JoSoTL
Journal of the Scholarship of Teaching and Learning

Volume 15, Number 1, December 2015 *www.iupui.edu/~josotl*

JoSoTL is a collaboration between
The Mack Center for Inquiry on
Teaching and Learning and the
Faculty Colloquium on Excellence
in Teaching, Indiana University

JOSOTL
JOURNAL OF THE SCHOLARSHIP OF TEACHING AND LEARNING

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 1 - 9.
doi: 10.14434/josotl.v15i1.13089

Learning to see the infinite: Measuring visual literacy skills in a 1st-year seminar course

Michael S. Palmer[1] and Tatiana Matthews[2]

Abstract: Visual literacy was a stated learning objective for the fall 2009 iteration of a first-year seminar course. To help students develop visual literacy skills, they received formal instruction throughout the semester and completed a series of carefully designed learning activities. The effects of these interventions were measured using a one-group pretest-posttest study design where students were asked to look at two different—but stylistically similar—paintings and write a response to the following two questions: what do you see and what do you think it means? Students' responses were analyzed based on the visual evidence they recorded and the strength of their arguments using Toulmin's argument model. After instructional interventions, paired t-tests indicate that students made significantly more basic ($t(15) = 4.291$, $p < .001$) and advanced observations ($t(15) = 2.440$, $p = .014$), offered more supporting visual evidence for their best-supported claims, and made stronger connections between their claims and the visual evidence.

Keywords: visual literacy, first-year seminar, Toulmin's argument model

Background

Visual literacy was first coined in 1969 by John Debes (Debes, 1969). An exact definition remains elusive because of the complexity and multidimensional nature of the skills involved, but at its core visual literacy refers to the ability to interpret, negotiate, and make meaning from information presented in an image (Avgerinou & Ericson, 1997). Drawing on the science of learning, James E. Zull argues in the *The Art of Changing the Brain* (2002, p. 146) that faculty should make "extensive use of images to help [students] learn," both by teaching with visuals and by requiring them to use visual forms to represent what they know. Little, Felten, and Barry (2010, p. 46) have argued that "visual literacy…is a critical skill for twenty-first-century students and ought to be a central component of liberal education." They go on to suggest that instructors teaching courses in academic disciplines outside those commonly affiliated with visual literacy (for example, art history and media studies) and those teaching first-year general education through capstone courses should explicitly help students develop visual literacy skills. This is not only because images saturate our daily lives but also because many fields, including those in the arts, sciences, social sciences, and health sciences, rely heavily on images as data sources and require a heightened level of visual literacy to understand and interpret the world (Brumberger, 2011).

Though an ever-expanding literature exists on teaching visual literacy skills in higher education (Felten, 2008), only a handful of studies measuring students' general visual literacy

[1] University of Virginia, mpalmer@virginia.edu
[2] University of Virginia, tem6c@virginia.edu

skills have been published (see, for example, Arslan & Nalinci, 2014; Brumberger, 2011). Even fewer studies exploring the impact formal instruction has on students' visual literacy skills have been reported (Hollman, 2014; Linenberger, & Holme, 2014; Yeh, 2010). The current study, which was embedded in my highly interdisciplinary, discussion-based, first-year seminar course known as *Falling from Infinity*,[3] helps to fill this void. Specifically, we examined whether after a series of instructional interventions focused on a subset of visual literacy skills my students were better able to 1) make more basic and advanced observations in an image, 2) offer more supporting visual evidence for a claim they made about that image, 3) and make stronger connections between their claim and the visual evidence.

Falling from Infinity drew on a diverse set of perspectives—literary, poetic, artistic, mathematical, scientific, religious, philosophical—and invited my students to spend a semester grappling with uncountable numbers, immeasurable spaces, and unending times. The diverse nature of the topic lends itself well to the use of visuals as an important and essential vehicle for exploring course content. Because researchers (Wineburg, 1999) have found that students often need significant structured guidance to form meaningful metacognitive routines when exploring and making meaning from images (for example, asking themselves "what else do I need to do to better understand the image"), visual literacy was as one of the explicit learning objectives of the course. I implemented a series of learning activities designed to help students practice foundational visual literacy skills (e.g. making observation) early in the semester to support more complex ones later on (e.g. critically analyzing images). Many of the visual literacy learning activities were developed in class through small group, collaborative exercises or large group discussions. Others were assigned as part of reflective journal assignments. For the latter, students completed a short task related to course content outside of class and then wrote a 1-2 page reflection on the experience. For example, students were asked to look at images outside of class and briefly reflect on how this new information complicated their understanding of the infinite. Other activities supported more complex assignments, such as creating a photo essay and writing a 3-4 page reflection evaluating the ways the photo was effective in representing the infinite and in what ways it fell short.

A thorough discussion of all the specific teaching interventions used throughout the course is beyond the scope of this work, but we highlight here two of the more substantial activities. Curious readers can explore the complete week-by-week timeline of the course elsewhere (Palmer, in-press). The first illustrative example occurred near the mid-point of the semester, after much of the visual literacy foundation had been laid. I began class by formally defining certain terms, like observation (a neutral, non-judgmental and verifiable statement) and inference (a meaning, an interpretation, or an assumption based on observation) and then systematically engaging my students in critical analysis of the images we were encountering. To start our in-class discussion, I displayed only a portion of Joel Sternfeld's photograph *Warren Avenue at 23rd Street, Detroit, Michigan*. I choose this particular photo, one that has nothing to do with infinity, to avoid having students overlay their prejudices and biases about infinity. They were first prompted to make as many observations about the visible part of the photograph as possible, preceding each observation with the statement "I see…" Next, I revealed another portion of the photograph and asked them to add to their list of observations. I continued to reveal the photograph slowly, piece-by-piece until the image was complete. Once the students were satisfied that they made all relevant observations, I prompted them to make inferences,

[3] The study was conducted during the 2009 fall semester. For details about the course, visit http://faculty.virginia.edu/infinity/.

beginning each with, "I infer…because…" In follow-up questions, I prompted them to consider the strength of their inferences, asking "How well supported is your inference based on the observations we just made?" After this exercise, students examined van Gogh's Starry Night, and again individually wrote down all their observations, compared notes with their neighbor, created a combined class list, and then made the best supported inferences possible.

In another example, students looked at Caspar David Friedrich's painting Monk by the Sea before class. In class, I asked them to respond to the question, "What do you see?" I then asked them to individually write down 3-5 adjectives that described their emotional response to the painting. Afterward, they shared these in the large group as I wrote and categorized their adjectives on the board. This particular painting is unique in that it typically evokes polar opposite responses: some viewers feel it evokes "awe" and a feeling of "tranquility" while other feel it evokes "dread" and "hopelessness." During the conversation, I asked the students, "What in the painting leads you to feel…?" After all responses were collected and reported out on the board, we discussed the question, "Which set of responses is correct?" What my students discovered through this activity is that some claims, when based solely on the observational data alone, are more easily supported than others.

Methodology

To help students develop visual literacy skills, they received formal instruction throughout the semester and completed a series of carefully designed learning activities like those described above. The effects of these interventions were measured using an IRB-approved, non-graded pre-/post-semester methodology where students (n=16) were asked to look at two different—but stylistically similar—paintings and write a response to the following two questions: what do you see and what do you think it means? The paintings were Salvador Dali's *The Persistence of Memory* and *The Disintegration of the Persistence of Memory*. Our hypothesis was that following the interventions students would:

- make more basic and advanced observations in an image;
- offer more supporting visual evidence for a claim they made about that image;
- make stronger connections between their claims and the visual evidence.

Toulmin's argument model (Toulmin, 1969), with particular focus on claim, supporting evidence, and warrant (the inferences or assumptions taken for granted by the writer that connect the claim and the supporting evidence), was used to analyze students' responses. This model was chosen based on its relative simplicity and because it mirrors the pedagogical strategy I used during classroom activities; that is, I routinely asked students to make observations and then support their claims based on those observations. Students were not introduced to or given formal instruction in Toulmin's argument model.

When analyzing students' responses, we defined supporting evidence as the observations of visual information they included and coded these observations as either basic or advanced. A basic observation described an object or feature of the painting without significant qualifiers, e.g. "I see a clock." An advanced observation described an object or feature of the painting beyond merely identifying it, such as the position of the object/feature relative to others or its location relative to the painting, e.g. foreground/background; the texture of objects or the texture of the painting itself; the contrast or juxtaposition of objects or features; the source and/or direction of light; a minute, easily overlooked detail of the piece; or, an observation the viewer made about

his or her own experience viewing the painting, e.g. "I looked at the painting in a counterclockwise manner."

Students' responses to the pre- and post-assessments were initially analyzed independently by two raters. Observations were coded as either basic or advanced. The raters' results were then compared to make certain every observation was identified and coded consistently. Observations were counted regardless of whether they were found in the response to the question "what do you see" or "what do you think it means" since students often included additional observations when supporting their claim.

A similar methodology was used to analyze the claim, supporting evidence, and strength of the warrant. After identifying the main claim, raters coded the supporting visual evidence, rated the warrant as strong, moderate, weak, or none, and then again compared results. When differences arose, raters discussed the student's response and came to consensus on the appropriate coding. A strong warrant indicated that links connecting the claim and the evidence were clear, all inferences were appropriate, and the assumptions reasonable. A moderate warrant indicated that *most* of the links connecting the claim and the evidence were clear, *most* inferences were appropriate, and *most* assumptions reasonable. A weak warrant indicated that there were no clear links connecting the claim and the evidence, inappropriate inferences, or faulty assumptions. If no claim was present, the warrant was coded as "none." If a student included more than one claim, only the best-supported claim was analyzed.

Findings

Student's responses to the pre- and post-semester assessments were coded and analyzed for the number of basic and advanced observations they made, the number of pieces of visual evidence supporting their main claim, and the strength of their warrant.

Basic Observations

The number of basic observations made per student in the pre- and post-semester assessments is shown in figure 1. At the start of the semester, before the instructional interventions, the median value for the number of basic observations made was 13 (range = 0-33), with three students making less than 10 basic observations. Near the end of the semester, following the instructional interventions, the median value for the number of basic observations made was 26 (range = 13-68). Two students made approximately the same number of basic observations in the pre- and post-semester assessments, but all other students made more in the post. A paired t-test found strong evidence that the interventions significantly improved the post basic observation score: $t(15) = 4.291$, $p < .001$.

Advanced Observations

The number of advanced observations made per student in the pre- and post-semester assessments is shown in figure 2. At the start of the semester, the median value for the number of advanced observations made was 2 (range = 0-7), with 11 of the 16 students making 2 or fewer advanced observations. Near the end of the semester, the median value was 3 (range = 2-14), and nine of the 16 students made more than 3 advanced observations. Two students made approximately the same number of advanced observations in the pre- and post- assessments, but

all other students made more in the post. A paired t-test found strong evidence that the interventions significantly improved the post advanced observation score: $t(15) = 2.440$, $p = .014$).

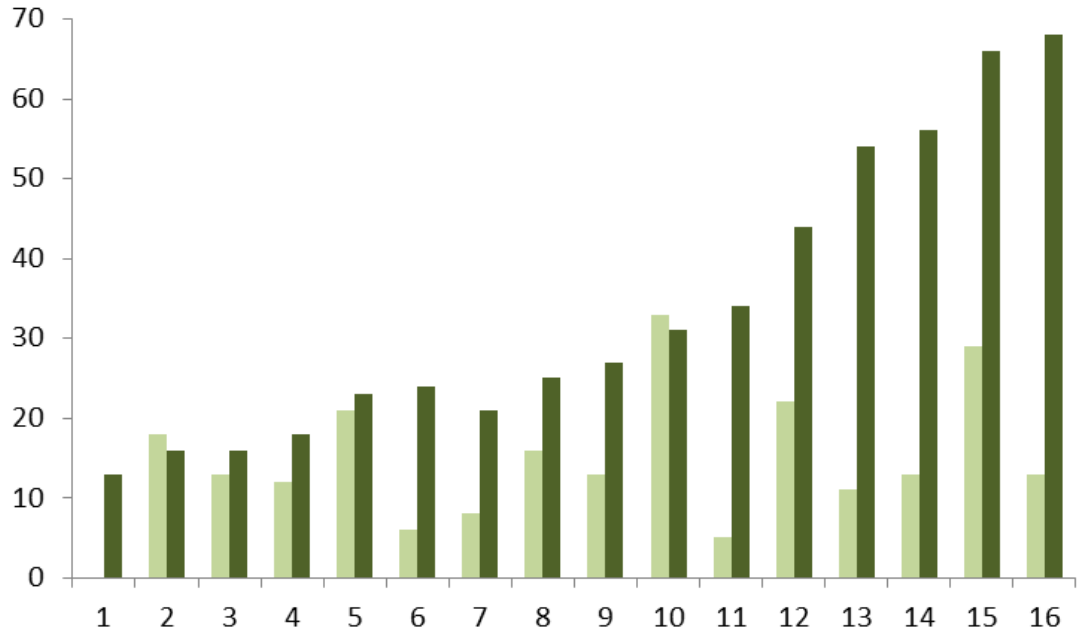


Figure 1. Basic observations made per student (n=16) in pre- (light green) and post-semester (dark green) assessments.

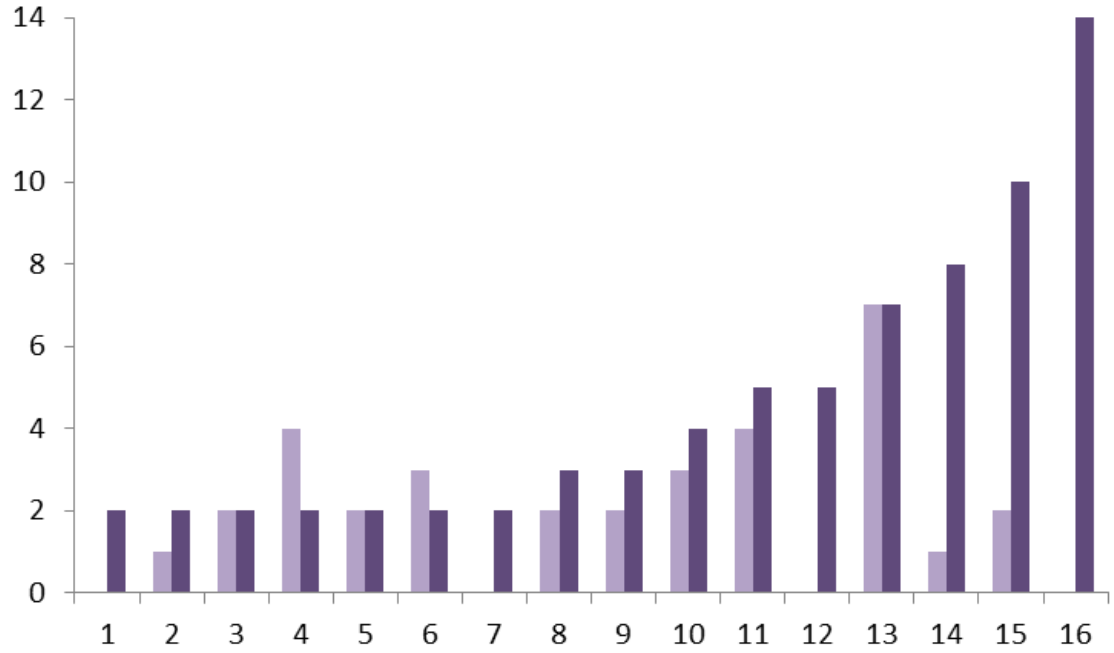


Figure 2. Advanced observations made per student (n=16) in pre- (light green) and post-semester (dark green) assessments.

