1994; Prieto et al., 2007), while the rest had no STEM GTAs (Burton et al., 2005; Liaw, 2004;

Tollerud, 1990). These demographic differences might have influenced the way teaching experience impacted teaching self-efficacy in these groups. In a detailed analysis, Liaw (2004) demonstrated that there was a high collinarity between years of teaching experience and the level of course (beginning or intermediate) the GTAs taught. Interestingly, for those GTAs who taught beginning courses there was no effect on teaching self-efficacy for teaching experience. Therefore the teaching experience effect on teaching self-efficacy may have been partially due to teaching advanced courses; the level of courses taught by this sample of STEM GTAs was not determined however it was fair to assume many taught in large introductory laboratory courses as was common in STEM departments. Finally, STEM GTAs taught many different courses during their teaching experience. In this sample the mean was 3.2 courses in 5.2 quarters of teaching. Since teaching self-efficacy was context specific, teaching a large number of different courses might have impacted and contributed to the lack of a correlation between the GTA-TSES and learning factor with quarters of teaching experience.

According to social cognitive theory (Bandura, 1997), professional development in a task should increase self-efficacy of the person in performing that task. Prior research has shown mixed results for effects of GTA professional development on teaching self-efficacy (Meyers et al., 2007; Prieto & Altmaier, 1994; Prieto & Meyers, 1999; Prieto et al., 2007; Tollerud, 1990), which may be because of the measures of professional development used in these studies, either time in professional development or simple presence/absence of professional development. In this study, GTA perception of professional development (DeChenne et al., 2012) correlated moderately with both subscales and the STEM GTA-TSES, whereas all three of these scales showed only a small correlation with hours of professional development. The mixed results from prior studies may be related to the quality of the GTA professional development received. Good GTA professional development would include mastery experiences, vicarious experiences, and verbal persuasions that should increase teaching self-efficacy. However, if the quality of the GTA professional development was poor, then there would be little or no correlation to teaching self-efficacy. In this study, this premise was demonstrated through the higher correlation to the GTA perception of professional development than to the hours of professional development.

*C. Implications for Practice.*

How do these results help us understand and improve the teaching of STEM GTAs? The STEM GTAs do have a relatively high teaching self-efficacy which is consistent with other GTA research (Prieto & Altmaier, 1994; Prieto et al., 2007) but it could be improved and that should improve the instruction of the STEM GTAs and then the learning of their undergraduate students. The STEM GTA-TSES can be useful in examining several aspects of professional development of STEM GTAs. It can be used by faculty working with GTAs to assess the impact of specific GTA professional development programs, supervision, and teaching experiences on GTA teaching self-efficacy. The *instructional* factor relates to activities needed to prepare and teach a class, and these are relatively concrete items. The *learning* factor focuses on more complex concepts involved in promoting and providing an active, positive, and respectful classroom environment, which can be more difficult to implement in an actual classroom or laboratory setting. This subscale will help GTA professional developers and supervisors evaluate their GTAs' readiness for classroom challenges.

Results of this study indicate that STEM GTA professional development needs to be increased and improved. The average STEM GTA had half a week of professional development

in teaching which they rated as average (3.2 out of 5). Yet the perception of learning in their GTA professional development was moderately correlated with teaching self-efficacy and only mildly correlated with hours spent in GTA professional development. This indicates that good (or at least better) professional development improves GTA teaching self-efficacy. The types of GTA professional development within this sample ranged from a short (two to three days) professional development which discussed how to teach and how students learn in the last half day, to a required one quarter course in teaching and learning in addition to weekly group GTA meetings with the course instructor for the laboratory the GTAs were teaching (personal communication with the departments in the study). One department offered a year-long (three quarter) series on teaching and learning for their GTAs (although only the first quarter was required). Since there was a correlation between time spent in professional development and perception of learning, increasing the time GTAs spend learning about teaching and learning should improve their teaching self-efficacy. Additionally, more time spent within the professional development on teaching and learning would improve their teaching self-efficacy.

The participants in this sample had little to no feedback (verbal persuasions) on their teaching (personal communication with departments in the study). None of the departments' video recorded the GTAs teaching, most of the GTAs were never observed by either the course instructor or other GTAs, and over half of this sample did not even receive student evaluations. Teaching experience should provide mastery experiences, however without any feedback or reflection, the teaching experience does not provide the mastery experiences required to affect teaching self-efficacy. This is alternative explanation for the poor correlation between teaching self-efficacy and quarters of teaching experience in this study. This research suggests that the main vehicle for gaining experience in teaching, the teaching assistantship, is not providing an effective experience for these STEM GTAs. The teaching assistant experience needs to be moved beyond the need for instruction coverage by the department to consider the needs of preparing future faculty and the needs of the current undergraduate students served by the department. This lack of correlation between teaching experience and teaching self-efficacy suggests that only the first consideration is being utilized by many STEM departments for their GTAs.

The literature contains various suggestions for GTA professional development best practices (for a review, see Park, 2004). However, social cognitive theory (Bandura, 1986, 1997) suggests that also including several video recorded teaching sessions with both peer and instructor feedback would greatly improve GTA teaching self-efficacy (and therefore teaching effectiveness). This activity encompasses mastery experiences, vicarious experiences, and verbal persuasions, all of which will improve teaching self-efficacy. In this one exercise, which could be done either through a GTA professional development experience and/or within the teaching requirements for the GTA (for example during weekly GTA meetings), there is a chance to significantly effect GTA teaching. Repeating this exercise would greatly improve the teaching of the GTAs. Having the GTAs record a teaching session, showing that session to the GTA group, soliciting feedback from the group and feedback from the professional development or course instructor, would provide the GTA a chance to reflect on his/her own teaching. If repeated, various aspects of teaching could be focused on for each session. It is not even necessary for the whole teaching session to be watched by the GTA group; instead the GTA could pick a small section for feedback. This also increases the reflection of the GTA, since they must determine what to show their fellow students/instructor; they may emphasize a question they have about their teaching or something in which they excelled. Providing repeated

experiences for mastery, vicarious, and verbal persuasions will greatly improve their teaching and turn the teaching experience into a true apprenticeship in teaching, one which improves the GTAs instructional abilities, prepares them for a future role as faculty, and improves the learning of their students.

This instrument will also be useful for research on STEM GTA teaching and can be used to study STEM GTA teaching self-efficacy. For any quantitative study in which teaching self-efficacy is a variable of interest, this could be a valuable instrument. The instrument could also be beneficial in research on longitudinal effects of GTA professional development and teaching experience on teaching self-efficacy. Henson (2001) expressed that it is time for teaching self-efficacy studies to move beyond correlations. Using this instrument, factors influencing the development of STEM GTA teaching self-efficacy have been modeled (manuscript submitted). Further research could explore the explicit relationships between STEM GTA teaching self-efficacy, teaching performance, and student achievement; determine other factors influencing STEM GTA teaching self-efficacy beyond teaching experience and professional development; and influence research into social cognitive theory based professional development for STEM GTAs.

## Acknowledgements

The authors would like to acknowledge the contributions of Donald Heer and Niki Schulz of the College of Engineering for their insightful comments on possible items in the STEM GTA-TSES. We would also like to thank Dr. Larry Flick of the Science and Mathematics Education Department for reviewing the STEM GTA-TSES. An earlier version of this article was presented at and greatly improved by comments from the 2010 annual conference of the American Educational Research Association in Denver, CO. Of course, this research would not have been possible without the participation of STEM GTAs from six universities; thank-you very much for participating.

## References

Abraham, M.R., Cracolice, M.S., Graves, A.P., Aldhamash, A.H., Kihega, J.G., Gil, J.G.P., et al. (1997). The nature and state of general chemistry laboratory courses offered by colleges and universities in the United States. *Journal of Chemical Education, 74*(5), 591-594. doi: 10.1021/ed074p591

Ashton, P.T., & Webb, R.B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.

Austin, A.E. (2002). Creating a bridge to the future: Preparing new faculty to face changing expectations in a shifting context. *The Review of Higher Education, 26*(2), 119 - 144. doi: 10.1353/rhe.2002.0031

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory* (1st ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W.H. Freeman and Company.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pagares & T. Urdan (Eds.), *Self-Efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age Publications, Inc.

Borko, H., & Putnam, R. (Eds.). (1996). *Learning to teach*. New York, NY: Macmillan.

Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: SAGE Publication, Inc.

Burton, J.P., Bamberry, N.J., & Harris-Boundy, J. (2005). Developing personal teaching efficacy in new teachers in university settings. *Academy of Management Learning & Education, 4*(2), 160-173. doi: 10.5465/AMLE.2005.17268563

Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS / Windows*. Thousand Oaks, CA: Sage.

Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chang, T.S., Lin, H.H., & Song, M.M. (2011). University faculty members' perceptions of their teaching efficacy. *Innovations in Education and Teaching International, 48*(1). doi: 10.1080/14703297.2010.543770

Chou, C.P., & Bentler, P.M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling* (pp. 37-55). Thousand Oaks, CA: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second ed.): Lawerence Erlbaum Associates, Publishers.