Evidence Supporting Main Claim

The number of pieces of evidence that students provided for their best-supported claim is shown in figure 3. For the pre-assessment, students provided on median 2.5 pieces of evidence (range = 0-6). Three students had no visual evidence to support their claim; one student made no substantive claim. For the post-assessment, students provided on median 8.0 pieces of visual evidence for their best-supported claim (range = 3-40) and every student provided at least three pieces of supporting evidence.

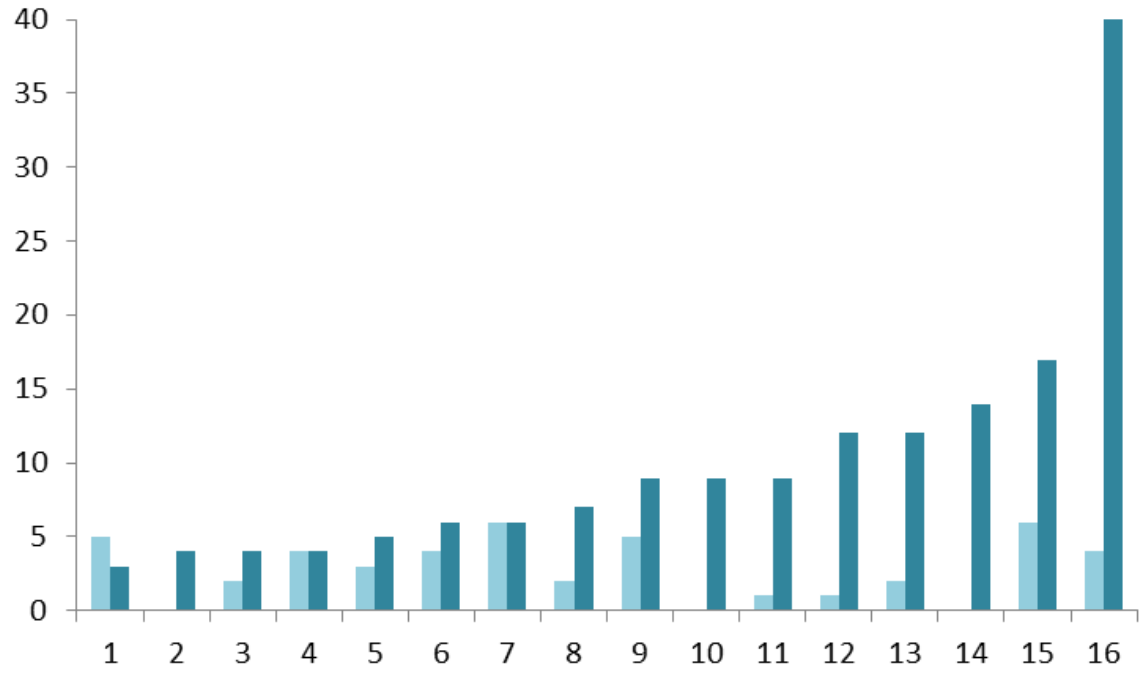


Figure 3. Number of total observations per student (n=16) supporting the primary claim of the pre- (light teal) and post-semester (dark teal) assessments.

Strength of Warrant

The strength of the students' warrant in the pre- and post-assessments is shown in figure 4. At the start of the semester, three students provided no warrant, six had a weak one, five a moderate one, and only two had a strong warrant. Near the end of the semester, all students provided a warrant: one weak, nine moderate, and six strong.

Conclusion

The results of this study suggest that the classroom interventions I incorporated into the first-year seminar course, *Falling from Infinity*, significantly improved students' ability to make necessary and appropriate observations in images and to develop stronger claims supported by their observations. Because I embedded numerous learning activities focused on improving visual literacy skills throughout the course, it is difficult to pinpoint which activities had the

greatest impact on my students' learning. A portion of every class period, whether explicitly or implicitly, was devoted to some aspect of visual literacy. Though we believe the two activities described in the background section of this paper were key interventions, most likely, it was the sum of activities that led to the positive gains rather than any one particular intervention. The fact that advanced observations weren't defined or categorized until late in the semester and that students exhibited lower gains in this area supports the idea that the interventions had a cumulative effect. Future studies could probes this hypothesis and also help identify the specific activities that have the greatest impact on improving the visual literacy skills explored herein. Such studies could also shed light on whether formal instruction in Toulmin's argument model might in and of itself lead to similar or possibly improved outcomes.

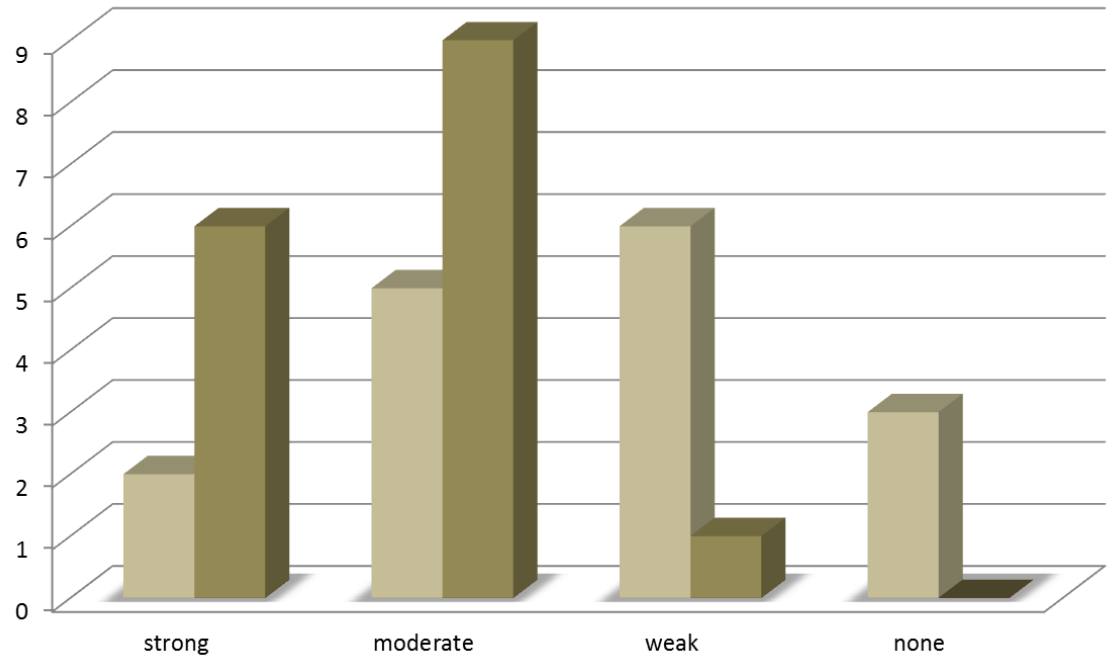


Figure 4. Strength of the warrant for students' primary claim for the pre- (light gold) and post- (dark gold) semester assessments. None indicates that no claim was present.

As a final comment, it is worth noting that my students' perceptions were consistent with the positive findings of this study. On the end-of-course evaluation statement, "Because of this course, I am better able to critically analyze visual images," students responded on average 4.31 out of 5.00 (5 = strongly agree). Select qualitative comments included:

> *The most important skill I learned [in this course] was analyzing works of art (painting/photography).*
>
> *...every time I look at a painting now, I'll think: 1) Observation! 2) Inference.*

Acknowledgements

I kindly acknowledge Melissa Hurst for help with the statistical analysis of the pre-/post-semester assessment data. I am forever grateful to Deandra Little for introducing me to the world of visual literacy, for mentoring me throughout the project, and for her uncommon colleagality.

References

Arslan, Rumiye, & Nalinci, G. Z. (2014). Development of visual literacy levels scale in higher education. *The Turkish Online Journal of Educational Technology*, *13*(2), 61-70.

Avgerinou, M., & Ericson, J. (1997). A review of the concept of visual literacy. *British Journal of Education Technology*, *28*(4), 280-291. doi: 10.1111/1467-8535.00035.

Brumberger, (2011). Visual literacy and the digital native: An examination of the millennial learner. *Journal of Visual Literacy, 30*(1), 19-46.

Debes, J. (1969). The loom of visual literacy: An overview. *Audiovisual Instruction*, *14*(8), 25-27.

Felten, P. (2008). Visual literacy. *Change: The Magazine of Higher Learning*, 40(6), 60-64. doi: 10.3200/CHNG.40.6.60-64

Hollman, V. (2014). Promoting visual literacy among undergraduate students in geography: teaching a visualized Latin America. *Journal of Geography in Higher Education, 38*(1), 136-147. doi: 10.1080/03098265.2013.836626.

Linenberger, K. J., & Holme, T. A. (2014). Biochemistry instructors' views toward developing and assessing visual literacy in their courses. *Journal of Chemical Education*. Advance online publication. doi: 10.1021/ed500420r

Little, D. Felten, P., & Berry, C. (2010). Liberal education in a visual world. *Liberal Education*, *96*(2), 44-49.

Palmer, M. S. (in-press). Learning to see the infinite: Teaching visual literacy in a 1st-year seminar course. *New Directions for Teaching and Learning, 141*, 19-29.

Toulmin, S. (1969). *The uses of argument*, Cambridge, England: Cambridge University Press.

Wineburg, S. (1999). Historical thinking and other unnatural acts. *Phi Delta Kappan, 80*(7), 488-500.

Yeh, H. (2010). Towards evidence of visual literacy: assessing pre-service teachers' perceptions of instructional visuals. *Journal of Visual Literacy, 29*(2), 183-197.

Zull, J. E. (2002). *The art of changing the brain: Enriching the practice of teaching by exploring the biology of learning*. Sterling, VA: Stylus.

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 10 - 19.
doi: 10.14434/josotl.v15i1.12905

Collaborative learning supported by rubrics improves critical thinking

Carlos Saiz[1], Silvia F. Rivas and Sonia Olivares

Abstract: In previous works we developed and assessed a teaching program, ARDESOS v.1, with which we aimed to improve the fundamental skills of critical thinking. The results obtained were positive, but modest. After analyzing the limitations of the program we introduced certain modifications and assessed the new version. The changes involved designing the activities programmed by means of rubrics and making the students perform them with less direct orientation from the instructor. In sum specificity and initiative proved to be the key variables in the improved program, ARDESOS v.2. Based on the data collected we have seen a significant improvement of the new version over the old one in the following aspects: a) version 2 improved all the fundamental dimensions, mainly in the pre- and post-test measurements, to a significant extent (Student's t test); b) the effect size (Cohen's d) was significantly higher, and finally c) these improvements in the program elicited better performance. Accordingly, an improvement in critical thinking can be achieved via an instruction design that addresses the factors that really induce change. Currently, with these results we have been successful in adding a new improvement to the instruction, which we have re-evaluated.

Keywords: Critical thinking, Instruction, Evaluation

Introduction

In two previous works (Olivares, Saiz & Rivas, 2013; Saiz & Rivas, 2011) we developed and assessed a program for the instruction of critical thinking (ARDESOS, first version- v.1). The successful functioning of this teaching methodology prompted us to develop a second version of the program (v.2), and also to improve the efficiency of the former version. In the two previous studies, the data obtained were reasonably satisfactory since they reflected important changes in many of the basic skills of critical thinking. This stimulated us to continue working on this ambitious teaching project. The changes observed were also challenging because there were some aspects of the program that did not lead to the expected changes. This is of course quite usual in any line of research: the presence of clearer and more shadowy areas, which should be strengthened and eliminated respectively.

Improving Critical Thinking has and continues to be the underpinning of our research efforts. In our earlier work, we followed several principles and used teaching resources that we have maintained in the present project, although complemented by others. In the first version of the program we used a) the importance of team work, b) direct teaching, c) the need to learn from deficiencies or limitations, and d) the advantages of learning based on problems that arise in people's everyday lives.

Currently, the teaching system has evolved with respect to the first version. A scheme could serve to clarify this. Figure one summarizes the essential features of the ARDESOS v.2

[1] Universidad de Salamanca (Spain), csaiz@usal.es. Web: www.pensamiento-critico.com

program as used in the present work. In this scheme we have integrated the working methods, tasks, materials and motivational factors. However, to all this we should add, and emphasize, the fact that the participants in the program must decide whether they wish to enroll or not. The students have two options: our instruction program or another conventional teaching program and they must decide which to choose. This choice is more important than it may appear to be. In our program the learning process is based on ideas developed in previous contributions. For the present study, it is appropriate to underscore learning from limitations and problem-based learning (PBL) as the main motor driving the change or improvement in critical thinking. Figure one contains some ideas that are in bold and others that are not. The words in bold differentiate our program from others. They are procedures that have not been implemented or have been used only sporadically. For example, unlike the generalized use of comprehension tasks it is very uncommon to use production tasks in teaching. It is common to use one or another task separately, but not together and the same importance must be given to both. This is one of the original characteristics of our program, at least as far as we know. Moreover, the instruction system based on deficiencies or limitations is certainly one of the most singular aspects of our methodology. Regarding the materials used, there are no studies that have used daily or professional problems, videos, opinion-oriented articles, working the fundamental skills of critical thinking in an integrated way in each of them. As indicated in figure one, these aspects affect and foster the essential motivational traits such as interest, utility, achievement and effort.

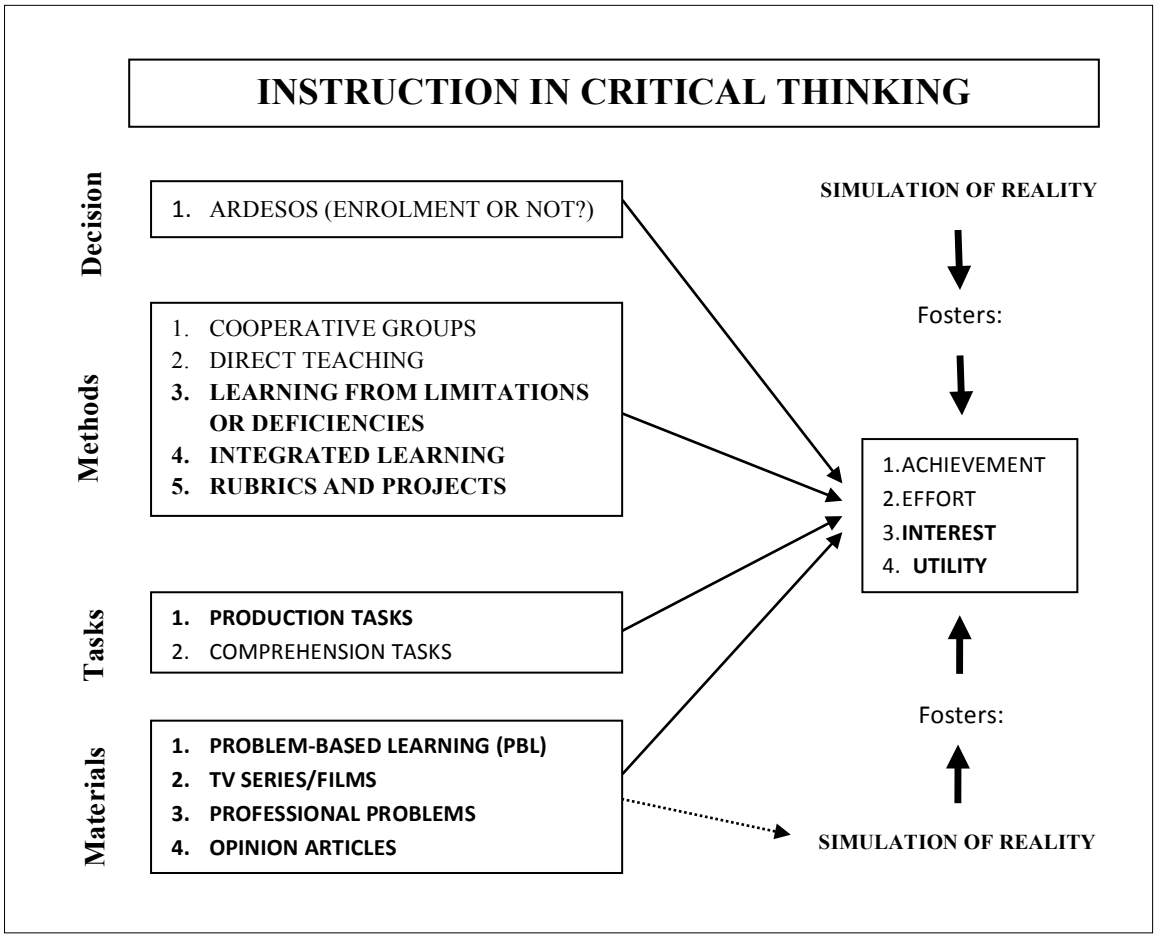


Figure 1. Main characteristics of the ARDESOS v.2. Program.

In the first version of our program, and to a certain extent in the second, the main effort was directed towards achieving efficiency in teaching. It was therefore also mainly directed at achieving an improvement in critical thinking, using strategies, tasks and materials that would guarantee a good result. This global effort to construct a system that would work was successful. However, we did not know which aspects of the program or which factors or variables were relevant or more relevant than the others. The crux of the matter in the present work is to determine whether there are aspects of the instruction that are more determinant than others. We believe that this is indeed the case. Here we isolated the factors involved in teaching: which of them really makes it work. Also, we wished to know whether it is the overall intervention that fosters the changes in the critical thinking skills. Apparently, this problem in education has not yet been addressed in the literature, but should have been tackled a long time ago. Knowing whether there are relevant factors in instruction is of great importance, both theoretically and practically. Furthermore, in our case, after many years of experience we have observed, but not confirmed, that there are some aspects of teaching that have a greater influence than others on the learning process. One of them has to do with the generalized assertion within the field of education that learning often depends more on what the student does than what the instructor does (Almeida, 2013). The *active participation*, in contrast to the passivity, of students seems to be an especially important factor in education. Nevertheless, there are no studies that have endeavored to check this. Here, we attempt to fill this gap; greater involvement or participation in the learning process must be guided or oriented. Accordingly, active participation by students must be accompanied by *specific instructions*. In the current version of our program the two main changes made are: greater activity or participation in tasks by students, and specificity in the performance of such tasks. How did we operationalize these variables? First by ensuring that the instructor would dedicate more time in directing and orienting the students' work and less time in solving the problems posed. Second, through the elaboration of specific rubrics for each of the tasks or problems posed. This teaching resource made students address the problems by following the indications specified in the method. Accordingly, their activity in the classroom would be focused and well oriented. They knew which aspects were to be worked, the relevance of each, the points they would earn, and the strategies required to apply them.

Thus, the two chief goals in the present work are: a) to determine whether greater activity or participation by the group in resolving the problems posed improves their thinking competencies, and b) whether a guide in the form of rubrics for performing the tasks also contributed to the improvement. Let us illustrate these factors with one of the tasks. One of the activities designed in our program addresses the development of competencies in argumentation. Chart one show one of the rubrics used. It may be seen that the parts to be taken into account and the aspects to be considered in any type of argumentation are detailed and assessed. The method employing rubrics is one of the most efficient ways of quantifying qualitative tasks and guiding learning in a highly concrete and specific manner.

Now, how can we test whether these factors produce change? We tested this by comparing the effect size in the test on critical thinking. The way chosen by us was to compare the effect size in the test on Critical Thinking PENCRISAL (Rivas & Saiz, 2012; Saiz & Rivas, 2008) with the assessment of the ARDESOS v.1 program and the current version. If, as we assumed, the factors introduced in our instruction program determined the improvement in the learning procedure, we expected that the effect size would be significantly greater in the current version than in the previous one. We also expected there would be significant improvements in some of the dimensions of critical thinking that we did not manage to achieve with the first

version. Finally, on comparing both programs we expected that performance would be significantly better in version 2 of the Ardesus program. All these changes are addressed in the section on methods.

CRITERIA	SCORE				TOTAL
Comprehension	+5	+5	+5		+15
	Precision in the drafting of ideas	Identification of what is fundamental	Relevant observations		
Argumentation					
	10	45	+10	+5	55/+15
Structure	Conclusion	5 mainreasons/ counterarguments	Another 3 reasons/ counterarguments	Restrictions or conditions	
	+5	+5	25	+5	25/+15
	Opinions, assumptions, conjectures,...	Facts	Relations	Other Considerations	
	5	10	5	+5	20/+5
Assessment	Acceptability	Relevance	Global	Falacies	
MAXIMUM TOTAL SCORE	15/+10	55/+10	30/+15	0/+20	100

Chart 1. Rubric Arg.1 Group Comprehension Task Argumentation.

Methodology

Participants

The sample of the present study comprised 144 students from the first year of the Degree in Psychology of the University of Salamanca. Of these, 82.6% (119) were women as compared with 25 men (17.4%). This difference is statistically significant (Chi^2= 16.531; 1 gl; p=.00). The mean age of the sample was 18.83 (s.d. 1.89) (CIat 95%: 18.51-19.14), within a range of 18-32 years. The distribution did not fit the normal model with p<.01 on the Kolmogorov-Smirnov goodness of fit test (p=.00) owing to a marked positive asymmetry (As= 4.00) and a clearly leptokurtic shape (K=20.40). The study sample of version v.1 is described in the paper by Saiz and Rivas (2011).

Instruments

The PENCRISAL Critical Thinking test.

This test comprises 35 situation-production problems, with an open format and is structured around 5 factors: Practical Reasoning, Deduction, Induction, Decision Making and Problem Solving (Cronbach alpha = .632; test-retest = r =.786, Rivas & Saiz, 2012). Each of the factors contains the most representative structures, thus enabling us to isolate the main skills of Critical Thinking and the most relevant methods of reflection and resolution of our daily lives. The PENCRISAL test has been described in detail in Saiz and Rivas (2008). This test was designed following the methodology of task analysis in order to uncover which processes or mechanisms of thinking are functioning on each of the 35 problems posed in the test. The

problems were designed in such a way that it was only posible to solve them by using a strategy or a mechanism. Thus, we know that on solving a problem, an item of causality, this can only be done using causal reasoning and not other mechanism. In other words, that if a problem needs to be solved using an identification strategy it cannot be done in any other way. What is more important, we can identify the mechanism in the open answers given on the test. For further information, the links to those works in English can be consulted:
Saiz and Rivas (2008):
http://www.pensamiento-critico.com/archivos/evaluationCTergoENGLSH.pdf
Rivas and Saiz (2012):
http://www.pensamiento-critico.com/archivos/validacionpencrieng.pdf

ARDESOS v.2 Program

As reported above, in comparison with the first version the instruction was improved. The duration of the program was 60 hours (face-to-face teaching) along 15 weeks and four hours of class per week. The instruction was given in classes of 30-38 students divided into four groups so that the students could work in teams. All activities were planned at the beginning of the course, with rubrics. The classroom work was directed towards the development of these activities, under the supervision of an instructor. The role of this latter consisted of orienting the students in each of the tasks and clarifying any doubts that might arise during their completion. Later, in the assessment of the activities the solution to each activity was explained to the students.

Assessment was performed on a weekly basis, with feedback facilitated 2 to 3 days later. The importance of the immediacy of the assessment should be noted in the sense of that it fosters a good development of learning. The assessment was quantitative, as specified in each rubric. Thus, students knew how much weight each part of the task carried and what was more or less important. For example, in an activity involving argumentation what was most important were the identification and relationships of the elements of an argument, while its evaluation was less important. It is important to recall that evaluation is an essential component of our program; the learning process would be impossible without it.

Procedure

The ARDESOS v.2 program was applied along one term at the School of Psychology of the University of Salamanca. One week before the start of instruction all students took the PENCRISAL test. Likewise, it was applied one week after the intervention to obtain a second measurement of the variables. The time elapsed between the pre-treatment and the post-treatment measurements was four months. The first version of the ARDESOS program was implemented using a procedure identical to that used in the application of the current one.

Design

In order to analyze the efficiency of the intervention we used a quasi-experimental design, with pre- and post-treatment measurements.

Statistical analysis

In the statistical analysis we employed the IBM SPSS Statistics 19 package. The tools and statistical techniques used were as follows: frequency tables and percentages for the qualitative variables, with a Chi-square test for homogeneity; exploratory and descriptive analyses of the quantitative variables with a test for goodness of fit to the normal Gaussian model and box diagrams for the detection of atypical values (outliers); statistical techniques (mean, standard deviation, median… etc.) for numerical variables; the *t* test for the value of a measurement, tests of the significance of differences of Student's *t* means, and calculation of Cohen's *d* to estimate effect size.

Results

Regarding the descriptive statistics of all the variables included in the study, we observed that most of them fit the model of normality adequately, although some had significant deviations, which were overlooked due to the size of the sample.

Table 1

Descriptive statistics of the variables

	N	M	SD	Minimum	Maximum
TOT_PencriPre	144	28.58	6.53	12	45
DR_PencriPre	144	3.98	2.00	0	9
IR_PencriPre	144	5.06	1.81	1	16
PR_PencriPre	144	6.31	2.53	0	12
DM_PencriPre	144	6.69	1.94	1	11
PS_PencriPre	144	6.53	2.19	1	10
TOT_PencriPost	144	31.70	6.49	14	44
DR_PencriPost	144	5.25	2.17	0	11
IR_PencriPost	144	5.48	1.67	2	9
PR_PencriPost	144	8.40	2.32	1	13
DM_PencriPost	144	7.01	2.08	2	13
PS_PencriPost	144	5.56	2.49	0	11

Below, the results of the statistical analyses performed are shown in order as a function of the above aims.

In order to assess the differential effect caused by the program over the two years and to determine in which factors the improvements introduced were affected the most, we performed tests on the significance of differences of Student's *t* means and calculated Cohen's *d* values to estimate the effect size.

As can be seen in table two, the results provided by the descriptive statistics indicate that the optimized v.2 ARDESOS program was more effective since it significantly improved the performance on the post- measurements across the whole scale and in all the factors, with the exception of decision making, whereas with v.1 a significant increase occurred only in the post-performance of the induction and decision-making factors.

With a view to analyzing the impact of the intervention of the two versions, we used the standard mean difference, *d*, of Cohen (1988) as an index of effect size. The data show that in v.2 of the program a significant increase occurred in the deduction, practical reasoning and problem-solving factors and in the overall score of the scale. It may be seen that regarding practical reasoning (*d*=.83) and deduction (*d*=.63) effect size has very high values. However, in v.1 these values are lower (Pract. Reasoning: *d*=.03; deduction, *d*=.45). Likewise, the total of the scale (*d*=.48) and the problem-solving factor (*d*=.44) had a moderate effect size in v.2 whereas in v.1 these values were very low, ranging around .10. In light of these results, it may be concluded that the improvements introduced are reflected in an increase in critical thinking skills and the skills with the greatest effect size are practical reasoning and deduction, followed by problem-solving and, to a lesser extent, induction skills.

Table 2

Differences in Student's t means and effect size- Cohen's d

	ARDESOS PROGRAM VERSION 1					**ARDESOS PROGRAM VERSION 2**				
	PRE	**POST**	**Student's *t* test**			**PRE**	**POST**	**Comparison**		
	M (SD)	**M (SD)**	**Diff. in means p-sig n**	**t**	**Effect size**	**M (SD)**	**M (SD)**	**Diff. in means p-sig n**	**t**	**Effect size**
DED	6.31 (2.47)	5.21 (2.21)	1.10** .000 97	3.83	.45	3.98 (2.01)	5.25 (2.17)	-1.27** .000 144	-6.57	.63
IND	3.74 (1.59)	4.69 (2.20)	-.95** .000 99	3.84	.60	5.06 (1.81)	5.48 (1.67)	-.41** .006 144	-2.51	.23
PR	6.37 (2.69)	6.47 (2.74)	-.10 .741 97	.33	.03	6.31 (2.53)	8.40 (2.32)	-2.09** .000 144	-9.08	.83
DM	6.08 (1.74)	6.64 (2.04)	-.56* .040 88	2.08	.32	6.60 (1.94)	7.01 (2.08)	-.31 .063 144	-1.53	.16
PS	3.75 (1.32)	3.53 (1.29)	.22 .135 94	1.51	.17	5.56 (2.19)	6.53 (2.49)	-.97** .000 144	4.72	.44
TOT	25.98 (6.27)	26.65 (7.35)	-.67 .448 88	.76	.10	28.58 (6.53)	31.79 (6.49)	-3.12** .000 144	-5. 87	.48

* Significant at 5% ** Significant at 1%

Since we were interested in checking whether the improvements might indicate better performance we decided to use the *t* test to see whether the values of the means were statistically

significant (see table 3). We then compared the means of the improved version with the average mean obtained in the v.1 sample. The difference between the means of the improved version and v.1 proved to be statistically significant at $p<.01$ on the whole scale and on all the subfactors, except decision making. This allowed us to conclude that the sample analyzed with the improved version of the program afforded a significantly better performance than the sample of v.1.

We observed that the sample analyzed had a significantly improved performance on all the skills of critical thinking than (as compared with) the sample from v.1, with a difference of 5.5. points (CI 95%: 3.98-6.12). Regarding the critical thinking skills variables, we noted that all of them but one underwent a statistically significant increase. Deduction rose from a mean of 5.21 in the first version to 5.25 in the second one (CI 95%: .31-.40). Although the means are fairly similar, it should be noted that in v.1 there was a problem in the pre- measurement because the instruction had already been followed, such that –as seen in Table 2- it was higher than the post- value. The result of the second version is therefore important since the increase from the pre- mean to the post- mean was more than almost a whole point, accounting for .63 of the effect size. Induction was affected to the same extent, with a significant increase in its mean of almost a whole point (CI 95%: .51-1.06). Practical reasoning also showed higher performance means in the second version, where an increase of almost two points was observed (CI 95%: 1.55-2.31). The decision-making variable evolved in a similar fashion to the others, although the analyses revealed a small increase (.367) in the mean of v.2 (CI 95%: .02-.71). Finally, problem solving had the strongest increase in its mean (CI 95%: 1.62-2.44). In the first version, the students obtained a mean score of 3.52 whereas in the second version the mean rose to 6.53).

Table 3

Student's t test for the contrast of hypotheses for the value of a mean

Variables	Contrast value for the mean	N	M	SD	Difference (CI 95%)	Student's t test	
						T	P-sign
DED	5.21	144	5.25	2.17	.40	.220	.413
IND	4.69	144	5.48	1.67	.789	5.663	.000**
PR	6.47	144	8.40	2.32	1.933	9.978	.000**
DM	6.64	144	7.01	2.08	.367	2.116	.018*
PS	3.53	144	5.56	2.49	2.033	9.769	.000**
TOT	26.65	144	31.70	6.49	5.505	9.336	.000**

* Significant at 5% ** Significant at 1%

Globally, the results support our predictions since we observed important changes with v.2 of our program. Properly directed, greater participation and more collaborative work mean that the improvement in critical thinking skills is substantially greater. We observed that the only change in instruction, with version 2 of the program, was greater *activity* and *specificity*; all the rest remained equal. Accordingly it would be reasonable to speculate that these variables would be responsible for the results. We go further into this in the Discussion section.