Commander, N.E., Hart, L., & Singer, M. (2000). Preparing tomorrow's faculty: An assessment model to determine institutional needs. *The Journal of Graduate Teaching Assistant Development, 7*(2), 93 - 111.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.

Costello, A.B., & Osborne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), Retrieved from: http://pareonline.net/pdf/v10n17.pdf.

Davis, S.F., & Kring, J.P. (2001). A model for training and evaluating graduate teaching assistants. *College Student Journal, 35*(1), 45 - 51.

DeChenne, S.E. (2012, April). *Towards Development of a Model of Graduate Teaching Assistant Teaching Effectiveness.* Roundtable presented at the annual meeting of American Educational Research Association, Vancouver, BC.

DeChenne, S.E., Anderson, S., Lesseig, K., Li, S., Staus, N., & Barthel, C. (2009). *Graduate student teaching training: Differences in graduate students' training needs.* Paper presented at the American Educational Research Association, San Diego. CA.

DeChenne, S.E., Anderson, S.M., Lesseig, K.R., Li, S.L., Staus, N.L., & Barthel, C. (2012). Towards a Measure of Graduate Student Teaching Professional Development. *Journal of Effective Teaching, 12(1),* 4-19.

Dellinger, A.B., Bobbett, J.J., Olivier, D.F., & Ellett, C.D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*, 751-766. doi: 10.1016/j.tate.2007.02.010

Fagen, A.P., & Wells, K.M. . (2004). *The 2000 national doctoral program survey.* In D. H. Wulff & A. E. Austin (Eds.), *Paths to the professoriate: Strategies for enriching the preparation of future faculty*. San Francisco, CA: Jossey-Bass.

Fencle, H., & Scheel, K. (2005). Engaging students: An examination of the effects of teaching strategies on self-efficacy and course climate in a non-majors physics course. *Journal of College Science Teaching. 35(1),* 20-24.

Gist, M., Schwoerer, C., & Rosen, B. (1989). Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology, 74*(6), 884 - 891. doi: 10.1037//0021-9010.74.6.884

Golde, C.M., & Dore, T.M. (2001). *At cross purposes: What the experiences of doctoral students reveal about doctoral education.* Retrieved from http://phd-survey.org/report%20final.pdf

Golde, C.M., & Dore, T.M. (Eds.). (2004). *The survey of doctoral education and career preparation: The importance of disciplinary contexts*. San Francisco, CA: Jossey-Bass.

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika, 19*, 149-161. doi: 10.1007/BF02289162

Hadre, P. L. (2003). The effects of instructional training on university teaching assistants. *Performance Improvement Quarterly, 16*(4), 23 - 39.

Haney, J.J., Czerniak, C.M., & Lumpe, A.T. (1996). Teacher beliefs and intentions regarding the implementation of science education reform strands. *Journal of Research in Science Teaching, 14*(3), 287 - 313. doi: 10.1002/(SICI)1098-2736(199611)33:9<971::AID-TEA2>3.0.CO;2-S

Henson, R.K. (2001). *Teacher Self-Efficacy: Substantive Implications and Measurement.* Paper presented at the Educational Research Exchange, College Station, Texas.

Henson, R.K. (2002). From adolescent angst to adulthood: Substantive implications and measurement dilemmas in the development of teacher efficacy research. *Educational Psychologist, 37*(3), 137 - 150. doi: 10.1207/S15326985EP3703_1

Hoy, A.W. (2003-2004). Self-Efficacy in College Teaching. *Essays on teaching excellence: Toward the best in the academy, 15*(8). Retrieved from http://www.podnetwork.org/publications/teachingexcellence/03-04/V15,%20N8%20Woolfolk-Hoy.pdf

Johnson, E. (1998). *Graduate teaching assistants' beliefs about teaching mathematics.* Ph.D. Dissertation, University of Nebraska - Lincoln.

Jones, J.L. (1993). TA training: From the TA's point of view. *Innovative Higher Education, 18*(2), 147 - 161. doi: 10.1007/BF01191892

Kim, E. (2009). Beyond language barriers: Teaching self-efficacy among East Asian international teaching assistants. *International Journal of Teaching and Learning in Higher Education, 21*(2), 171-180.

Komarraju, M. (2008). A social-cognitive approach to training teaching assistants. *Teaching of Psychology, 35*, 327-334. doi: 10.1080/00986280802374344

Liaw, E.C. (2004). *"How are they different?" A comparative study of native and nonnative foreign language teaching assistants regarding selected characteristics: Teacher efficacy, approach to language teaching/learning, teaching strategies and perception of nativeship.* Doctoral Dissertation, The Ohio State University, Columbus. Retrieved from http://etd.ohiolink.edu/send-pdf.cgi?osu1092324144

Lindblom-Ylanne, S., Trigwell, K., Nevgi, A., & Ashwin, P. (2006). How approaches to teaching are affected by discipline and teaching context. *Studies in Higher Education, 31*(3), 285 - 298. doi: 10.1080/03075070600680539

Luft, J.A., Kurdziel, J.P., Roehrig, G.H., & Turner, J. (2004). Growing a garden without water: Graduate teaching assistants in introductory science laboratories at a doctoral/research university. *Journal of Research in Science Teaching, 41*(3), 211 - 233. doi: 10.1002/tea.20004

Luo, J., Grady, M.L., & Bellows, L. H. (2001). Instructional issues for teaching assistants. *Innovative Higher Education, 25*(3), 209 - 230.

Meyers, S.A., Lansu, M.L., Hundal, J.S., Lekkos, S.K., & Prieto, L.R. (2007). Preparing new psychology instructors to teach undergraduates: Developing confidence and competence. *Journal of Faculty Development, 21*(1), 45 - 54.

Miller, S., Pfund, C., Pribbenow, C.M., & Handelsman, J. (2008). Scientific teaching in practice. *Science, 322,* 1329- 1330. doi: 10.1126/science.1166032

Mills, N., & Allen, H.W. (2007). Teacher self-efficacy of graduate teaching assistants of French. In J. Siskin (Ed.), *From thought to action: Explaining beliefs and outcomes in the foreign language program* (pp. 213-234). Boston, MA: Heinle and Heinle.

Moore, T., & Diefes-Dux, H. (2004). *Developing model-eliciting activities for undergraduate students based on advanced engineering content*. Paper presented at the ASEE/IEEE Frontier in Education Conference, Savannah, GA.

Morris, D. B., & Usher, E. L. (2011). Developing teaching self-efficacy in research institutions: A study of award-winning professors. *Contemporary Educational Psychology, 36*, 232-245. doi: 10.1016/j.cedpsych.2010.10.005

O'Neal, C., Wright, M., Cook, C., Perorazio, T., & Purkiss, J. (2007). The impact of teaching assistants on student retention in the sciences: Lessons for TA training. *Journal of College Science Teaching, 37,* 24 – 29.

Pajares, F. (1996a). *Assessing self-efficacy beliefs and academic outcomes: The case for specificity and correspondence.* Paper presented at the American Educational Research Association, New York, NY.

Pajares, F. (1996b). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578. doi: 10.2307/1170653

Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and Evaluation in Counseling and Development, 33*, 214-221.

Park, C. (2004). The graduate teaching assistant (GTA): Lessons from North American experience. *Teaching in Higher Education, 9*(3), 349 - 361.

Piccinin, S., & Fairweather, D. (1996-97). A local analysis of TA training needs: Implications for the planning and development of TA training programs. *The Journal of Graduate Teaching Assistant Development, 4*(1), 23 - 34.

Pomalaza-Raez, C.A., & Groff, B.H. (2003). Retention 101: Where robots go...students follow. *Journal of Engineering Education, 92*(1), 85 - 90.

Prieto, L.R., & Altmaier, E.M. (1994). The relationship of prior training and previous teaching experience to self-efficacy among graduate teaching assistants. *Research in Higher Education, 35*(4), 481 - 497. doi: 10.1007/BF02496384

Prieto, L.R., & Meyers, S.A. (1999). Effects of training and supervision on the self-efficacy of psychology graduate teaching assistants. *Teaching of Psychology, 26*(4), 264 - 266. doi: 10.1207/S15328023TOP260404

Prieto, L.R., & Scheel, K. (2008). Teaching assistant training in counseling psychology. *Counseling Psychology Quarterly, 21*(1), 49 - 59.

Prieto, L.R., Yamokoski, C.A., & Meyers, S. A. (2007). Teaching assistant training and supervision: An examination of optimal delivery modes and skill emphasis. *Journal of Faculty Development, 21*(1), 33 - 43.

Prieto Navarro, L. (2005). *Las creencias de autoeficacia docente del profesorado universitario*. Madrid: Universidad Pontifica Comillas.

Rushin, J.W., De Saix, J., Lumsden, A., Streubel, D.P., Summers, G., & Berson, C. (1997). Graduate teaching assistant training: A basis for improvement of college biology teaching and faculty development? *American Biology Teacher, 59*(2), 86 - 90.

Sargent, L.D., Allen, B.C., Frahm, J.A., & Morris, G. (2009). Enhancing the experience of student teams in large classes: Training teaching assistants to be coaches. *Journal of Management Education, 33*(5), 526-552. doi: 10.1177/1052562909334092

Shannon, D.M., Twale, D.J., & Moore, M. . (1998). TA teaching effectiveness: The impact of training and teaching experience. *The Journal of Higher Education, 69*(4), 440-446. doi: 10.2307/2649274

Skaalvik, E.M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*(3), 611-625. doi: 10.1037/0022-0663.99.3.611

Sundberg, M.D., Armstrong, J.E., & Wischusen, E.W. (2005). A reappraisal of the status of introductory biology laboratory education in U.S. College and Universities. *The American Biology Teacher, 67*(9), 525-529. doi: 10.1662/0002-7685(2005)067[0525:AROTSO]2.0.CO;2

Tanner, K., & Allen, D. (2006). Approaches to biology teaching and learning: On integrating pedagogical training into the graduate experiences of future science faculty. *CBE Life Sciences Education, 5*(1), 1-6. doi: 10.1187/cbe.05-12-0132

Taylor, R.L., Heer, D., & Fiez, T.S. (2003). Using an Integrated Platform for Learning (TM) to reinvent engineering education. *IEEE Transactions on Education, 46*(4), 409 - 418.

Tollerud, T. (1990). *The perceived self-efficacy of teaching skills of advanced doctoral students and graduates from counselor education programs.* Doctoral Dissertation, The University of Iowa, Iowa City, IA.

Torvi, D.A. (1994). Engineering graduate teaching assistant instructional programs: Training tomorrow's faculty members. *Journal of Engineering Education, 83*(4), 375 - 381.

Tschannen-Moran, M., & Hoy, A.W. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education, 17*, 783 - 805. doi: 10.1016/S0742-051X(01)00036-1

Tschannen-Moran, M., Hoy, A.W., & Hoy, W.K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202 - 248. doi: 10.2307/1170754

Weber, R.J., Gabbert, A., Kropp, J., & Pynes, P. (2007). Creating the teaching professor: Guiding graduate students to become effective teachers. *The Journal of Scholarship of Teaching and Learning, 7(1),* 45 – 63.

Verleger, M., & Velasquez, J. (2007, October). *An engineering teaching assistant orientation program: Guidelines, reactions, and lessons learned from a one-day intensive training program.* Paper presented at the ASEE/IEEE Frontiers in Education Conference, Milwaukee, WI.

Young, S.L., & Bippus, A.M. (2008). Assessment of graduate teaching assistant (GTA) training: A case study of a training program and its impact on GTAs. *Communication Teacher, 22*(4), 116-129. doi: 10.1080/17404620802382680

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 124 – 135.

# Factors affecting timely completion of a PhD: a complex systems approach

**Jegar Pitchforth[1], Stephanie Beames[1], Aleysha Thomas[1], Matthew Falk[1], Charisse Farr[1], Susan Gasson[2], Sri Astuti Thamrin[1] & Kerrie Mengersen[1]**

*Abstract: Completing a PhD on time is a complex process, influenced by many interacting factors. In this paper we take a Bayesian Network approach to analyzing the factors perceived to be important in achieving this aim. Focusing on a single research group in Mathematical Sciences, we develop a conceptual model to describe the factors considered to be important to students and then quantify the network based on five individual perspectives: the students, a supervisor and a university research students centre manager. The resultant network comprised 37 factors and 40 connections, with an overall probability of timely completion of between 0.6 and 0.8. Across all participants, the four factors that were considered to most directly influence timely completion were personal aspects, the research environment, the research project, and incoming skills.*

*Keywords: bayesian network, graduation, skills, environment, personal, project*

Timely completion of a PhD is an important outcome for the student, the host university and the economy. However, completion of this programme in the required timeframe is dependent on many interacting factors. In this study, we develop a statistical complex systems approach to identify and quantify the important factors and their interactions that are perceived to impact on timely completion of a PhD in Statistics in an Australian university. We define timely completion to be within 3.5 years. We construct a Bayesian Network (BN) to describe these inter-relationships (Pearl, 1985); the construction and interpretation of a BN is described in more detail below. The conceptual model for the BN was developed collectively and then quantified by five candidates: three students, a supervisor and a university research students centre manager.

Australian universities receive competitive funding for PhD enrolments and successful completions, yet completion rates are well below 100% (Jiranek, 2010). It is therefore of interest to institutions and government bodies if predictive or causal factors can be identified which may assist students to progress through their studies, or to better prepare for and support the postgraduate supervision of students. Gaining an understanding of factors affecting timely completion, and providing such information to prospective students to better equip them, could assist with attrition rates.

We were interested in three main questions. First, what is the overall perceived probability of timely completion of a PhD in Statistics at QUT? Second, what factors were most influential in timely completion, and how do these differ between the five candidates? Third, what is the change in the probability of timely completion under specified scenarios? The scenarios chosen for evaluation are detailed below.

[1] Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane 4001
[2] Research Students Centre, Queensland University of Technology, GPO Box 2434, Brisbane 4001

## I. Background.

There is a substantial and growing literature that identifies important factors associated with completion of a PhD research project. In a meta-analysis based on over 160 references, Bair and Hanworth (2005) associate persistence rates with funding and socialization, and completion rates with positive and supportive mentor relationships. Maher, Ford and Thompson (2004) list a suite of factors frequently linked to completion time of doctoral degrees, including availability of funding resources, the nature of the advising relationship, the extent to which students receive research preparation and opportunities, and individual student concerns about marital, family or health problems. Seagram, Gould and Pyke (1998) also list several potential factors that impact on timely completion, including gender, discipline, supportive relationship, financial situation and enrolment status. A linear regression analysis of the results of a survey of 154 graduates of doctoral programs in three discipline areas at York University revealed that beginning a dissertation early in the program, remaining with the original topic and supervisor, meeting frequently with the supervisor and collaborating with supervisor on conference papers were important indicators, but only explained 30% of the total variance.

The role of supervisors has also been examined by other authors; see, for example, Zhao (2007). A Procrastination Inventory proposed by Muszynski and Akamatsu (1991) revealed that demographic and situational variables, including a supportive advisor, finding a topic of interest, making the dissertation a top priority and living close to the university were predictive of success, but that specific research interests or measures of needs or values were not significant predictors. Psychological factors have also been investigated by other authors; see, for example, Kearnes, Gardiner and Marshall (2009) who focused on the important issue of self-sabotaging behaviour due to over-committing, procrastination and perfectionism.

Another important domain that has been considered in the literature is the role of cohort partnerships and groups (Witte & James, 1998) and peer-to-peer support (Devenish, Dyer, & Jefferson, 2009). Race (Ellis, 2001), type of attendance (Rodwell, 2008) and gender (Maher, Ford, & Thompson, 2004) have also been discussed.

A variety of perspectives about the issue of timely completion have also been considered. Barnes and Austin (2009) considered the role of doctoral advisors from the advisors' perspective. Isaac, Quinlan and Walker (1992) have examined faculty perceptions of the doctoral dissertation, noting in particular field-related differences with respect to characteristics, content and purpose of the doctoral dissertation. The impact of departmental factors has also been identified by other authors; see, for example, de Valero (2001).

The importance of this topic and the intense interest in it is underscored by the large, high profile Council of Graduate Schools Ph.D. Completion Program (2009), conducted in the USA and Canada, and the citations and references therein. The study profiles the following key factors influencing PhD completion: selection, mentoring, financial support, program environment, research mode of the field, and processes and procedures.

There is now a large literature on the underpinning theory and methodology of BNs as well as their application to a wide range of problems. We have previously employed them to address environmental and health outcomes (e.g., Johnson et al., 2009, 2010; Waterhouse et al., 2010), among other areas. They have also been used for over a decade in the education field; see for example the student models of Millan et al. (2010) and Carmona *et al.* (2008), models for assessing diagnostic performance considered by Almond et al. (2007) the general discussion of BNs in educational assessment by Mislevy et al. (2000), and the references therein.

In this study, we focus on a single discipline area, Statistics, in the Mathematical Sciences Discipline at Queensland University of Technology (QUT), Australia. This focus is based on the findings of Seagram et al. (1998), Muszynski and Akamatsu (1991), Isaac et al. (1992) and de Valero (2001), among others, that there are discipline-related and institutional differences in PhD completion time itself, and the factors that potentially impact on it.

## II. Methods.

### *A. Bayesian Networks.*

The first step in constructing a Bayesian Network is the development of a conceptual model of the factors and their interactions. This is depicted as a graphical model, or network, of nodes (representing the factors) and directed arrows (representing the interactions between the nodes). The final outcome (timely PhD completion) is called the terminal node.

The second step of the BN typically involves categorizing each node into a (small) number of states, for example high/low, 0-10/10-20/20+, good/medium/bad. The thresholds for the states are chosen to be meaningful in the context of the problem.

In the third and last step of the construction of the BN, each node is quantified by attaching probabilities to each state of the node. The probabilities are conditional on the states of the nodes feeding into it (as determined by the directed arrows in the network).

A characteristic of the BN is that quantification of a node in the BN depends only on a subset of the network. Thus the whole problem is collapsed into a series of local analyses. Moreover, a variety of sources can be used for quantification of a node, including data, simulation models, statistical or mathematical models, results from literature or previous studies, expert knowledge, and so on. This ability to integrate diverse data is arguably one of the strengths of the BN approach. An iterative approach to designing a BN is described by Johnson et al. (2010).