Discussion and conclusions

Having discussed the analyses, we are now in a better position to assess the progress made in our second version of the ARDESOS program (v.2). Previously we stated that we were seeking to determine whether a change in critical thinking had occurred from one type of instruction to the other in three ways: 1) comparing the effect size in the test of critical thinking; 2) observing whether an improvement had been achieved in the dimensions of critical thinking for which satisfactory results were not obtained with the first version (v.1), and 3) observing whether performance was better with the new version of the program. The above analyses show that with the new version of the program the effect size was considerably improved, leading to a change in all the dimensions of critical thinking. However, in decision-making a positive improvement was observed as regards trend but not with respect to significance. Additionally, the very high values obtained for the effect of the practical reasoning and deduction dimensions are promising. We believe that obtaining these values with the changes introduced into the program means that we should be optimistic or expect similar results in the other dimensions. These, argumentation and deduction, are the dimensions best delimited conceptually, and decision making and problem solving are the least well delimited. Accordingly, once greater precision has been achieved in these latter two, we expect to obtain similar results in these four dimensions.

Across the whole scale and in problem solving the values were moderately high. Regarding the improvements with respect to the pre-post differences, we obtained the same pattern of changes, namely an improvement in all dimensions. However, despite the observation of a positive trend decision-making did not reach statistical significance. Finally, performance in critical thinking improved across the whole scale. This was especially the case of induction, practical reasoning and problem-solving with respect to the first version of the program. Concerning decision making, performance was moderate but acceptable. However, performance in deduction did not improve owing to an anomaly in the procedure used in the first version (see above).

From the foregoing, our conclusions are clear. The results expected from our approach are very positive, with the observation of an effect size, pre-post differences and performance that were quite high across the scale and in some of the dimensions. Only decision making failed to meet our expectations, this dimension showing modest and in some cases non-significant values. By contrast, the problem-solving dimension improved considerably. To understand this lack of consistency in the data -a slight change in decision making and a large change in problem solving- it should be recalled that both dimensions share general items, two and four respectively. The instruction in the current version of the program works the general process of problem solving much more intensely and places less emphasis on specific strategies. A possible explanation for this may lie in the fact that decision making does not benefit from the change in instruction, unlike problem solving. It should also be noted that there is a conceptual difficulty involved in separating these general strategies from these dimensions. The difference between these two dimensions is not clear, because both of them have general items, and it is difficult to know whether they are general items of problem solving or decision making. This is essentially a conceptual problem that we are currently trying to solve.

From the modifications in the instruction corresponding to the current version of the program it may be suggested that in part the problem could be solved by approaching these

strategies based on a single factor (efficiency). This means that they would be used in a context of choice or of solution to obtain the best result possible.

The current improvements in our instruction program partly contribute to solving this conceptual problem. One way of solving it is to use strategies guided by a factor common to the general strategies of problem solving and decision making. This factor is efficacy, which will drive all the strategies, in order to obtain the best result possible or the best solution to the problem approached.

Our prognosis is that these conceptual and empirical difficulties will disappear. In fact, we already have one result pointing in this direction, since having the best explanation of a problem guarantees maximum efficacy and with this many action strategies become superfluous. However, will be addressed in a future work.

References

Almeida, L. (2013).*Ajustamento e sucesso acadêmico no Ensino Superior: Um roteiro para a avaliação das suas dimensões psicológicas.* IV Congresso Brasileiro de Avaliação Psicológica. Maceio 4-7 de Junio de 2013.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Olivares, S., Saiz, C. & Rivas, S.F. (2013). Encouragement for thinking critically. *Electronic Journal of Research in Educational Psychology, 11*(2), 367-394. doi: 10.14204/ejrep.30.12168

Rivas, S.F. & Saiz, C. (2012). Validación y propiedades psicométricas de la prueba de pensamiento crítico PENCRISAL [Validation and psychometric properties of the test of critical thinking PENCRISAL]. *Revista Electrónica de Metodología Aplicada, 17*(1), 18-34.

Saiz, C. (2013). *Pensamiento crítico como reto personal y social*. Palestra en el Programade Pos-graduaçao Stricto Senus en Psicología da Universidade Sao Franciso. 13 de Novembro, 2013. Itativa, Sao Paulo (Brasil).

Saiz, C. & Rivas, S.F. (2008). Evaluación del pensamiento crítico: una propuesta para diferenciar formas de pensar [Assessment of critical thinking: a proposal to distinguish ways of thinking]. *Ergo, Nueva Época, 22-23* (marzo-septiembre), 25-66.

Saiz, C. & Rivas, S.F. (2011). Evaluation of the ARDESOS program: an initiative to improve critical thinking skills. *Journal of the Scholarship of Teaching and Learning, 11*(2), 34-51.

Saiz, C. & Rivas, S.F. (2012). Pensamiento crítico y aprendizaje basado en problemas [Critical thinking and problem-based learning]. *Revista de Docencia Universitaria, 10* (3), 325-346.

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 20 - 37.
doi: 10.14434/josotl.v15i1.12761

Adding confidence to knowledge

Ludwika Aniela Goodson[1], Don Slater[2], and Yvonne Zubovic[3]

Abstract: A "knowledge survey" and a formative evaluation process led to major changes in an instructor's course and teaching methods over a 5-year period. Design of the survey incorporated several innovations, including: a. using "confidence survey" rather than "knowledge survey" as the title; b. completing an instructional task analysis with an instructional designer's perspective of the Gagné framework rather than Bloom's taxonomy; and c. using a rating scale based on established measurement practices for self-efficacy surveys. Results included increased instructor-student interactions; gains in confidence scores from before to after study of course units; high value of the survey for students; changes in grades and confidence scores across teaching methods; and advancement of Chickering and Gamson's principles of good practice in undergraduate education.

Keywords: knowledge survey, confidence, self-efficacy, active learning, assessment, pedagogy, teaching methods, instructional task analysis, cognitive task analysis, formative assessment, formative evaluation, undergraduate education

Introduction

"Love the confidence surveys; more classes need them…"—Student Comment

The Classroom Problem

Most students in the senior-level course planned to become construction site managers, a career path in which the theme for success was "If you don't know the dirt, you'll lose your shirt." Yet, they could barely care about learning how to use soils as construction material. With unwitting complicity, the instructor's own previous instructors had contributed to this problem with their use of traditional lectures and homework. The instructor had used these same methods for twenty years of classroom teaching until the day he decided to leverage his experience in teaching, over forty years in Civil Engineering, and five academic degrees toward making changes. He replaced homework with a presentation project and written report in order to require application of knowledge to actual construction and engineering realities rather than artificial routine problems. But students failed to write articulate reports. What he called a "colossal disaster" became an opportunity as he walked across campus to an instructional designer's "Active Learning" workshop. At the conclusion of this workshop, the instructor made an appointment to discuss further what was happening in his classroom. In subsequent meetings,

[1] Instructional Consultant and Designer, Center for the Enhancement of Learning and Teaching, IPFW, goodsonl@ipfw.edu
[2] Professor, Department of Construction Management and Civil Engineering, Georgia Southern University, fiverings100@gmail.com
[3] Associate Professor, Department of Mathematics, IPFW, zubovic@ipfw.edu

collaborative decisions would soon take his students to destinations never before seen in his course.

With only a couple of months before the semester began, the design and management of group work, simulations, and guided inquiry would wait for later semesters. Although a course is more than a set of learning objectives, the explicit learning outcomes expected for students would help in considering what active learning strategies would align well with those outcomes and be practical for implementing in the coming semester. The instructional designer began with document analysis. She found that the syllabus listed course goals and general statements of learning outcomes and the lecture outlines showed the main topics—suitable for a lecture guide, but indefinite about explicit outcomes. However, she found that the test items required explicit application of specific concepts and principles to realistic engineering challenges. These types of items revealed that the instructor expected deeper levels of learning than were listed in the syllabus and lecture outlines. The instructor confirmed that the test items signaled best the expectations for learning, yet students would not see those expectations until the test administration time arrived.

Connecting Concepts

Before selecting a strategy, the instructor and instructional designer began reviewing several areas of research that seemed relevant to student engagement and motivation. These areas included the rationale for completing an instructional task analysis (Feldon & Stowe, 2009; Smith & Ragan, 2005); use of knowledge surveys in formative assessment (Nuhfer & Knipp, 2003; Wirth & Perkins, 2005); how formative evaluation can help improve instruction (Shepard, 2005, Dunn & Mulvenon, 2009; Herman, 2013); and Keller's ARCS motivation model with its elements of attention, confidence, relevance, and satisfaction (Keller, 2000, 2010).

Instructional Task Analysis: To reframe expectations for students into explicit statements of outcomes, the surest method would be to proceed with an instructional task analysis. Similar to a cognitive task analysis, this process would produce more effective instruction than other ways of identifying content (Clark, Feldon, van Merriënboer, Yates, & Early, 2007; Feldon & Stowe, 2009; Gagné, 1974, 2000; Gagné, Wager, Golas, & Keller, 2005; Smith & Ragan, 2005; Jonassen, Tessmer, & Hannum, 1999; Smith & Ragan, 2005). However, conducting such analysis would require intensive instructor effort and commitment, collaboration in collecting and analyzing knowledge through interviews and documents, and cycles of review and revision until the instructor could clearly map learning outcomes to student performance and back to content.

Knowledge Surveys: The value of the knowledge survey was highlighted when it was cited as one of the best practices reported in the 2001 National Survey of Student Learning (Nuhfer & Knipp, 2003). Such a survey adds value to assessment by providing greater reliability compared to faculty-made tests such as quizzes, midterm exams, and final exams (Nuhfer, n.d.; Nuhfer & Knipp, 2006). It makes a valuable addition to multiple measures, triangulation of data, and ongoing assessment with greater validity, as well as reliability (Sawchuk, 2013; Strayhorn, 2006; Yeasmin & Rahman, 2012). It helps students improve their own learning by engaging them in self-evaluation and self-monitoring (Nilson, 2013; Panadero & Alonso-Tapia, 2014).

Requirements for producing a knowledge survey appeared to be dependent on completing an instructional task analysis. Those requirements include: (1) items that cover all learning outcomes and course content in the same sequence as presented in the course; (2) major themes

in the course; (3) questions like those a student might encounter on a quiz, such as "How do you perform Gaussian elimination?" (Clauss, 2006), or task statements, such as "Make a contour plot that shows the locus of points with a single value of the function" (Frary, 2009); and (4) a way for students to rate their confidence to answer a question or perform a task both before and after instruction. The instructor could look for patterns in survey results and investigate possible changes needed in the course and students could see areas on which to focus their study time.

Nuhfer and Knipp (2003) had provided a blueprint for creating and using a knowledge survey and explained how it could advance Chickering and Gamson's (1987) seven principles for good practice in undergraduate education. This explanation carried high promise for more student-faculty contact, cooperation among students, active learning, prompt feedback, time on task, and addressing diverse ways of learning. Wirth and Perkins (2005) gave an account of how a survey provides full disclosure of course content to students before instruction, the value of a survey as a learning guide, and how the survey process helps students develop self-assessment skills. Wirth and Perkins, like Nuhfer and Knipp (2003), elaborated on how the process of constructing a survey and analysis of data can lead to improved course design and teaching, and their data and analysis showed that knowledge survey scores provided reliable and meaningful measures of learning gains.

Formative Evaluation: Formative evaluation involves assessments during instruction that help instructors make changes before the end of a course (Herman, 2013; Johnson, 2009; Nuhfer & Knipp, 2003; Wirth & Perkins, 2005), and they have a positive impact on student achievement, a practice that itself engages more active learning (Herman, 2013). Whether during or at the end of a course, student feedback helps to inform changes to make in the future (Gilpin, 2013; Nuhfer & Knipp, 2003; Wirth & Perkins, 2005). A knowledge survey is one example of formative assessment.

Motivation Model: Important relationships unfolded during review of Keller's ARCS model (Keller, 2000; Keller, 2010): (1) a knowledge survey at the beginning of the course could gain student *attention;* (2) it would focus on student *confidence*, and furthermore, the pre- and post-survey process and use of the survey as a study guide could support confidence as students focus on what to study; (3) with survey items reflecting all of the course content and organized into thematic units, students could see the *relevance* of what they study; and (4) completing the survey with gains in confidence could allow students to experience *satisfaction* in their accomplishments.

Survey Design

Notwithstanding the *knowledge survey* nomenclature in published studies, the survey was re-titled as a *confidence survey*. Students would see "confidence" in the title and be less likely to think of the survey as a knowledge test. Besides, the survey would be gathering ratings of confidence, not answers about knowledge levels.

Instructional Task Analysis

The instructor listed the major topics for the course, in the sequence taught, and learning outcomes for each topic. Weekly review meetings identified ambiguities or gaps in conceptual and procedural content. Between meetings, the instructor filled in gaps or clarified expected learner performance. These cycles of review and revision are frequently needed with experts

because they typically possess the desirable professional quality of "automated, unconscious knowledge" (Clark, Feldon, Yates, & Early, 2007, p. 590), yet this same quality often leads them to understate the conceptual and/or procedural knowledge in a content domain (Clauss & Geedey, 2010; Feldon & Stowe, 2009; Frary, 2009; Merrill, 2009).

Answers to the instructional designer's questions produced explicit statements of learning tasks. For example, for the topic of *Soils, Investigation, Testing and Classification*, questions included: "Would students need to define soils? What kinds of testing would they need to do? What are the possible classifications of soil?" The resulting statements of learning tasks used performance verbs to signal levels of learning as in the Gagné taxonomy of learning (Gagné, 1977; Gagné, Wager, Golas, & Keller, 2005; Smith & Ragan, 2005), such as: (1) "Describe the mechanical analysis test" (verbal information), (2) "Identify two problem soils" (concept classification), and (3) "Plan and execute a preliminary site investigation" (rule using). These learning tasks then became items in the survey.

This approach to analysis differs from Nuhfer's model, which persistently uses Bloom's taxonomy (Bloom, 1956) with varied methods of application, and varied results (Bell & Volckman, 2011; Bowers, Brandon, & Hill, 2005; Clauss & Geedey, 2010; Marshall & Nuhfer, 2013; Nuhfer & Knipp, 2003; Wirth & Perkins, 2005). The alternative Gagné taxonomy worked well with student outcomes in this course because it makes distinctions between verbal information, attitudes, and psychomotor skills, as well as the hierarchy of intellectual skills defined as concept learning, rule using, and problem solving.

Through the analysis process, the instructor had identified eleven major topic categories and listed explicit learning tasks for the first nine. The number of learning tasks within each ranged from as few as three to as many as forty-one. Because of pressing deadlines to copy the survey for the first day of class, the instructor began analysis of the remaining two topics later in the semester and completed it before the next semester began, at which time he also updated the corresponding survey. Thereafter, he gathered formative evaluation data each semester and continually reviewed and refined the analysis and parallel confidence survey items. The instructional task analysis became his course map for teaching.

Response Scale, Terms, and Layout

On any scale, numbers and their anchoring labels should ascend in the same pattern, from low to high (Bandura, 2006; Kasunic, 2005) such as "cannot do at all" to "highly certain can do" (Bandura, 2006, p. 312), and scales should "measure what they purport to measure" (p. 318). Bandura's scale examples tend to go from 0% to 100%. "Readability" is the critical feature, not so much the number of numbers; that is, the rater needs to see at a glance the construct and how to indicate the strength of belief depicted in the scale, whether as a checklist or a ratings continuum (Bandura, 2006; Kasunic, 2005; Tullis & Dumas, 2009).

Yet, most knowledge surveys have used complex double-barreled directions with mixed constructs in a multiple choice format (Bowers, Brandon, & Hill, 2005; Nuhfer & Knipp, 2003; Wirth & Perkins, 2005). For example, Nuhfer and Knipp's middle choice on a 3-point scale could allow a rating to show 50% confidence in answering a question or in knowing where to find the information needed within 20 or 30 minutes. But to "know" and to "find" are different constructs. Several subsequent studies followed the same complex multiple-choice approach to rating confidence (Bell & Volckman, 2011; Clauss & Geedey, 2010; Fleisher, 2008; Frary, 2009; Price & Randall, 2008).

In contrast, the aim in this project would be to assess ONE construct, namely *confidence to do a task*, and to make the directions and scale easy to interpret by following the standards for self-efficacy rating scales (Bandura, 2006; Kasunic, 2005). This approach would avoid unnecessary cognitive load and split attention (Sweller & Chandler, 1994) because students would not need to focus on interpreting a complex scale while also trying to estimate confidence. The survey question for each topic focused student attention: "How much confidence do you have in your ability to accomplish each of the following objectives and tasks?" and the scale ranged from 1 (no confidence at all) to 5 (complete confidence). This scale was used in all semesters. Students were asked to enter a *score* rather than a *rating*, and instead of asking about confidence before and after *instruction*, students marked scores before and after *study*. Tasks were listed under the directions and scales.

Students could add up the total score for all tasks in a topic category before study, after study, and compute differences in their before-study and after-study scores. The bottom of the form allowed room for students to make comments. Other areas allowed room for ancillary information, such as topic name, student's name, and survey date. Figure 1 illustrates a sample survey for one of the topic categories. Regardless of number of tasks, this same format was used for all topics in the confidence surveys in all semesters.

Score Before Study	**Task Number**	**How much confidence do you have in your ability to accomplish each of the following objectives and tasks?** **5 = complete confidence** **4 = moderate confidence** **3 = some confidence** **2 = a little confidence** **1 = no confidence at all**	**Score After Study**	**Before /After Score Difference**
	1	List two duties of the construction surveyor.		
	2	Sketch two ways to mark grade stakes.		
	3	Use a hand level, Jacob staff, and folding rule to conduct a vertical survey.		
	4	Use a tape to conduct a horizontal survey.		
Total Score _____		**Please make any comments here.**	**Total Score** _____	**Total Score Difference** _____

Figure 1. Survey for the topic of "Layout and Grade Staking"

Formative Evaluation Measures

Over the five-year period of this project, the instructor also added a *formative evaluation* questionnaire, a *pre-evaluation* questionnaire, and a *pretest*. Altogether these measures, along with the *confidence survey*, contributed to the practice known as *formative evaluation* or *formative assessment*.

End-of-Course Formative Evaluation: The end-of-course formative evaluation questionnaire collected information anonymously about the students' learning experiences. The first question asked: "How valuable were the confidence surveys for objectives and tasks?" with a response scale from 1 (no value) to 5 (very valuable). Other questions inquired about what students thought were the most and least important things they learned and what changes they would make.

Pre-Evaluation: After the first semester of using the confidence survey, the instructor reframed the same questions from the end-of-course formative evaluation to ask students at the beginning of the course what they *anticipated* happening.

Pretest: In the next year, the instructor copied the survey items and put them in a pretest format, thus keeping tight alignment with the survey. That is, the instructor took the same questions from the survey and simply added space to write answers under each question, added "Pretest" as the title, and provided directions for how to complete the pretest.

Classroom Procedures

The instructor duplicated and handed out the pre-evaluation, pretest, and survey at the beginning of the course, and the formative evaluation questionnaire at the end of the course. Survey directions explained expectations for students to use the survey throughout the semester, and the survey was posted online for students to review, print, and use as a study guide. Before completing a unit of study, students entered their before-study confidence scores. After completing a unit, they entered their after-study scores and calculated differences in their scores before and after study of the unit. Students handed in the survey score sheet each time they completed it, but could retrieve the survey at any time.

Figure 2 shows the learning methods and formative assessments for five years, eleven semesters, eighteen sections, and 428 students. The confidence survey and formative evaluation were introduced in Year 3 when the instructor began his collaboration with the instructional designer. The pre-evaluation and pretest were introduced in Year 4. Years 3, 4, and 5 each had more teamwork and more quizzes. Year 5 had a process-oriented guided inquiry learning (POGIL) project. Coding for semesters shows number of students in different classes (n_A, n_B, n_C). Updates noted in Figure 2 occurred because of the instructor's ongoing analysis of formative evaluation measures and student performance.

Data Analysis

Course data included survey scores, quiz points, test points, student ratings of the value of the survey, points on projects, total points earned toward final grades, and final grades.

Students' Evaluation of the Confidence Survey: The instructor reviewed students' comments about the survey, observed students using the survey as a study guide, and tabulated value ratings. To determine if there were differences in students' value ratings for the survey across teaching methods, pairwise comparisons were made with the Kruskal-Wallis test, which is a nonparametric equivalent for a one-way analysis of variance (ANOVA).

Confidence Scores: For confidence surveys in Years 3, 4, and 5, the instructor entered before-study and after-study confidence scores in Excel. The survey data pool for statistical analysis included items with slight variations in wording across surveys, for example, a change from *soil* to *soil mass,* and excluded topics with missing data in any teaching year and any item

that had not been used in every survey. As a result, the data pool provided 83 learning tasks for analysis, about 86% of the whole set used in the surveys.

As in previous knowledge survey studies, data analysis compared the average before-study and after-study confidence scores per student per topic. Years 1 through 5 were analyzed separately. Within each year, analysis was completed for all tasks using paired t-tests to test whether the average total score after study exceeded the average total before study. A one-way ANOVA compared changes in the size of gains in confidence scores from before-study to after-study across teaching years. A Kruskal-Wallis test was used to examine the total after-study final confidence scores across the five years.

Year 1 "Old Methods"
♦Lecture ♦Homework 1, 2, 3 ♦ Quizzes 1-5 ♦ Exams 1, 2 ♦ Extra Credit Project, 1-3 points
Total 26: Semester 1 n_A=26
Year 2 "Written Report, Presentations, Peer Review"
♦Lecture Updated ♦Semester Project (replaced homework sets) ♦Written Report ♦Presentations 1, 2 ♦Peer Review ♦Quizzes 1-5 ♦Exams
Total 81: Semester 1 n_B=15 n_C=21; Semester 2 n_A=17 n_B=28
Year 3 "1st Confidence Survey, Team Building, Formative Evaluation"
♦Lecture Updated ♦Confidence Survey ♦Lab Report ♦Team Building Activity ♦Four More Quizzes♦Formative Evaluation
Total 81: Semester 1 n_A=22 n_B=21; Semester 2 n_A=22 n_B=16
Year 4 "Completed Survey, Pretest, Pre-Evaluation, Formative Evaluation"
♦Lecture Updated♦Pre-Evaluation (Aligned with Formative Evaluation) ♦Pretest (Aligned with Survey) ♦Confidence Survey (Updated) ♦Team Project ♦Two More Quizzes♦"Bonus" Quiz♦Three Presentations with Peer Review ♦Formative Evaluation (Updated)
Total 105: Semester 1 n_A=17 n_B=24; Semester 2 n_A=15 n_B=*17*; Semester 3 n_A=32
Year 5 "Addition of POGIL"
♦Lecture Updated♦Pre-Evaluation (Aligned with Formative Evaluation, Updated) ♦Pretest (Aligned with Survey, Updated)♦Confidence Survey (Updated) ♦Kept Teamwork♦Two More Quizzes♦Three Presentations with Peer Review ♦Process Oriented Guided Inquiry Learning Team Project ♦Student Reviews of Team Projects♦Formative Evaluation (Updated)
Total 135: Semester 1 n_A=28 n_B=32; Semester 2 n_A=44; Semester 3 n_B=31

Figure 2. Learning methods and formative assessments

Grades and Points Earned in the Course: The student-topic-average was matched to instructor records of final grading points and letter grades. (Final grading points=numerical total of points earned in the semester. Letter grades=A, B, C, D, or F assigned based on a percent of total possible points, that is, 90-100%=A, 80-89=B, 70-79=C, 60-69=D, less than 60=F.) The number of letter grades earned in each year was counted to examine differences in the grade distribution across teaching methods. Final points earned toward the assigned grades were compared using the Kruskal-Wallis test. To examine the correlation between final points earned

in the course and after-study confidence scores, correlation coefficients were computed for each task separately.

Pre-Evaluation and Formative Evaluation: The instructor reviewed students' comments on the pre-evaluation and formative evaluation questions. No statistical analysis was applied to those evaluations.

Results

Students' Evaluation of the Confidence Survey

Students indicated they gained *relevance* and *satisfaction* with comments such as "…Showed me how much I really learned;" "Gives you a degree of accomplishment;" "…showed a difference of what I thought I knew versus what I learned." Correlations of after-study confidence scores with final grading scores would be expected to further increase student satisfaction. Students also gave high value ratings for the confidence survey as shown in Figure 3. Other charts generated for each assigned grade showed a similar high-value pattern.

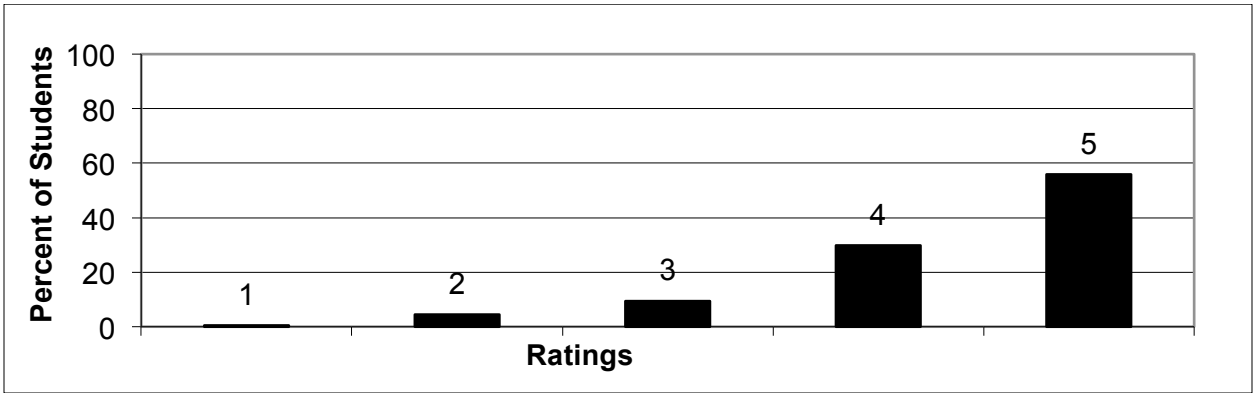


Figure 3. Student ratings for value of confidence surveys; 1=No Value; 5=Very Valuable

The Kruskal-Wallis test showed differences in the value ratings across semesters ($H=61.43$ with $p < .001$). Pairwise comparisons based on Mann-Whitney tests, showed higher value ratings in Year 5 (POGIL added) than in the two previous years—Year 3 (incomplete confidence survey without pretest; $p < .0001$) and Year 4 (complete survey, pretest added; $p < .0001$). There was no significant difference between Years 3 and 4 ($p = .1239$).

Confidence Scores

Charts like the one shown in Figure 4 allowed easy visual comparison of before-study and after-study confidence scores. In this example, before-study was higher for Topic 4 than for the other three topics, signaling the need for some investigation and possible re-evaluation of the planned teaching on this topic. This investigatory process is similar to descriptions by Nuhfer and Knipp (2003) and Wirth and Perkins (2005). In this project, however, separate formative evaluation comments from students also helped to guide the instructor's investigations.

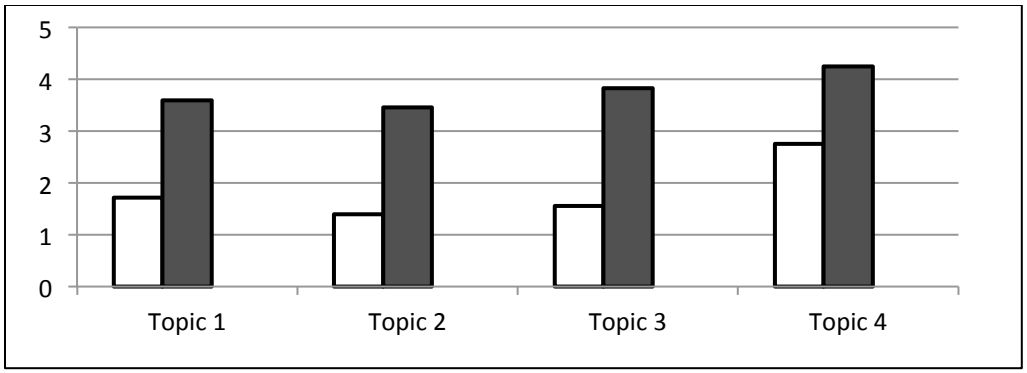


Figure 4. Before-study and after-study average confidence scores for first four tasks, Year 5, Semester 3; White=Before Study, Black=After Study

Besides patterns in records and charts, statistical analysis confirmed the significance of increases in confidence scores from before to after study and showed their correlation with final grading points. For total confidence scores summed across topics, results of paired t-tests for after study minus before study were T=28.4, 30.41 and 42.33 for Years 3, 4 and 5, respectively with $p < .001$ in each case. Similarly, within each topic, the confidence scores increased from before to after study with paired t-test statistics ranging from T=14.77 to T=38.40 with each $p < .001$. Spearman's correlation coefficients for after-study total confidence score and final grading points were significant for Year 3 ($r = .294$ with $p = .008$) and Year 4 ($r = .421$ with $p < .0001$), but not for Year 5 ($r = .025$ with $p = .773$).

In Years 3, 4, and 5, gains in average confidence ranged from 1.92 to 2.76 on the 5-point scale, with the Years 4 and 5 (pretest years) showing the greatest increase ($p < .001$). The size of gains for half of the learning tasks was greatest in the year when the pretest was first introduced (significant differences across years for Task 1 with $p = .040$, Task 4 with $p = .036$, Task 5 with $p = .006$, and Task 6 with $p = .011$), but otherwise showed no differences. The same high-level total after-study final confidence scores occurred across Years 3, 4, and 5.

Grades and Points Earned in the Course

Grades and final points shifted across methods. In Years 1 and 2, before introducing the confidence survey, grades were 30-35% A's, 40% B's, and 20-25% C's. In year 3 when the confidence survey was introduced, grades shifted by 10% from A's to B's, with 25-30% A's, 50% B's. In years 4 and 5, with the introduction of the pretest, grades shifted to more A's and B's and fewer C's: approximately 60% A's, 30% B's and 5% C's. The percentage of students earning below a C was less than 4% in any year, with the highest percentage of D's occurring in Years 1 and 2. The distribution of A, B, and C grades differed significantly across years (Chi Square =72.6 for d.f. = 8 with $p < .0001$). Figure 5 shows the percentage and counts of the grades.

In Years 4 and 5 with the completed revised survey and the pretest, final points earned toward the assigned grades were significantly greater than in Years 1, 2, and 3 (using ANOVA with $F=23.65$, $p < .001$). No statistically significant difference was found in Years 2 and 3. No topic had a statistically significant correlation of confidence scores to final points in Year 5 when the POGIL project was added.