Once completed, the conditional probabilities 'flow through' the BN to provide an overall probability for each level of the terminal (outcome) node. The network can then be interrogated to identify the major factors influencing the outcome. Moreover, it can be employed to assess the impact of 'evidence' and evaluate scenarios, where these are represented by setting one or several of the nodes in the BN to specified levels.

### *B. Conceptual BN model.*

The structure of the Bayesian Network was developed during a series of meetings with a focus group comprising postgraduate students in Statistics at QUT from December 2010 to January 2011. The focus group comprised 10 unincentivised volunteers, representing approximately 25% of all postgraduates enrolled in the Discipline at the time. While this sample was not probabilistically drawn, it was broadly representative with respect to personal demographics (age, gender, cultural background) and stage of completion. Based on the focus group meetings, a list of all possible factors was created, then those that were similar were merged and those that were deemed to be beyond the scope of the study, namely were removed. Factors were then classified into groups, which became the nodes of the network. Each of these nodes was then assigned binary states and operational definitions (Table 1). These factors related to external political and financial environments, including the following: government attitudes to higher education, government funding for postgraduate students, global financial status, national

financial status. Subjects were unable to quantify the impact of environments other than the current one, based on their own experience.

*C. Quantification.*

The conceptual BN was translated to the software package Genie for probabilistic quantification. Five participants quantified the BN model. The first (A1) was a former domestic doctoral student, the second (A2) a current domestic PhD student, and the third (A3) a current international student. The fourth participant (B1) was a supervisor of these doctoral students, and the fifth (C1) was the manager of the university research students' centre.

The network was quantified by each participant independently, with guidance from two of the coauthors. The guidance provided was in the form of a structured statement describing a Bayesian Network, giving definitions of the nodes, and providing an example of how to complete the required conditional probability tables. The statement was provided to all participants.

Participants were taken through each external node and asked to quantify the probability of each node being in the positive state. For internal nodes, participants were asked to complete the underlying conditional probability tables. For illustration, an example question that was asked of a subject in order to quantify the network is as follows: 'if the factors that directly influence this node are all conducive, what is the probability of this factor being conducive". As previously stated, the definitions of relevant nodes and states (conducive, not conducive) had been defined for the subject. The probabilities provided by the subject were then confirmed through statements such as, 'this value would indicate that in x out of 10 times, or for x out of 10 students, this factor would be conducive, given that all of the input nodes are conducive'. Similar questions were asked for the other combinations of states making up the conditional probability tables. The subject was then invited to evaluate the full set of probabilities for consistency and relative magnitude. This process was repeated for all nodes in the network. Where subjects found this process difficult, they were alternatively asked to weight the importance of each of the input nodes. The weights were then standardized to equal 1, and used as coefficients in a linear regression with indicator variables representing the input nodes. The outputs of the regression model were used as inputs into the conditional probability tables for the node under consideration.

*D. Analysis and Interrogation.*

Final probabilities depicted in the output node were recorded as representations of each participant's perceived probability of timely completion. Internal nodes feeding directly into the model were independently interrogated to determine their effect on the stated probability in the output node. Finally, each node was interrogated independently to determine its final effect on the output node in order to determine any unexpected effects.

**III. Results.**

*A. Overall network structure.*

Figure 1 depicts the conceptual BN model developed in this study. The network includes 37 nodes and 40 connections, indicating that three nodes in the network connect to two other nodes

each. Four internal nodes feed directly into the outcome node, each with their own network of factors influencing their state. Table 1 provides a full list of all included nodes with their possible states and operational definitions.

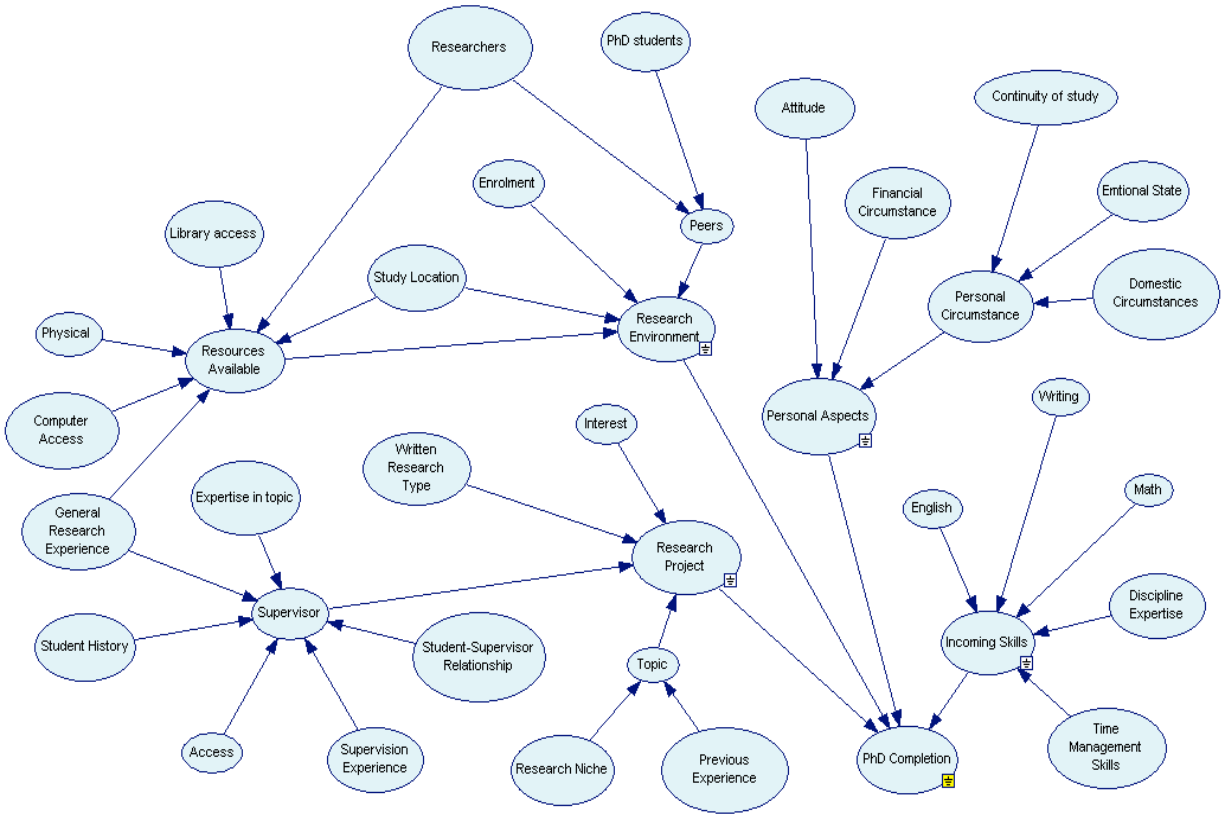


**Figure 1.** Overall structural diagram of the Bayesian Network.

**Table 1.** Structure nodes, states and operational definitions.

| Node | Levels | Definition |
|---|---|---|
| Time Management Skills | Adequate/ Inadequate | The ability of the student to plan and prioritise tasks to meet deadlines set by the university or supervisors. |
| Discipline Expertise | Adequate/ Inadequate | The knowledge of the student regarding their discipline at the time of enrolment. |
| Math | Adequate/ Inadequate | The student's general ability to understand and use mathematical logic. |
| Writing | Adequate/ Inadequate | The student's general ability to clearly communicate their thoughts in writing. |
| English | First Language/ Not First Language | The student's general level of skill with the English language |
| Incoming Skills | Adequate/ Inadequate | The research and management skills of the student at the time of enrolment. This is broadly defined as English, Writing, Math, Discipline Expertise and Time Management Skills |
| Domestic Circumstance | Conducive/ Non-conducive | The living arrangement of the student. This may vary, but whether it is conducive depends on the student. |
| Emotional State | Positive/ Negative | How the student feels about life in general at any period during their degree. |
| Continuity of Study | Conducive/ Non-conducive | Whether the student is returning to study after a period of time, or is continuing on directly after a different degree. |
| Personal circumstance | Adequate/ Inadequate | The family and social circumstances of the student. Broadly defined as the continuity of study, emotional state and domestic circumstances |
| Financial circumstance | Adequate/ Inadequate | The financial position of the student. This is defined as their ability to meet their financial obligations. |
| Attitude | Conducive/ Non-conducive | The student's perspective of how to approach challenges relating to their degree. |
| Personal Aspects | Conducive/ Non-conducive | The collection of all factors related to a student's non-academic life. These are broadly defined as Attitude, Financial Circumstance and Personal Circumstance |
| PhD Students | Useful/ Not Useful | The presence and helpfulness of other PhD students. This might include their ability to resolve academic, administration or personal issues. |
| Researchers | Useful/ Not Useful | The presence and helpfulness of relevant researchers. This might include their ability to resolve academic, administration or personal issues. |
| Peers | Useful/ Not Useful | The presence and helpfulness of other PhD students and Researchers collectively. This might include their ability to resolve academic, administration or personal issues. |
| Enrolment | Full Time/ Part Time | Whether the student is enrolled full time or part time. A full time load is 20 hours per week, whereas a part time load is 10 hours per week |
| Study Location | Internal/ External | Whether the student is based on or off campus. This is defined by whether they have a designated workspace on the University campus. |
| Research Environment | Conducive/ Non-conducive | The general culture of research and physical environment in which the student exists. This might include whether the student is encouraged to attend conferences, or whether the campus (or home if the student studies externally) is safe and comfortable to work in |
| Library Access | Adequate/ Inadequate | The resources and access provided by the University Library. This would include books and journal subscriptions, and access to outside |