Years	Letter Grades				Totals
	A's	B's	C's	D's	
Year 1	34.62% (9)	42.31% (11)	19.23% (5)	3.85% (1)	100.00% (N=26)
Year 2	31.76% (27)	40.00% (34)	25.88% (22)	2.35% (2)	100.00% (N=85)
Year 3	21.25% (17)	50.00% (40)	27.50% (22)	1.25% (1)	100.00% (N=80)
Year 4	62.50% (65)	33.65% (35)	3.85% (4)	0.00% (0)	100.00% (N=104)
Year 5	62.96% (85)	31.85% (43)	3.70% (5)	1.485% (2)	100.00% (N=135)

Figure 5. Percentage and counts of A, B, C, and D grades by teaching year

Pre-Evaluation and Formative Evaluation

Multiple channels of student input from the confidence survey and the other formative evaluation measures helped to guide changes in the course. Students' responses allowed the instructor to clarify misconceptions or affirm alignment with what was actually planned for course content and activities. Where student confidence scores seemed pervasively low, or students wrote comments such as "I have no clue about this topic," the instructor would expand instruction; where confidence scores were high, the instructor would lean it up. Added dialogue with students about their formative evaluation responses in Year 3 led the instructor to envision the pre-evaluation questionnaire launched in Year 4. Like his use of the pre-study confidence scores, he used pre-evaluation responses as the basis for discussions with students and to make adjustments early in the course.

Discussion

This discussion includes a recap of the context of change, how this project advanced the principles of good practice in undergraduate education, the power of instructional task analysis to clearly identify learning outcomes and improve course alignment, alternative learning taxonomies for coding levels of learning, alternatives for survey implementation, value of the confidence survey, and the use of a confidence survey and other formative assessments as a catalyst for scholarship.

The Context of Change

The initial classroom problem concerned lack of student engagement and motivation to learn. Teaching methods at that time were limited to "old methods" of traditional lecture, homework, and tests. A first attempt at change, though a "colossal disaster," prompted investigation of active learning methods in an instructional designer's workshop. From there, the instructor embraced research-based concepts including instructional task analysis to identify major themes and explicit learning outcomes in his course, the confidence survey, and an end-of-course-formative evaluation. The instructional task analysis required for survey construction produced an organized content map for teaching and learning. Over time, the instructor added more formative assessments and realized his vision for adding more active learning methods. None of the added assessment methods replaced formal grading, but all provided insights to improve instruction and learning.

Principles of Good Practice in Undergraduate Education

Integration of the confidence survey process with the other formative assessments and new teaching methods further advanced the principles for good practice in undergraduate education (Chickering & Gamson, 1987), similar to what Nuhfer and Knipp (2005) advocated for the use of the knowledge survey.

- **Student-faculty contact:** Students' survey ratings and comments inspired the instructor to initiate dialogue on many topics. For example, the instructor discussed with students the reasons for confusion of the terms "stabilization" and "modification" and made subsequent changes on the survey and in the course. In response, students communicated more frequently with the instructor about other issues. Student input on the pre-evaluation and end-of-course formative evaluation also prompted more instructor-student interactions. Students engaged in face-to-face discussions with the instructor before class, during class, after class, and in his office. They sought clarifications and expressed how much more aware they had become of additional knowledge they needed to explore as a result of the before-study survey and pretest.
- **Reciprocity and cooperation among students:** Team-building activities and projects added by the instructor advanced this principle and supported particular learning outcomes the instructor had identified in the instructional task analysis. Teams formed in class became study and peer tutoring groups. In peer-review activities, students evaluated each other's presentations and provided positive communications.
- **Active learning:** Students not only completed the confidence survey, but used it as a study guide. With the pre-evaluation, students wrote about what they anticipated learning, and with the end-of-course formative evaluation, they wrote about their learning experiences and changes they would make in the course. The whole experience of teaching seemed to improve when the instructor went to team projects with in-class presentations instead of written reports. This began in Year 3, which was the first year of the confidence survey and continued in Years 4 and 5.
- **Prompt feedback:** Completing the pretests and survey items gave prompt feedback to students about their knowledge and skill levels. By using the survey as a study guide they could track and record their progress. Their added interactions with the instructor also gave them immediate responses to their questions.
- **Time on task:** The instructor observed students using the survey throughout the course, indicating task-focused time. Student comments disclosed they used the survey to prepare for class, as well as for quizzes and tests. In addition, as the instructor began to teach the content in a more organized and logical manner, students indicated they learned and retained more.
- **Communicating high expectations:** The pretest, with items identical to the survey, contained the same explicit expectations for learning as the survey items. Students reported they could see in the confidence survey what they were expected to study and the different types of tasks for the whole course.
- **Diverse ways of learning:** With the instructional task analysis, the instructor reflected upon the teaching and learning activities that work well for the types of students in his course. For example, he found that the presentations and dialogue played to students' strengths in ways that matched up with the roles they would encounter in their future jobs. His reflections led to deployment of an increasing variety of activities, giving

students the opportunity to learn in more diverse ways.

Instructional Task Analysis

Before developing a survey, standards of practice require an analysis to be completed in such a way that different levels of task demands within a domain become clear (Bandura, 2006; Nuhfer, 2003; Nuhfer & Knipp, 2003). Instructional task analysis, the process used by instructional design practitioners (Gagné, Wager, & Golas, 2005; Smith & Ragan, 2005), fulfills this survey analysis requirement. The instructional task analysis produced the content for the confidence survey and pretest, and prompted changes in the instructor's organization of the course. The analysis process included: document analysis, unstructured interviews, learning hierarchy analysis, and "a multi-stage interview technique that captures the automated and unconscious knowledge" of the content expert (instructor) (Clark, Feldon, van Merriënboer, Yates, & Early, 2007, p. 106).

As a result, the instructor reflected upon and articulated the significant learning in the course (Fink, 2007) along with implications for what teaching and learning activities should be developed. Other published studies, with variable methods of analysis, have reported similar impact of the survey development process on instructor reflections about course content and learning outcomes (Bell & Volckman, 2011; Bowers, Brandon, & Hill, 2005; Clauss & Geedey, 2010; Frary, 2009; Wirth & Perkins, 2005). Furthermore, survey design was not a one-shot process. For example, after the instructional task analysis in Year 3, the instructor prepared for Year 4 by using student input and further reflections to refine his analysis and reorganize the survey. Years 4 and 5 both had updates, but Year 4 had the greatest changes in survey items with the re-organization of tasks in the first topic and the completed analysis for the last two topics.

In this project, the pretest and survey had tight alignment with each other, and the final exam and projects required students to integrate the knowledge and skills articulated therein. This alignment could account for the positive shift in grades, a persistent phenomenon reported in a much earlier review of studies by Cohen (1989). In several knowledge survey studies, the alignment of analysis with surveys and exams has varied. On the one hand, Bell and Volckman (2011) reported that authors of the survey also were authors of the tests, and the resulting survey aligned well with the same distribution of topics and levels of learning as taught in the course. On the other hand, Bowers, Brandon, and Hill (2005) reported that different instructors developed different exams for their different sections of the course. And, Clauss and Geedey (2010) reported that faculty wording of survey items produced some confusion for the research assistants who did the actual coding of levels of learning. While dividing the analysis workload may be practical, it may not produce tight alignment of the analysis with the survey and assessments.

To support instructional task analysis, this project used the Gagné framework. The learning task statements simply signaled explicit learning outcomes. Some were information level, such as "Name the four cycles of particle angularity." Others were higher order, such as "Draw and interpret a PI/LL chart." The focus was on explicit clarity without manipulation to add more high-level learning outcomes beyond those identified through the analysis process. This has not been the case in other studies which reported the practice of adding more high-level learning outcomes after the analysis was completed (Bell & Volckman, 2011; Bowers, Brandon, & Hill, 2005; Clauss & Geedey, 2010; Frary, 2009; Nuhfer & Knipp, 2003; Wirth & Perkins, 2005). In those instances, the researchers may have used the process to prompt consideration of

how high-level outcomes might fit with the course, but sometimes it seemed the push was to add more simply for the sake of having more.

Learning Taxonomies

Coding of levels of learning might be distinctively different with different taxonomies. The Gagné framework (Gagné, 1977; Gagné, Wager, & Golas, 2005) could be compared to Bloom's taxonomy or to the "new Bloom" (Anderson & Krathwohl, 2001; Krathwohl, 2002). Future research could advance the coding of learning levels by making a logical choice among taxonomies to fit the type of content in a course. For example, learning that involves safety, health, and life as in the case of nursing (Harper, 2007), often depends on psychomotor skills and affective dispositions not addressed in Bloom's cognitive taxonomy. However, if choosing Bloom's taxonomy, using the full set of levels and sub-levels from the original publication (1956) could produce greater reliability of coding than by using only its major categories and general descriptions. Other choices include content-specific taxonomies like an *engineering taxonomy* (Girgis, 2010) or a pocketful of other taxonomies identified by Anderson and Krathwohl (2001) and Moseley, Baumfield, Elliott, Higgins, Miller, Newton, and Gregson (2005).

Survey Design

Confidence survey items tightly aligned with the instructional task analysis. The confidence survey used a Likert-type scale for the single construct of *confidence* to do each task within each topic category. Future studies could compare results for a single construct on a simpler 3-point continuum, a checklist, or the longer 0 to 100% continuum often used for efficacy scales (Bandura, 2006). The 0 to 100% scale could yield visual displays in line with final student scores based on a 100-point scale. If an instructor prefers the Nuhfer model and thinks getting the information is an important level of "confidence," he or she could split out the mixed construct item format into separate items—one for each construct.

Survey Implementation

As predicted by Nuhfer (2003), the survey offered time-efficient comprehensive assessment, allowing students to score confidence for many items in a "very short time span" (p. 59). Frary (2009) suggested: "The instructor's comfort level with each medium and the length of the survey will determine the best method" (p. 8) for implementation. But methods could be compared, such as the pen and pencil ratings on printed surveys, as in Bell and Volckman's (2011) study and in this project, versus surveys completed on an Apple-based mobile device, like an iPad, on a personal computer, a smart-phone, or with other technology tools.

Value of Confidence Survey

Developing and using the confidence survey was of high value to the instructor and using it was of high value to students. Students used it as a study guide, an observation similar to other anecdotal reports (Bowers, Brandon, & Hill, 2005; Clauss & Geedey, 2005) and student confidence increased from before to after study. However, changes in points earned in the course

and redistribution of grades may be due to the combination of formative assessments and more active learning methods. Although Year 5 showed higher value ratings, it had lower confidence correlation with grades. But final points and grades were better, and the lower correlation might be due to the added effort that often accompanies POGIL (Vanags, Pammer, & Brinker, 2013).

Another consideration is that taking the pretest before the survey in Years 4 and 5 may have allowed students to make a more honest appraisal of their before-study confidence levels than in Year 3. In other words, students may have had improved metacognitive confidence (better evaluative judgment) which was the knowledge survey focus for Bell and Volckman (2011). The pretest, survey, and study process together may have contributed to improved judgment of after-study confidence. Changing the time and place for taking the survey also might improve metacognitive confidence as suggested by Nuhfer and Knipp (2003), namely: "The best results occur when survey items clearly frame specific content, and students take the survey home to complete it with plenty of time for self-reflection" (p.5). A useful future study would be to find verification for this proposition compared to other methods.

A Catalyst for Scholarship

Nuhfer and Knipp (2003) and Wirth and Perkins (2005) reported that the process of making a confidence/knowledge survey improves course organization and preparation. This study shows the same impact. Whether with an instructional designer or a colleague, the process can engage reflection about what content should be taught, in what sequence, how students learn, and what teaching methods to use. Integrating design and use of a confidence survey with other formative assessments allows the instructor to gain more knowledge about students and to accommodate teaching methods to address their needs. For instructors who care about the scholarship of teaching and learning, this endeavor can be a worthy investment, as it was for this instructor and for all of his students.

References

Anderson, L.W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Addison, Wesley, Longman, Inc.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan & F. Pajares (Eds.), *Self-Efficacy Beliefs of Adolescents*, (pp. 307-337). Charlotte, NC: Information Age Publishing.

Bell, P., & Volckmann, D. (2011). Knowledge surveys in general chemistry: Confidence, overconfidence, and performance. *Journal of Chemical Education, 88*(11), 1469-1476. doi: 10.1021/ed100328c

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives, the classification of educational goals—Handbook I: Cognitive domain*. New York: David McKay Company, Inc.

Bowers, N., Brandon, M., & Hill, C.D. (2005). The use of a knowledge survey as an indicator of student learning in an introductory biology course. *Cell Biology Education: a Journal of Life Science Education, 4*(4), 311-322. doi: 10.1187/cbe.04-11-0056

Chickering, A. W. & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *American Association of Higher Education Bulletin, 39*, 3-7.

Clark, R. E., Feldon, D. van Merriënboer, J. J. G., Yates, K. & Early, S. (2007). Cognitive task analysis. In J.M. Spector, M.D. Merrill, J.J.G. van Merriënboer, & M.P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Clauss, J. (2006). *Moodle questionnaire*. Retrieved from http://dmc.augustana.edu/KS/KSFinMA329.pdf

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

Dunn, K. E. & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7).

Fink, D. L. (2007). The power of course design to increase student engagement and learning. *Peer Review*, *9*(1).

Fleisher, S.C. (2008). *Knowledge survey: Psychology 211 Survey of Psychology (Cognition and Learning)*. Retrieved from http://familymedicine.medschool.ucsf.edu/paetc/resources/asilomar_2009/psychology.pdf

Feldon, D. F. & Stowe, K. (2009). A case study of instruction from experts: Why does cognitive task analysis make a difference? *Technology, Instruction, Cognition and Learning, 7,* 103-120.

Frary, M. (2009). *Knowledge surveys*. Retrieved from http://scholarworks.boisestate.edu/cgi/viewcontent.cgi?article=1000&context=ctl_teaching

Gagné, R. M. (1974). Task analysis—its relation to content analysis. *Educational Psychologist*, *11*(1). doi:10.1080/00461527409529118

Gagné, R. M. (1977). Analysis of objectives. In L.J. Briggs, (Ed.), *Instructional Design: Principles and Applications*. Englewood Cliffs, NJ: Educational Technology Publications.

Gagné, R. M., Wager, W. W., Golas, K., Keller, J. M. (2005). *Principles of instructional design,* (5th Ed). Belmont, CA: Wadsworth/Thomson Learning.

Gilpin, L. (2013). Enhancing teaching and learning. *Mountain Rise: The International Journal for the Scholarship of Teaching and Learning, 8(1).*

Girgis, M. (2010). A new engineering taxonomy for assessing conceptual and problem-solving skills. Proceedings of the American Society for Engineering Education Conference, USA, AC 2010-1793.

Harper, S. P. (2007). *Instructional design for affective learning in online nursing education*. (Doctoral dissertation). Capella University. Retrieved from http://udini.proquest.com/view/instructional-design-for-affective-goid:304722378/

Herman, J. L. (2013). Formative assessment for next generation science standards: A proposed model. Resource Paper No. 16. National Center for Research on Evaluation, Standards, and Student Testing. Retrieved at https://www.cse.ucla.edu/products/resource/cresst_resource16.pdf

Johnson, E. (2009). Formative and summative assessment. Retrieved at http://www.education.com/reference/article/formative-and-summative-assessment/

Jonassen, D. H., Tessmer, M., & Hannum, W. H. (1999). *Task analysis methods for instructional design*. NJ: Lawrence Erlbaum Associates.

Kasunic, M. (2005). Designing an effective survey. Carnegie Mellon University. Retrieved at http://www.sei.cmu.edu/reports/05hb004.pdf

Keller, J. M. (2000, February). *Applying the ARCS model of motivational design*. Paper presented at VII Semanario, Santiago, Cuba.

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. New York, NY: Springer.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4).

Marshall, S. & Nuhfer, E. B. (2013, November). *Knowledge surveys and course-based learning outcomes for soil science courses*. Presentation at Soil Science Society of America Symposium, Tampa, Florida.

Merrill, M.D. (2009). *What makes e3 (effective, efficient and engaging) instruction*? Presented at World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009. Retrieved from http://www.editlib.org/p/31461

Moseley, D., Baumfield, V., Elliott, J., Higgins, S., Miller, J., Newton, D. P., & Gregson, M. (2005). *Frameworks for thinking: A handbook for teaching and learning*. Cambridge, UK: Cambridge University Press.

Nilson, L. B. (2013). *Creating Self-Regulated Learners: Strategies to Strengthen Student Self-Awareness and Learning Skills*. Sterling, VA: Stylus Publishing, LLC.

Nuhfer, E. (n.d.). Knowledge surveys (KS). *MERLOT ELIXR*. California State University Channel Islands. Retrieved from http://elixr.merlot.org/assessment-evaluation/knowledge-surveys/knowledge-surveys2

Nuhfer, E. & Knipp, D. (2003). The knowledge survey: A tool for all reasons. *To Improve the Academy*, *21*, 59-78.

Nuhfer, E. B. & Knipp, D. (2006). Re: The use of a knowledge survey as an indicator of student learning in an introductory biology course. *CBE Life Science Education, 5*(4): 313–314. doi: 10.1187/cbe.06-05-0166

Panadero, E. & Alonso-Tapia, J. (2014). How do students self-regulate? Review of Zimmerman's cyclical model of self-regulated learning. Anales de Psicología, 30(2), 450-462.

Price, B. A. & Randall, C. H. (2008, May/June). Assessing learning outcomes in quantitative courses: Using embedded questions for direct assessment, *Journal of Education for Business*, 288-294. doi: 10.3200/JOEB.83.5.288-294#preview

Sawchuk, S. (2013). Combined measures better at gauging teacher effectiveness, study finds. *Education Week*, *32*(17), 1-16.

Shepard, L. A. (2005, October). *Formative assessment: Caveat emptor*. Presented at ETS Invitational Conference: The Future of Assessment: Shaping Teaching and Learning, New York.

Smith, P. L. & Ragan, T. J. (2005). *Instructional design* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Strayhorn, T. L. (2006). *Frameworks for assessing learning and development outcomes*. Retrieved from Council for the Advancement of Standards in Higher Education website: http://nau.edu/Student-Affairs/_Forms/FALDO/

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition & Instruction, 12*(3), 185-233.

Tullis, T. & Dumas, J. (2009, May). *Rating scales: What the research says*. Presented at the Boston Usability Professionals' Association Mini-UPA Conference. Retrieved from http://www.measuringuserexperience.com/Mini_UPA-Dumas-Tullis.ppt

Vanags, T., Pammer, K., Brinker, J. (2013). Process-oriented guided-inquiry learning improves long-term retention of information. *Advances in Physiology Education, 37*, 233-241. doi: 10.1152/advan.00104.2012

Wirth, K. R. & Perkins, D. (2005). Knowledge surveys: An indispensable course design and assessment tool. *Innovations in the Scholarship of Teaching and Learning*. Retrieved from http://www.macalester.edu/geology/wirth/WirthPerkinsKS.pdf

Yeasmin, S. & Rahman, K. F. (2012). 'Triangulation' research method as the tool of social science research. *Bangladesh University of Professionals Journal, 1*(1), 154-163.

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 38 - 56.
doi: 10.14434/josotl.v15i1.12984

Periodic review sessions contribute to student learning across the disciplines in Pharmacology

Orla P. Barry[1], Eleanor O' Sullivan[2], and Marian McCarthy[3]

*Abstract: Background: The teaching of the discipline of pharmacology is in constant flux. In order to meet the challenges of teaching pharmacology effectively we investigated a new teaching and learning strategy. Aim: Our aim was to investigate whether structured periodic review sessions (RS) could improve teaching and learning for students in a multidisciplinary undergraduate pharmacology module. Methods: Following each lecture students were asked to identify topics of difficulty in pharmacology using the one minute paper classroom assessment technique (CAT). Three review sessions were then introduced based on the problematic issues identified by students. They completed a pre- and post-review session multiple choice question (MCQ) examination to gauge improvements in their learning. Feedback was obtained from students at the end of the module regarding the acceptability, advantages and limitations of the CATs and the review sessions. Results: There was active participation by students in all thirteen CATs (71.15% ± 1.2%), three review sessions (78.3% ± 1.6%) and the end of module (EOM) questionnaire (81%). A significant increase in student learning across all disciplines was observed in all three review sessions (**, $p<0.01$; ***,$p<0.001$ and **, $p<0.01$ for review sessions 1-3 respectively). The majority of students (99%) expressed the opinion that the review sessions enhanced their learning of pharmacology. A limitation expressed by some students was the necessity to complete multiple CATs to support each of the three review sessions. Conclusion: Strategically structured student directed review sessions which are carefully designed and executed are an interesting and effective educational tool for improving and complementing student learning across the disciplines in pharmacology.*

Keywords: Review sessions, classroom assessment technique, undergraduate pharmacology, questionnaire, multidisciplinary.

Introduction

The teaching of pharmacology has progressed substantially in recent years from a passive traditional didactic lecture format to a more interactive inclusive student/teacher model (Rogers, 2012; Sekhri, 2012; Zgheib et al., 2010). Of note, however, the delivery of pharmacological information in a conventional lecture is still required prior to engaging with new teaching and learning strategies (Fry et al., 1999; Rogers, 2012). In this way students can understand and familiarise themselves with the topic(s) before re-addressing, re-enforcing and supplementing key principles using alternate strategies. In addition they are a fact of reality given growing

[1] Department of Pharmacology and Therapeutics, Room 3.89 Western Gateway Building, Western Road, University College Cork,Cork, Ireland. o.barry@ucc.ie
[2] Oral Surgery, Cork University Dental School & Hospital, Cork, Ireland. eleanor.osullivan@ucc.ie
[3] Teaching and Learning Centre, Ionad Bairre, University College Cork, Ireland. mmcarthy@ucc.ie

student numbers as well as financial pressures (Brown & Gamber, 2002). The error lies, however, when only didactic lectures are utilised which alone does not give the students the opportunity for deep learning (Bligh, 2000; Handelsman et al., 2004; McKeachie et al., 2006; Udovic et al., 2002).

The challenge to teach pharmacology effectively continues on a daily basis (Joshi & Trivedi, 2010). The direct consequence of increased drug discovery has led to a continuous increase in the associated scientific information. Thus, it has become an almost impossible task to teach students all pharmacological facts relating to individual drugs (Achike & Ogle, 2000). Only recently has significant consideration been given to the complexities of teaching and learning of pharmacology. Novel teaching techniques are now in vogue aimed at promoting a more integrated and efficient programme as well as improving the quality of teaching in the discipline and in turn student learning. These include wet practicals where students apply their knowledge in a practical context in the laboratory (Rogers, 2012; Sekhri, 2012), computer assisted learning which provides a dry laboratory experience (Coleman et al., 1995; Dewhurst and Norris, 2003), problem based learning (Gregson et al., 2010), small group learning (Morgan, 2003; Sekhri, 2012), team-based learning (Zgheib et al., 2010), use of structured role play (Morgan, 2003), case studies (Routledge 1999; Rogers, 2012), collaborative online learning, (Wiecha, 2003), audio visual aids (Banerjee et al., 2013), drug flash cards (Rogers, 2012) and electronic hospital libraries (Brewer & Hiscock, 2001). Despite such interventions, however, many students continue to find it difficult to master core knowledge in pharmacology. They are often left memorizing unfamiliar terms, drug names and pathways with a very limited frame of reference.

To-date there is a paucity of studies investigating effective pharmacology teaching for multidisciplinary student learning within the same class (Darling-Hammond, 2008; Gardner & Boix Mansilla, 1994; Kwan, 2002). Different topics and methodologies have been reviewed for teaching pharmacology to medical, dental, and nursing students (McBane & Mesaros, 2010). However, pharmacological knowledge is essential not only for students who will be involved in clinical practice but also for those involved in the biological sciences (Kwan, 2002; Lymn & Mostyn, 2010). The composition of the class chosen for this research project consists of chemistry, biochemistry, physiology, medical and visiting Erasmus (European Union student exchange programme) students. The last group (Erasmus students) have traditionally been part of the cohort of the class taking this module. Their backgrounds are closely aligned with other student groups in the class. While pharmacology is strictly related to physiology (and pathology), the challenge in teaching this class is to define a core of knowledge in pharmacology for each discipline according to the practice of each profession. Furthermore, the teaching of pharmacology to medical students differs from that to science students as medical students not only have to master core concepts in pharmacology but must also be able to relate this information to the relevant management of the underlying disease.

The purpose of this study was to explore a new teaching and learning strategy in pharmacology for students from multidisciplinary backgrounds. Previously the module consisted of didactic lectures with three associated laboratory practicals. We have now replaced some lectures with periodic RS which complement the traditional instructive lectures. Importantly, the teaching in the RS did not occur at the expense of students acquiring a lesser amount of essential knowledge in the lectures. The RS were designed to focus and highlight relevant pharmacological material for each discipline without providing excessive information. Individual student conducted CATs guided the design and activities of the RS. CATs are formative

evaluation methods whose main aims are to permit assessment of student understanding and learning and to allow for timely feedback about the effectiveness of particular teaching techniques. Approximately fifty different CATs have been described to-date (Angelo, 1991a,b; 1998; Angelo & Cross, 1993; Byon, 2005; Rouseff-Baker & Holm, 2004) with limited qualitative and quantitative research assessing their value in improving student learning (Cottell & Harwood, 1998; Simpson-Beck, 2011). Some CATs access student prior knowledge, recall and understanding (minute paper, muddiest point, background knowledge probe and memory matrix), while others assess student's skill set in analysis and critical thinking (pro-and-con grid, categorizing grid, defining features matrix and content, form and function outlines). Furthermore, CATs permit assessment of students' skills in synthesis and creative thinking (concept maps), problem solving (problem recognition tasks), application and performance (application cards) and student self-awareness as learners (autobiographical sketches). Although the various CATs differ in their complexity and the time necessary for their preparation, administration and final analysis, they can all be utilised to improve course content, teaching methodologies and ultimately student learning. It is important to be aware, however, that irrespective of the CAT employed as with all educational interventions, due consideration must be given to the different learning styles of the students from the different disciplines. The approach taken in our CATs and the RS may not appeal to each and every student given different learning preferences (Gardner, 1995, 1999). Thus, in an effort to accommodate different learning strategies we incorporated different strategies into our RS including visual and auditory techniques. In brief, we observed that the RS aided students' learning and deepened their understanding in pharmacology. They permitted students to evaluate, synthesise and apply their knowledge of pharmacology more readily (Bloom, 1956; Cronin Jones, 2003; Ebert-May et al., 1997; Litke, 1995). Thus, periodic RS are an interesting and cost-effective tool worth exploring for effective teaching and learning to a disparate group of undergraduate students studying pharmacology.

Method

Participants

Research was conducted with an undergraduate pharmacology class of eighty nine students, forty eight of whom were male (54%). Students were from four different disciplines i.e. chemistry, biochemistry, physiology, medicine as well as visiting Erasmus students. All students had completed at least one pharmacology module prior to enrolling in this module (PT2101/PT3005). The module was taught in three distinct sections namely (i) inflammation and anti-inflammatory drugs, (ii) antimicrobial chemotherapy and (iii) cancer chemotherapy. Students were delivered didactic lectures (as was traditionally done within this module) but also engaged in CAT, MCQs and RS which were newly introduced for our research purposes.

Design, conducting and analysis of student CATs

(i) *Design*: Students were asked to complete a one minute paper (Angelo & Cross, 1993) at the end of each lecture stating (i) two pharmacological aspects that they understood well in the lecture and (ii) two pharmacological aspects that they found difficult to understand in the lecture. Topics listed in the latter question provided the basis for the RS. (ii) *Conducting*: At the end of each fifty minute didactic lecture students were given sufficient time to complete a CAT (this

normally takes longer than one minute so students can process the lecture material and provide a response (Stead, 2012)). CATs are an invaluable learning tool not only for students but for the teacher as well. The advantages and disadvantages of the one minute paper have recently been reviewed (Stead, 2012), (iii) *Analysis*: CATs from all participating students were carefully read to ascertain the areas of student difficulties as well as ease in the preceding lecture. CATs were pooled from the first four lectures for RS1, the next four lectures for RS2 and the remaining five lectures for RS3. Topics of difficulty were entered into an excel sheet and formed the basis for each of the three subsequent RS. Each RS was based on six pharmacological topics most frequently identified as difficult by students.

Description of the three Review Sessions

RS1 and RS2 were conducted after a block of four lectures each covering inflammation and antimicrobial chemotherapy respectively. RS3 occurred following the remaining five lectures in the module covering cancer chemotherapy.

Design and conducting of Review Sessions including MCQ examinations

(i) Review topics: Six topics of difficulty identified most frequently by students formed the basis of the material to be re-addressed in the RS as well as the pre- and post-review MCQ questions. A single MCQ topic was assigned to each of the six problematic areas with five different questions of format T/F/D (true/false/don't know) associated with the topic. *(ii) Format of the RS*: Students completed the MCQ questions (6 topics x 5 T/F/D MCQ questions =30 questions in total) in the first fifteen minutes. The next thirty minutes was devoted to revisiting the six problematic areas identified by students in their CATs. The review time was used to reinforce the important and fundamental aspects of the six different pharmacological topics using an interactive teacher/student powerpoint presentation. The remaining fifteen minutes was allocated to the students re-sitting the same MCQ questions presented in the pre-review MCQ. Both pre- and post-review MCQs were printed on different colored paper and stapled together to allow analysis of student learning from individual students, while maintaining student anonymity. All MCQs were conducted anonymously except through identification of student discipline.

Strategies to design the Review Sessions and MCQs

Key challenges in designing and conducting the RS (as well as teaching the module in general) was to keep to the forefront the knowledge that the student group was diverse and from different disciplines. Thus, different criteria were adhered to in designing the MCQ examination questions (which did not differ in the pre- and post-review MCQ examination) as well as how the material identified by students was to be re-visited during the RS. Four different strategies were employed: (i) relevant; questions were designed and material was reviewed at an appropriate level to reflect the backgrounds, needs and diversity of students. All material was closely aligned with learning outcomes of the module; (ii) realistic; the level of pharmacology that the students required at this point in their training was taken into account (iii) engaging; both MCQ questions and material reviewed incorporated a mix of straight forward knowledge based enquiry as well as more challenging aspects of pharmacology (Bloom's Taxonomy, 1956) and (iv) instructional;

the questions and material readdressed helped to inform and assess students' core and applied pharmacological knowledge.

Analysis of pre- and post-review MCQ results

Each participating student (identified by discipline only) submitted a pre- and post-review MCQ sheet for analysis, totalling sixty answers per student per RS. The correct answers to the thirty MCQ questions were firstly entered into an excel sheet, followed by student T/F/D answers to the pre- and post-review MCQs. A plus one value was assigned to each correct answer and a zero value to an incorrect answer or an unanswered question. Both the pre- and post-review MCQ results were analysed for each student. Lastly, the change in each student answer for all thirty MCQ questions was calculated between the pre- and post-review answers as follows; a value of zero was assigned to no change to a correct or incorrect answer, plus one for an incorrect to a correct answer and minus one for a correct to an incorrect answer.