| | | |
|---|---|---|
| | | libraries. |
| Physical | Adequate/ Inadequate | The physical resources of the University, such as car parks, lecture halls and study space. |
| Computer Access | Adequate/ Inadequate | The availability and appropriateness of computer -based resources and assistance. This includes physical hardware such as desktop and laptop computers as well as software licences. |
| General Research Experience | Adequate/ Inadequate | The supervisor's previous experience in academic research at the time of enrolment. This could be defined by the number of publications produced or the length of time actively involved in research. |
| Resources Available | Adequate/ Inadequate | The general availability of resources related to the completion of a Research Higher Degree. This is broadly defined as Library Access, Physical resources, Computer access and General Research Experience. |
| Interest | High/ Low | The student's interest in their thesis topic |
| Written Research type | Publication/ Standard Report | The type of thesis submission the student nominates. Publication required that all sections of the thesis consist of published papers, while a standard report is approved by a panel. |
| Expertise in topic | Adequate/ Inadequate | The knowledge of the supervisor regarding their expertise in the specific thesis subject at the time of enrolment. This might be determined by number of papers published on the subject, or length of time spent researching the substantive area. |
| Student History | Mostly Successful/ Mostly Unsuccessful | The success record of the supervisor regarding previous postgraduate students. This is determined by the number of students completing on time divided by the number of students supervised. |
| Access | Adequate/ Inadequate | The availability of the supervisor for meetings, comments and feedback. This is determined largely by the student's need to access the supervisor. |
| Supervision Experience | Adequate/ Inadequate | The experience of the supervisor with supervising postgraduate students. This may be judged by the number of students previously supervised or the length of time spent actively supervising students. |
| Student-supervisor history | Positive/ Negative | The relationship and history between the student and supervisor prior to enrolment. This may include any personal or academic relationships within or without the context of the research higher degree. |
| Supervisor | Helpful/ Not Helpful | The helpfulness and timeliness of the supervisors comments and feedback. This may be judged by the comprehensiveness, relevance and correctness of comments. |
| Research Niche | Specific/ General | The specificity of the student's chosen thesis topic. This may be determined by the number of substantive areas in which the student considers their work relevant or the breadth of literature review (as judged by the number of publications and journals included) required to establish a theoretical base. |
| Previous Experience | Adequate/ Inadequate | The student's previous experience with their research topic. This may include study in the area, but may also include relevant research or industry roles previously held by the student. |
| Topic | Conducive/ Non-conducive | The topic of the student's thesis in relation to their experience. This is broadly defined as the Research Niche of the thesis and Previous Experience of the student |
| Research Project | Conducive/ Non-conducive | All aspects of the student's degree related to the specifics of their research project. This is broadly defined as their Interest, Written Research Type, Supervisor and Topic. |

*B. Overall probability of timely completion.*

The output showed a perceived probability of timely PhD completion in Statistics at QUT ranging from 68% to 79% (Table 2). Amongst students, the domestic current student perceived the highest probability of timely completion (79%), followed by the current international student. The supervisor perceived the second lowest probability of timely completion (70%), and the former domestic student perceived the lowest probability (68%). The research manager (C1) held the most optimistic overall view of the probability of timely completion (80%); although this was still within keeping with the other estimates, it was higher by 10% than the supervisor (B1).

**Table 2.** Final outcome probability of timely completion of a PhD based upon user beliefs.

| Network | A1 | A2 | A3 | B1 | C1 |
|---|---|---|---|---|---|
| Probability | 0.68 | 0.79 | 0.72 | 0.70 | 0.80 |

*C. Most influential factors.*

The most influential factors were found to be those feeding directly into the terminal node (timely completion). Results of the interrogation of these nodes are presented in Table 3 and depicted as radar plots in Figure 2.

These analyses revealed that all four factors contribute substantially to the probability of timely completion. Moreover, while low levels of one or two of the identified factors can deplete the probability to around 0.5 (a 50/50 chance of timely completion), there is almost unanimous agreement that low levels of more than two factors reduces this probability to less than 0.5.

The largest differences in the probabilities awarded to the different combinations of factors were observed between the supervisor (B1) and research manager (C1). Compared with the research manager, the supervisor showed much greater concern about timely completion for low levels of the research project, either alone when all other factors were at high levels, or with low levels of personal aspects when the other two factors were at high levels. In contrast, the research manager showed greater concern than the supervisor when the research environment was a low level, either alone with all other factors were at high levels, or paired with low levels of incoming skills and/or personal aspects.

The strength of influence of the different factors was also evaluated for each respondent. All three students and the supervisor identified availability of resources and presence of other researchers or PhD students. In addition, the former domestic student identified the research topic, and the current domestic identified attitude, financial and personal circumstances. The PhD supervisor also identified the importance of attitude, emotional state, maths background, previous experience, the research topic and the student-supervisor relationship. While the BN constructed by the research manager similarly revealed the importance of other researchers, other PhD students and the candidate's attitude, it also highlighted continuity of study, previous experience and the research niche.

**Table 3.** Relative influence of direct internal nodes on outcome of interest (timely completion), scaled to range between 0 and 1.

| Personal Aspects | Research Environment | Research Project | Incoming Skills | A1 | A2 | A3 | B1 | C1 |
|---|---|---|---|---|---|---|---|---|
| High | High | High | High | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| | | | Low | 0.8 | 0.8 | 0.8 | 0.8 | 0.7 |
| | | Low | High | 0.7 | 0.7 | 0.7 | 0.6 | 0.9 |
| | | | Low | 0.6 | 0.5 | 0.5 | 0.4 | 0.6 |
| | Low | High | High | 0.7 | 0.7 | 0.8 | 0.8 | 0.5 |
| | | | Low | 0.6 | 0.5 | 0.5 | 0.6 | 0.3 |
| | | Low | High | 0.6 | 0.5 | 0.5 | 0.3 | 0.5 |
| | | | Low | 0.3 | 0.3 | 0.3 | 0.2 | 0.2 |
| Low | High | High | High | 0.7 | 0.7 | 0.7 | 0.8 | 0.8 |
| | | | Low | 0.6 | 0.5 | 0.5 | 0.5 | 0.5 |
| | | Low | High | 0.5 | 0.5 | 0.5 | 0.4 | 0.7 |
| | | | Low | 0.3 | 0.3 | 0.2 | 0.2 | 0.4 |
| | Low | High | High | 0.5 | 0.5 | 0.5 | 0.6 | 0.3 |
| | | | Low | 0.7 | 0.2 | 0.3 | 0.4 | 0.1 |
| | | Low | High | 0.2 | 0.3 | 0.2 | 0.1 | 0.2 |
| | | | Low | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

## IV. Discussion.

This BN analysis revealed the following answers to the three main questions posed in this study. First, despite their different perspectives, there was general agreement among the participants in our study that the overall likelihood of timely PhD completion was around 0.7 to 0.8; that is, that on average just under one student in four will not graduate within the given period. The current domestic student (A2) rated their probability of timely completion as the highest, followed the current international student (A3) and the supervisor (B1). The former domestic PhD student (A1) was the most pessimistic about timely completion.

Across all participants, the Research Project was the most important factor impacting on timely completion, followed by the Research Environment. Interestingly, Incoming Skills and Personal Aspects were judged to be equally the least important.

Of course, it is not possible to make any general statements or inferences based on this small sample. However, the study does highlight that students and engaged staff can indeed develop a complex systems model for timely PhD completion, and then quantify it based on their expert judgment. The study also demonstrates that the quantitative outputs are useful for answering questions about PhD completion, including the likelihood of timely completion and the impact of factors contributing to this outcome. Finally, the outputs of a BN can facilitate understanding and decision-making about PhD matters by students, supervisors and university management.

Bayesian Networks based on a person's opinion are difficult to validate externally. Internal validation can proceed via cross-referencing of probabilities, inspection of consistency of probability statements in sub-nodes, and so on. However, nodes like a person's emotional

state cannot be objectively measured. Notwithstanding this, these are important factors, and a BN approach allows these to be included and quantified at a high level.

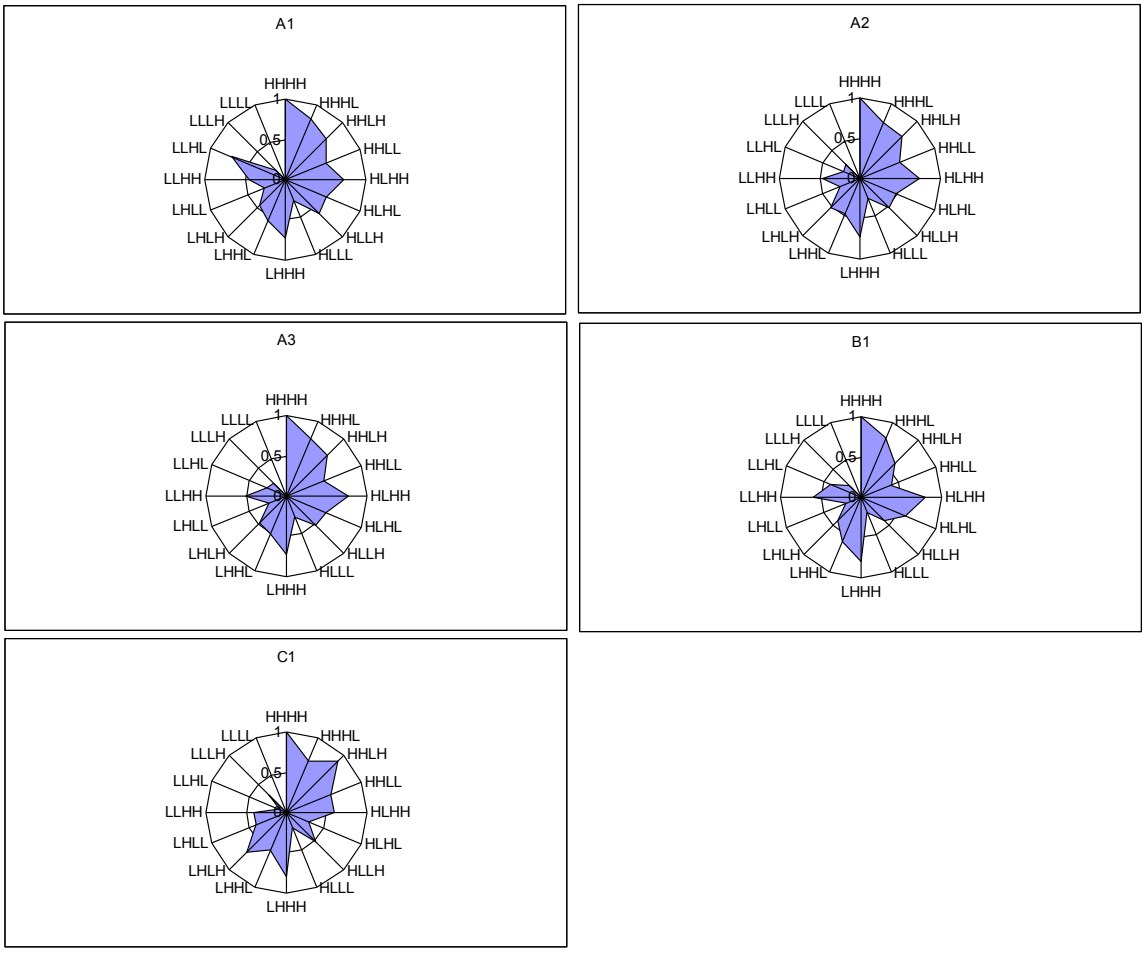


**Figure 2.** Radar plots of relative influence of factors directly influencing the target outcome (timely completion); probabilities are as displayed in Table 1. Factors are Personal Aspects, Research Environment, Research Project and Incoming Skills. Hence 'HHHH' refers to high levels of all factors, HLLL refers to high level of Personal Aspects and low levels of other factors, and so on.

It is noted that the study reported in this paper has focused on factors perceived to be important contributors to timely PhD completion and consequently has provided perceived probabilities of completion through the Bayesian Network analysis. These perceptions could lay the groundwork for further modeling of actual factors and completion rates, and the parallels between the perceived and actual Networks. There were three reasons why this was not pursued as part of the present study. First, perceptions are important in their own right, since they lead to a deeper understanding of the human aspects of the problem and can thus contribute strongly to behavioural and management change frameworks. Second, not all of the factors identified in the study have unequivocal objective metrics that are routinely collected by Universities. Third, confidentiality concerns constrained a more objective analysis, particularly for the defined group of interest in this study. This motivates a larger future study that would address all three of these

issues. Such a study could comprise students and supervisors from a wider range of disciplines, to both generate the network structure and quantify them.

Almost 30 years ago, Abedi and Benkin (1984) described research into reasons contributing to timely completion of degrees as "charitably sparse" (p.4). Twenty years later, Maher, Ford and Thompson (2004) argued that empirical research in this field could still be described as such. There has been considerable literature on the topic in the intervening years, and it is hoped that the present study contributes to our growing understanding of timely completion as a complex system.

## Acknowledgement

The models described in this paper were created using the GeNIe modelling environment developed by the Decision Systems Laboratory of the University of Pittsburgh and available at http://genie.sis.pitt.edu/.

## References

Abedi, J., & Benkin, E. (1984). The effects of students' academic, financial, and demographic variables on time to the doctorate. *Research in Higher Education, 27*(1), 3-14.

Almond, R.G., DiBello, L.V., Moulder, B., & Zapata Rivera, J.D. (2007). Modeling diagnostic assessments with Bayesian networks. *Journal of Educational Measurement, 44* (4), 341-359.

Bair, C., & Hanworth, J. (2005). Doctoral student attrition and persistence: A meta-synthesis of research. *Higher Education: Handbook of Theory and Research, 19*, 481-534.

Barnes, B.J., & Austin, A.E. (2009). The role of doctoral advisors: A look at advising from the advisors perspective. *Innovative Higher Education, 33*(5), 297-315.

Carmona, C., Castillo, G., & Millan, E. (2008). Designing a dynamic Bayesian network for modeling students' learning styles. *ICALT*, 346-350.

Council of Graduate Schools Ph.D. Completion Program (2009). Ph.D. Completion and Attrition: Findings from Exit Surveys of Ph.D. Completers. Monograph. Released Sept. 2009.

Devenish, R., Dyer, S., Jefferson, T., Lord, L., van Leeuwen, S., & Fakerley, V. (2009). Peer to peer support: The disappearing work in the doctoral student experience. *Higher Education Research and Development, 28*(1), 59-70.

Isaac, P.D., Quinlan, S.V., & Walker, M.M. (1992). Faculty perspectives of the doctoral dissertation. *The Journal of Higher Education, 63* (3), 241-268.

Jiranek, V. (2010). Potential predictors of timely completion among dissertation research students at an Australian faculty of sciences. *International Journal of Doctoral Studies* 5, 1-13.

Maher, M.A., Ford, M.E., & Thompson, C.M. (2004). Degree progress of women doctoral students – factors that constrain, facilitate and differentiate. *The Review of Higher Education, 27* (3), 385-408.

Millan, E. Loboda, T., & Perez-de-la-Crus, J.L. (2010). Bayesian networks for student model engineering. *Computers and Education, 55* (4), 1663-1693.

Mislevy, R.J., Almond, R.G., Yan, D., & Steinberg, L.S. (2000). Bayes nets in educational assessment: Where the numbers come from. In K.B. Laskey & H. Prade (Eds), *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence,* pp. 437-446, San Francisco: Morgan Kaufmann.

Muszynski, S.Y., & Akamatsu, T.J. (1991). Delay in completion of doctoral dissertations in clinical psychology. *Professional Psychology: Research and Practice, 22* (2), 119-123.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference.* Cambridge University Press.

Rodwell, J. (2008). Predictors of timely doctoral student capabilities by type of attendance; The utility of a pragmatic approach. *Journal of Higher Education Policy and Management, 30* (1), 65-76.

Seagram, B.C., Gould, J., & Pyke, S.W. (1998). An investigation of gender and other variables on time to completion of doctoral degrees. *Research in Higher Education, 39* (3), 319-335.

de Valero, Y.F. (2001). Departmental factors affecting time-to-degree and completion rates of doctoral students at one land-grant research institution. *Journal of Higher Education, 72* (3), 341-367.

Witte, J., & James, W.B. (1998). Cohort partnerships: A pragmatic approach to doctoral research. *New Directions for Adult and Continuing Education, 79*, 53-62.

Zhao, C.M. (2007). More than a signature: How advisor choice and advisor behaviour affect doctoral student satisfaction. *Journal of Further and Higher Education, 31* (3), 263-281.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 136 – 137.

**Book Review**

# Learner-Centered Curriculum: Design and Implementation

**Maria B. Peterson**[1]

Citation: Cullen, R., Hill, M., & Reinhold, R. (2012). *Learner-Centered Curriculum: Design and Implementation*. Hoboken, NJ: John Wiley & Sons. ISBN: 9781118171004

Publisher Description: Most of the scholarship on learner-centeredness is focused on individual classroom pedagogy, but this book takes learner-centeredness beyond the classroom and asks academic leaders to consider the broader implications of making their institutions fully learner-centered. Systemic change is needed, and curriculum is at the heart of what higher education does. To truly effect change, the curriculum needs to be examined and aligned with learner-centered practices. In this book, the authors offer both design specifications for a learner-centered approach to curriculum as well as practical recommendations for implementation and assessment. The book covers the need for redesigning curriculum, curriculum design in the instructional paradigm, learner-centered design in practice, implementation, program assessment (including a helpful rubric for this), innovation through technology, and construction of learning spaces that support learner-centered curricula.

Curriculum restrictions have been inhibiting faculty members from fully embracing learner-centered practices. Because learners' needs are quickly changing in regards to diversity and technology, classrooms need to redesign programs that meet the needs of "Generation We." Society needs learners who are autonomous, engaged in lifelong learning, flexible, and who can adapt quickly to new situations. Students who are intrinsically motivated will begin to independently present themselves with more opportunities to connect through curriculum and their learning, as well as creating individual self-efficacy (Cullen, Hill, & Reinhold, 2012).

*The Learner-Centered Curriculum: Design and Implementation* is directed at teachers (primarily higher education) who are confronted with the diverse and ever changing needs of undergraduate classes. The book is organized into chapters that go deeply into defining learner centered curriculum, how it can be set up, and how the logistics of grading and classroom arrangement can play a role in proper classroom facilitation. Allowing for student creativity and self-efficacy as part of successful student improvement are common themes throughout the entirety of this book.

The beginning of *The Learner-Centered Curriculum: Design and Implementation* explores the importance of redesigning curriculum and how it plays a role in creating a learner-centered classroom. The authors discuss the increased need and importance for individuals to be resourceful and adaptable within work environments. The authors also reiterate that students should develop a skill set to promote social adjustment with various cultures, as well as become personally and socially responsible and integrative in their learning. The authors suggest that by

[1] Graduate Faculty at Texas Women's University

supplementing current undergraduate courses to foster creativity and learner autonomy, transposition of information occurs that in turn stimulates true, authentic learning.

The text continues to deliberate the implications of learner-centered curriculum and how educators can effectively facilitate this type of learning environment. The authors discuss the importance of learners being actively engaged in the strategy of their learning experiences, creating individual goals, and having the ability to deliberate existing knowledge into newly learned knowledge. Lifelong learning strategies occur when autonomy is developed and students learn how to become more self-directed and determined in their learning.

The authors describe learner-centered practices in three domains: (1) creation of community, (2) sharing of power, and (3) use of assessment for continuous improvement. A useful table for faculty to appraise their current curricular design is included. The table creates a focus on separating the design element of the curriculum into various, evaluative parts and specifically examines: recursion, rigor, richness, relations, community building, power sharing, and assessment. The authors find that when examining learning outcomes for specific programs, faculty members must develop a plan of how to achieve and measure those outcomes and use the rubric to brainstorm possible changes and specific strategies that may be implemented (Cullen et al., 2012).

The conclusion examines the expected outcomes from using a learner-centered curriculum. The authors conclude that the use of valuable assessments both direct and indirect, facilitate the assessor as learner. The importance and need for using technology to support instruction is also discussed. The authors stated that as learners become more adept at monitoring and taking responsibility for their own learning, the use of technological tools will become more effective.

*The Learner-Centered Curriculum: Design and Implementation,* provides a solid foundation in understanding learner centered curriculum and its importance in working with today's students in today's classrooms. The book would be most appropriate for undergraduate faculty members; however, it has promising and feasible structures that any classroom would benefit from, specifically in terms of preparing autonomous, self-efficacious learners.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 138 – 140.

**Book Review**

# Engaging in the Scholarship of Teaching and Learning: A guide to the process, and how to develop a project from start to finish

**Annalee Kodman**[1]

Citation: Bishop-Clark, C. and Dietz-Uhler, B. (2012). *Engaging in the scholarship of teaching and learning: A guide to the process, and how to develop a project from start to finish*. Sterling, Va.: Stylus Pub.

Publisher Description: This is a book for anyone who has ever considered engaging in the scholarship of teaching and learning – known familiarly as SoTL – and needs a better understanding of what it is, and how to engage in it. The authors describe how to create a SoTL project, its implications for promotion and tenure, and how it fosters:

- Increased satisfaction and fulfillment in teaching
- Improved student learning
- Increased productivity of scholarly publication
- Collaboration with colleagues across disciplines
- Contributing to a growing and important body of literature

This guide provides prospective SoTL scholars with the necessary background information, foundational theory, tools, resources, and methodology to develop their own SoTL projects, taking the reader through the five stages of the process: Generating a research question; Designing the study; Collecting the data; Analyzing the data; and Presenting and publishing your SoTL project. Each stage is illustrated by examples of actual SoTL studies, and is accompanied by worksheets to help the reader refine ideas and map out his or her next steps. The process and worksheets are the fruit of the successful SoTL workshops the authors have offered at their institution for many years.

In *Engaging in the Scholarship of Teaching and Learning: A guide to the process, and how to develop a project from start to finish* Cathy Bishop-Clark and Beth Dietz-Uhler provide the academic "for dummies" equivalent of a "how to" book for conducting research in the Scholarship of Teaching and Learning (SoTL). This is by no means a disparagement, but a compliment to a book that is brief, user friendly, and accessible to a range of teachers from pre-K through college in any discipline. The authors write with a teaching audience in mind who may be interested in but hesitant about conducting research on a more formal level. The audience is evident in the tone of the text, which is casual in a way that suggests the authors are sharing their insight and expertise with readers. Most especially, the book is helpful for teachers whose academic backgrounds do not include research training. The authors strive to show the value of this work and to set it apart from reflective teaching practices by providing some basic information in a step-by-step guide that shows how to elevate the research process to achieve

[1] Ph.D. Candidate in Education at the University of Delaware, akkodman@udel.edu

SoTL quality work with the ultimate goal of presenting results in a public forum such as a peer-reviewed publication.

Chapters 1 and 2 lay the groundwork for the book's premise. In Chapter 1, the authors introduce themselves and share their individual stories relating how they became interested in the field of SoTL. They also provide a brief rationale as to the benefits of and need for SoTL. In Chapter 2, Bishop-Clark and Dietz-Uhler quickly sketch the history of SoTL and explore the various definitions of SoTL as well as its distinction from scholarly teaching. Specifically, they speak to newcomers of the field by articulating how SoTL is different: "The difference between what we do informally and what we do in SoTL rests on the formality of our reflections, observations, and sharing of our findings" (p. 9). This formality is the five-step research process described in the next several chapters.

Chapter 3 is an introduction to the five-step research process which, again, Bishop-Clark and Dietz-Uhler differentiate from less rigorous kinds of reflection and learning: "SoTL projects are well defined, carefully and systematically studied, and put in the context of others' research, and the results are shared with a broader community" (p. 18). These continued distinctions are helpful for new researchers to keep in mind as the chapters progress. The authors embed the explanation of the research process in several examples from their personal experiences and illustrate a variety of research designs. The illustrations are clear and jargon free, but the authors are careful to be explicit about the reciprocal nature of the research process, which can deter, distract, and detour new researchers.

Chapters 4 through 8 cover the five-step research process. Chapter 4 "Generating the Research Idea" provides two frameworks for categorizing the research question. Both are helpful given the myriad of questions readers could pose when beginning SoTL research in their classrooms. The chapter also gives several examples of research questions from a variety of disciplines. Additionally, the chapter introduces the concept of the literature review and discusses its importance to the research question and overall research process. This section, however, is extremely brief and may not provide sufficient information for engaging in the often complicated process of researching and writing a literature review. Chapter 5 "Designing the Study" introduces two general categories of research, quantitative and qualitative, as well as mixed-methods. Bishop-Clark and Dietz-Uhler outline the variety of forms these categories take such as experimental, quasi-experimental or descriptive and case study respectively. Again, several examples from their own work illustrate the methods, and they provide annotations for other resources to consult for further information. They also highlight the benefits of conducting a pilot study.

In Chapter 6 "Collecting the Data," Bishop-Clark and Dietz-Uhler discuss several key considerations when collecting the data including determining reliability and validity, defining and operationalizing variables, and choosing a preexisting instrument or creating one. They then break down the collection plan according to the methodological categories listed in the prior chapter such as experimental or case study. In addition, Chapter 6 discusses ethical considerations such as those designated by the local Institutional Review Board (IRB), but this information comes a bit late in the book as well as its consideration within the overall research process. Because IRBs protect participants, researchers, and institutions, are often required, and may significantly affect the research process, this information would be of more use if incorporated sooner. Ethical considerations addressed include informed consent, the right to privacy, the risk of harm, and debriefing. Bishop-Clark and Dietz-Uhler also list several IRB considerations.

Chapter 7 "Analyzing the Data" begins with a few paragraphs that argue why readers should know about data analysis, in other words, why systematic and rigorous quantitative or qualitative analysis is more convincing than experiential or anecdotal evidence. The authors explain how understanding and conducting data analysis is key to contributing to the field as consumers and, presumably after reading this book, as producers of SoTL knowledge. Chapter 7 is then broken into steps toward analyzing qualitative data and quantitative data. The qualitative data section is helpful in that it provides a five-step process, multiple tables with data transcription and qualitative analysis software resources, and two examples of qualitative data analysis: narrative and rubric (textual) analysis. The quantitative data section gives a brief primer in statistics—both descriptive and inferential and also provides a thorough table defining a variety of statistical terms and listing helpful resources for more information. Both sections show the reciprocal nature of data analysis and the research question.

Chapter 8 "Presenting and Publishing Your Results" discusses the importance of making research public and publishable for reasons including expanding the field, achieving a promotion or tenure, and bringing SoTL work full circle. The chapter gives helpful advice about the kind of forum scholars should seek and the medium in which to publish, but also highlight the value of sharing work in different stages to different audiences and in varied forums.

Finally, Chapter 9 "Challenges of and Solutions for Doing Research on Teaching and Learning" focuses largely on the college instructors conducting SoTL research and whether this will help or hinder their careers. This shift in audience is inconsistent with the rest of the book; nevertheless, the considerations are valid to college-level instructors.

A few other highlights of *Engaging* are the detailed appendices listing resources, SoTL conferences and journals. There are several completed worksheets accompanying each chapter that demonstrate the narrated examples. Additionally, blank worksheets throughout the chapters are available for readers' own work. Many chapters include a section encouraging readers not to neglect the "student learning" aspect of SoTL and show how readers can incorporate students into their work. Also, several quotes from teachers engaged in SoTL in their own classrooms are embedded throughout the book. Though visually they break up the page, they are of little substance and a little self-congratulatory.

*Engaging* is a concise but complete introduction to the field of Scholarship of Teaching and Learning and to the process of formally conducting research to contribute to this field.

JOSOTL
JOURNAL OF THE SCHOLARSHIP OF TEACHING AND LEARNING

# Mission

Founded in 2001, the Journal of the Scholarship of Teaching and Learning (JoSoTL) is a forum for the dissemination of the Scholarship of Teaching and Learning in higher education for the community of teacher-scholars. Our peer reviewed Journal promotes SoTL investigations that are theory-based and supported by evidence. JoSoTL's objective is to publish articles that promote effective practices in teaching and learning and add to the knowledge base.

The themes of the Journal reflect the breadth of interest in the pedagogy forum. The themes of articles include:

1. Data-driven studies: formal research projects with appropriate statistical analysis, formal hypotheses and their testing, etc. These studies are either with a quantitative or qualitative emphasis and authors should indicate the appropriate domain. Acceptable articles establish a research rigor that leads to significant new understanding in pedagogy.

2. Reflective essays: integrative evaluations of other work, essays that challenge current practice and encourage experimentation, novel conclusions or perspectives derived from prior work

3. Reviews: Literature reviews illuminating new relationships and understanding, meta-analysis, analytical and integrated reviews, etc.

4. Case studies: These studies illustrate SOTL and its applications, usually generalizable to a wide and multidisciplinary audience.

5. Comments and communications: Primarily, these are comments based on previously published JoSOTL articles, but can also include book reviews, critiques and evaluations of other published results in new contexts or dimensions

**Submissions**

Authors are encouraged to submit work in one of the following categories:

- **Traditional Research Reports: data driven studies with either a quantitative or qualitative emphasis**
- **Reflective Essays on SoTL**
- **Reviews of current themes in SoTL research including meta-analysis**
- **Case studies illustrating SoTL and its applications**
- **Comments and Communications on previous Journal articles, or book or software reviews**

All submissions for JoSoTL should be submitted using the online submission process on our website beginning on July 1, 2012.

Please follow this link to get more detailed information on the submission process for the Journal of the Scholarship of Teaching and Learning:
http://josotl.indiana.edu/about/submissions#onlineSubmissions

If you have additional questions or run into problems please contact the journal at josotl@iu.edu.

## Editorial Board

All members of the **JoSoTL Editorial Board** are affiliated with FACET, the Faculty Colloquium on Excellence in Teaching, at Indiana University.

*Journal of the Scholarship of Teaching and Learning*, Vol. 12, No. 4, December 2012, pp. 144 – 146.

# Style Sheet for the *Journal of the Scholarship of Teaching and Learning*

**John Dewey**[1] **and Marie Curie**[2]

*Abstract: This paper provides the style sheet for the Journal of the Scholarship of Teaching and Learning. Manuscripts submitted for publication should adhere to these guidelines.*

*Keywords: radiation, metacognition, identity theory, constructivism, educational philosophy.*

## I. General Guidelines for the Manuscript.

The final manuscript should be prepared in 12-point, Times New Roman, and single-spaced. Submissions should be double-spaced. All margins should be 1 inch. The text should be fully left- and right-justified. The title (in 16 point bold) and author's name (in 12 pt. bold) should be at the top of the first page. The author's name should be followed by a footnote reference that provides the author's institutional affiliation and address. The abstract should be indented 0.5" left and right from the margins, and should be in italics.

Except the first paragraph in a section subsequent paragraphs should have a 0.5" first line indent. Use only one space after the period of a sentence (word processors automatically adjust for the additional character spacing between sentences). The keywords should be formatted identically to the abstract with one line space between the abstract and the keywords. Authors should use keywords that are helpful in the description of their articles. Common words found in the journal name or their title article are not helpful.

Pages should be unnumbered since they will be entered by the Journal editorial staff. We will also insert a header on the first page of the article, as above.

References should be incorporated in the text as authors name and date of publication (Coffin, 1993), with a reference section at the end of the manuscript (see below for the desired format for the references). Titles of articles should be included in the references in sentence case. Unless instructed otherwise in this Style Sheet, please use APA style formatting. Footnotes should incorporate material that is relevant, but not in the main text.

### *A. Plagiarism.*

It is essential that authors refrain from plagiarism. Plagiarism is a violation of ethics and, in serious cases, will lead to a manuscript being rejected by this journal. No future manuscripts will be accepted from authors who have submitted a plagiarized manuscript.

[1]Department of Educational Philosophy, Indiana University Northwest, 3400 Broadway, Gary, IN 46408, jdewey@iun.edu.
[2]Institut Pasteur, University of Paris, 75015 Paris, France.

*B. Unique work.*

This journal does not accept previously published work. We also do not accept work that is being considered for publication by another journal. If your manuscript is accepted, you will be required to sign a form stating that your manuscript has not been previously published.

**II. Section and Sub-Section Headings.**

*A. Major Sections.*

Major section headings should be flush-left, bold-faced, and Roman numeral numbered. Major section headings should have one-line space before and after. The first paragraph(s) of the article do not require a major heading.

*B. Sub-Sections.*

Sub-section headings should also be flush-left, in italics, and alphabetically numbered. Sub-section headings should have a one-line space before and after. Sub-sub-sections should appear at the beginning of a paragraph (i.e., with an 0.5" indent, followed immediately by the text of the sub-sub-section), with the heading also in italics.

**III. Tables and Figures.**

Tables and figures should be inserted in the text where the author believes they best fit. They may be moved around a little to better correspond to the space requirements of the Journal. If necessary, tables and figures may occupy an entire page to ensure readability and may be in either portrait or landscape orientation. Insofar as possible, tables should fit onto a single page. All tables and figures should be germane to the paper. Tables should be labeled as follows with the title at the beginning (in bold), with data entries single-spaced, and numbered. Column labels should be half-line spacing above data.

**Table 1. The title of the table.**

| Unit | Length, inches |
|---|---|
| Point | 1/12 |
| Pica | 1/6 |

Figures should have their captions follow the image. Captions should be single-spaced, with title in bold. Additional text should not be in bold. The Editorial staff may adjust layout to allow optimal use of space.

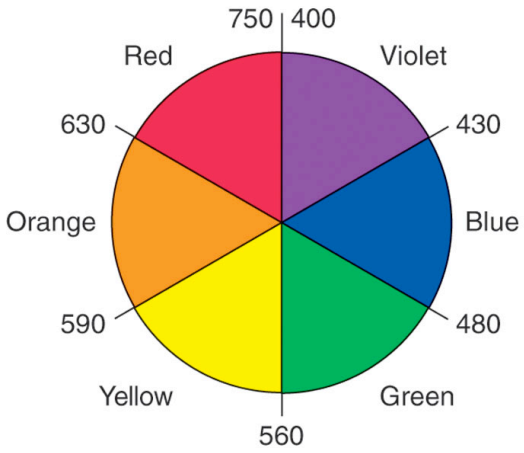


**Figure 1. Color wheel with wavelengths indicated in millimicrons.** Opposite colors are complementary.

## Acknowledgements

Acknowledgements should identify grants or other financial support for this research by agency (source) and number (if appropriate). You may also acknowledge colleagues that have played a significant role in this research.

## Appendix

Please insert any appendices after the acknowledgments. They should be labeled as follows:

**Appendix 1. The Title of the Appendix.**

## References

Coffin, D.A. (1993). Using the competitive edge. *Journal of Economic Education*, *24* (1), 62-69.

Garcia, J. and Rodriguez, P. (2002). The determinants of football match attendance revisited: Empirical evidence from the Spanish football league. *Journal of Sports Economics*, *3* (1), 18-38.

Hamilton, S. J. (1995). My name's not Susie: A life transformed by literacy. Portsmouth, NH: Boynton/Cook Publishers.

Pappas, D. (2004). Fixing the fan cost index: A more realistic view. Retrieved April 21, 2004, from http://www.baseballprospectus.com/article.php?articleid=2790.

## Contact Info for the Journal

JoSoTL Editorial Office

Indiana University Purdue University Indianapolis
755 W. Michigan St, UL 1180D
Indianapolis, IN 46202

josotl@iupui.edu

ISSN: 1527-9316