Questionnaire

The written questionnaire is a powerful tool in research (Oppenheim, 1992, Wellington, 2000). An EOM questionnaire was designed to assess student feedback on teaching, learning and assessment methods in the RS. Students from all five disciplines completed the questionnaire in the last fifteen minutes of a two hour tutorial. The first half of the questionnaire consisted of seven statements with students indicating whether they strongly agreed (SA), agreed (A), neutral (N), disagreed (D) or strongly disagreed (SA) with each of the statements. This part of the questionnaire was designed to determine quantitative data on student views of the RS as an unorthodox teaching, learning and assessment tool. The second half of the questionnaire invited qualitative comments from students on their experiences of the RS in an effort to identify advantages and disadvantages of this novel technique. Similar to the student CATs and pre- and post-review MCQ submission, the questionnaire was anonymous except for identification of the student discipline.

Statistical Analysis

Results are expressed as mean ± S.E. Statistical comparisons were made by using analysis of variance with subsequent application of Student's *t* test, as appropriate. GraphPad InStat 3 software was used for statistical analysis also.

Ethical approval

All students provided their signature at the commencement of the module indicating their willingness to participate in the different teaching, learning and assessment aspects of the RS.

Results

CAT participation and analysis

All students present actively participated in the CATs following each lecture. Mean total class participation for all thirteen CATs was 71.15% ± 1.2%. Interestingly there was no statistically significant difference in student attendance and participation (there was 100% correlation) in all thirteen CATs despite frequent EOM examinations in other subjects towards the end of the teaching period (data not shown).

The most frequently identified areas of difficulty outlined by students in their CATs can be summarized as follows; (i) new terminology, (ii) cellular pathways, (iii) drug pharmacodynamics i.e. the mechanism of action of drugs, (iv) drug inducers and inhibitors, (v) cell cycle control and (vi) multi-drug resistance. More specifically students outlined their misunderstandings and knowledge gaps relating to pro- and anti-inflammatory pathways including the arachidonic acid pathway, the peroxisome proliferator-activated receptor and the hypothalamic pituitary adrenal axis (Table 1). Clarification of drug pharmacodynamics was outlined in student CATs across all thirteen lectures (Table 1). A gap in their knowledge regarding drug inducers and inhibitors related to the pharmacology of the cytochrome P450 family of microsomal enzymes (Table 1). Students also requested revision of cell cycle control in the context of cancer proliferation and apoptosis, anti-microbial resistance and anti-cancer multi-drug resistance (Table 1).

The CATs aided teaching as well as student learning as the module progressed. Students expressed (in their EOM questionnaire) that they paid significantly more attention and concentrated for a longer period of time (more than the average twenty minutes, McBane & Mesaros, 2010) in the lectures in order to be able to complete the CATs. As each CAT was performed following a single lecture they provided an opportunity to access the standard of teaching and student learning in real time. Timely analysis of the CATs offered a window of opportunity to correct some teaching aspects which may have required adjustments prior to the subsequent lecture.

Review session teaching and student learning

The purpose of the RS (designed and conducted as outlined in *Methods)* was to complement the teaching in the lectures and to aid student learning. A significant proportion of the class attended and participated in each review session (78.8%, 80.9% and 75.3% for RS1, 2 and 3 respectively). The highest participation in all three RS was seen in the medical student group with a mean value of 86.7% ± 2.7%, the lowest observed in the physiology student group with 65.10 ± 1.6%. This is an interesting observation as the module is an elective special study module for medical students but a compulsory module for physiology students.

Teaching: In RS1 students were taught how to distil down the various pathways into their associated elements focusing on the sites that are specifically targeted with clinically available drugs. Teaching was initially centered on pathway umbrella terms with more specific terms being introduced afterwards (Table 2). Teaching (partly in RS1) and mainly in RS2 focused on drug classes rather than on multiple individual drugs. A key difficulty in teaching pharmacology is dealing with the sheer volume of drugs that are clinically available. Students find it extremely difficult to remember drug names, associate each drug with a particular drug class and to memorize and retain the pharmacological components of each

Table 1

Student CAT analysis

Lecture number	Analysis of student CATs	Six main areas of student difficulty
1-4	Pro- and anti-inflammatory pathways	(i) AA, (ii) PPAR, (iii) HPA
	Cox-2 specific inhibitors	(iv) Mechanisms of drug specificity
	Paracetamol	(v) Mechanism of drug poisoning and treatment
	Glucocorticoids	(vi) Mechanisms of action
5-8	Antibiotics	(i) β-lactam cell wall synthesis inhibitors
	Drug inducers and inhibitors	(ii) Cytochrome P450 family
	Folate as a pharmacological tool	(iii) Folate synthesis, metabolism and utilisation
	Anti-virals	(iv) Mechanism of action of acyclovir
	Antifungals	(v) Mechanism of action of flucytosine
	Anti-malarials	(vi) Mechanisms of anti-malarial drug resistance
9-12	Cancer genetics	(i) Prot-oncogenes versus oncogenes
	Anticancer drugs	Mechanisms of action of (ii) topoisomerase inhibitors, (iii) aromatase inhibitors, (iv) EDGF inhibitors
	Cancer proliferation and apoptosis	(v) Cell cycle control
	Drug resistance	(vi) Cancer multi-drug resistan

All CATs were a one minute paper completed by students at the end of each lecture. Six main areas of student difficulties were identified from lectures 1-4 for RS 1, lectures 5-8 for RS2 and lectures 9-13 for RS 3. AA; arachidonic acid, PPAR; peroxisome proliferator-activated receptors, HPA; hypothalamic pituitary axis.

individual drug. Thus, we employed a teaching technique using prototypical drugs i.e. a standard or typical drug which is a lead agent in a drug class (family). These are common agents used in exam questions. Knowledge of essential pharmacological principles of prototype drugs will later permit students to apply this knowledge instantaneously to similar drugs within the same drug class. Armed with core information using the prototype approach provides an adequate learning basis and knowledge store for both non-clinical and clinical years ahead. Thus, teaching in RS2 focused on drug classes using prototypical drugs rather than individual drugs. Key drug information (using the prototype drug) including name, how it works and major side effects were taught. Application of core knowledge was reinforced as opposed to the expectation that students should simply memorize vast volumes of drug data (Table 2). Also in RS2 students were taught to recognize potential drug-drug interactions using the cytochrome P450 (Cyt P450) family of liver microsomal enzymes as examples (Table 2). Drug classes were identifying as either substrates, inducers or inhibitors of Cyt P450 members initially. Individual drugs falling outside these three categories were outlined. Finally in RS3 anti-cancer drug pharmacology was

reinforced by dividing drug classes into older non-targeted drugs versus newer specific targeted drugs. Commonly observed side-effects depending on whether the drug was a new or old chemical was highlighted to students. Again teaching in the RS complemented rather added to the material delivered in the lectures. Targeted review of student focused topics allowed for pharmacology to be revised in a cohesive, clear and concise manner. Reinforcement and in some instances repetition of some pharmacological topics helped deepen student understanding.

Table 2

Teaching and learning in the three review sessions

Review session number	**Teaching**	**Student learning**
1	Inflammatory pathway(s) steps and terminology	(i) Umbrella terms (ii) Specific terms
2	Key aspects of drug classes Drug-drug interactions	(i) Pharmacology of drug classes using prototype drugs (ii) Apply knowledge to clinically relevant drugs (iii) Typical drug combinations
3	Cancer	(i) Cell cycle control (ii) Anti-cancer drugs old versus new (iii) Side effects (iv) Multi-drug resistance

Teaching is student driven based on the various CATs. Teaching emphasis is placed on key pharmacological aspects. Students learn core knowledge initially followed by its application to applied pharmacology.

Learning: The focus of the RS was to aid student learning by giving students the opportunity to revise pharmacological material a few days after the original lecture material was delivered when information and knowledge was still recent. This offered the students an opportunity to correct any misconceptions in a timely manner. The student/teacher interactive style used for the RS, as opposed to a teacher led didactic lecture aided student learning. Students learned to focus initially on pathway umbrella terms and then to concentrate on specific terms (Table 2). They saw the importance of learning pharmacology using the drug class prototype model initially in order to be able to apply their knowledge afterwards to specific drugs (Table 2). Students also learned how to identify important drug-drug interactions, learned how cancer arises and how drugs can tackle this global health issue (Table 2).

It was extremely important during the RS to keep to the forefront the knowledge based requirements of the different student cohorts from various disciplines. Thus, careful strategic planning of the teaching of pharmacological drug aspects was necessary for it to be applicable to chemistry, biochemistry, physiology, medical and visiting Erasmus students. The teaching topics utilized during the RS are outlined in Table 3. These included teaching students the importance of chemical structures in predicting structure/activity relationships and how to interpret pharmacological data in terms of drug-target pharmacodynamics (chemistry students), application of biochemical pathways and molecular pharmacology to understanding how drugs work (biochemistry students), the close association between physiology/pathophysiology and pharmacology and how some common drugs relate to these processes (physiology students), clinical application of basic pharmacology and how the mechanisms of drug actions relate to the

management of specific diseases (medical students) and focusing on core pharmacological aspects (Erasmus students). The structured RS permitted more discipline driven teaching of pharmacology that was not always possible in the lectures. This was mainly driven by the fact that more time could be devoted to individual discipline related topics. Topics requiring further revision or clarification were clearly outlined by students' CATs and so could be specifically targeted to individual disciplines. Also, a higher level of student/teacher interactive discussion created an environment of a questions and answer session specifically targeted to individual student disciplines and requirements.

Table 3

Teaching strategies employed per discipline in each of the three review sessions

Individual disciplines	**Topics to improve student learning**
Chemistry	Structure/activity relationships
Biochemistry	Biochemical and molecular emphasis
Physiology	Physiological and pathophysiological emphasis
Medicine	Clinical emphasis using real-world clinical situations
Erasmus	General pharmacology

Teaching in each of the three RS was tailored where possible to each of the five different student disciplines in the class.

MCQ examination

The objectives of the three MCQ examinations were to test students' core as well as applied knowledge in (i) inflammation, (ii) anti-microbials and (iii) cancer. Analysis of individual student pre- and post-RS MCQs readily indicated whether or not there was an improvement in student learning. There was a significant increase in mean MCQ results between the pre- and post-RS MCQ examinations for each of the three RS (Figure 1). Greatest improvement in overall student performances was observed in RS2 ($p^{***} < 0.001$) with similar values observed in RS1 ($p^{**} < 0.01$) and RS3 ($p^{**} < 0.01$). A similar significant increase in student learning was observed for all five disciplines in RS1-3 (data not shown).

On closer analysis of the data, it was interesting to observe that there was a significant improvement in student learning in all areas reviewed in the three RS i.e. cellular pathways, pharmacodynamics (PD), cancer and multi-drug resistance (Figure 2). In fact, the data permitted identification of individual questions and thus pharmacological topics that students from the different disciplines demonstrated the largest as well as the smallest improvement in learning (data not shown). For instance understanding of pro- and anti-inflammatory pathways (in RS1, in particular peroxisome proliferator-activated receptors pathway) and their overall pharmacological regulation remained challenging for the physiology cohort within the class. This was not evident in the biochemistry group of students for example. This may be explained by the fact that biochemistry students may have a stronger background in a range of biological and biochemical pathways (within the human body) taught in different modules within their

discipline. Moreover, the physiology group of students found understanding the genesis of cancer and its pharmacological control more difficult than all other groups. While

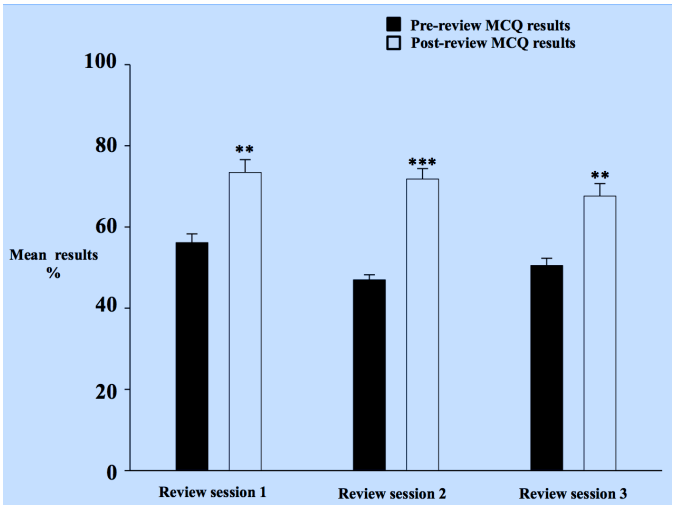


Figure 1. Students increase their mean MCQ examination results following review sessions. *Mean class MCQ examination results in pre- and post-review MCQs carried out in review sessions 1-3. The results shown are mean ± S.E. of the entire class performance. Significant (**, p<0.01; ***,p<0.001) changes from pre-review MCQ results.*

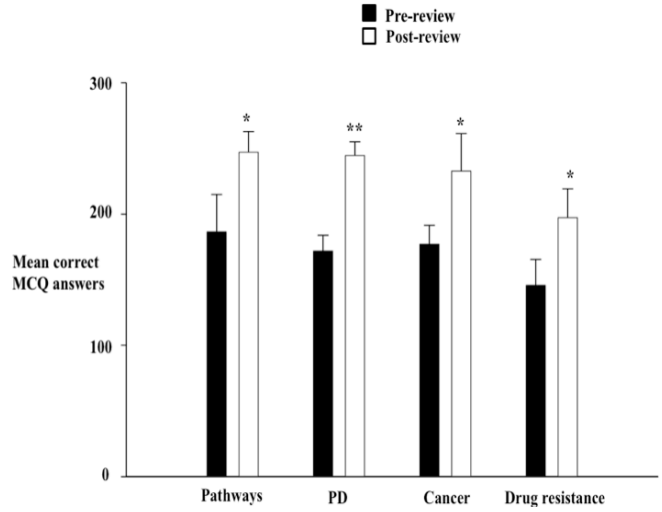


Figure 2. Students demonstrate improved learning in different pharmacological topics. *The mean correct MCQ examination results in pre- and post-review MCQs for the entire class are shown for five independent pharmacological topics. The results shown are mean ± S.E. of the entire class performance. Significant (*, p<0.05; **, p<0.01; ***,p<0.001) changes from pre-review MCQ results. PD; pharmacodynamics.*

analysis of the pre- and post-MCQs demonstrated a significant increase in their understanding in this topic the MCQ mean results for the physiology group were significantly lower (p**< 0.01) than all other groups. It was also interesting to observe that all groups except the visiting

Erasmus student group demonstrated a significant increase in their understanding of drug pharmacodyamics across a range of drug classes in all three RS (data not shown). This was not evident in the Erasmus group as they performed particularly well in this area of pharmacology in both their pre- and post-review MCQ examinations. This may be related to the fact that 80% of the visiting Erasmus student group was undertaking a pharmacy degree in their home country whereupon they may have previously undertaken a pharmacology module covering the pharmacodymanics of commonly used drugs. Lastly, of note it was the area of drug resistance in particular that proved most difficult for all student groups as evident from the lower mean correct MCQ answers in both the pre- and post-review MCQ (Figure 2). This was somewhat surprising as students are taught that in general drug resistance is related to drug pharmacodymanics. Once knowledge and understanding of a drug's mechanism of action is obtained, then by inference, so too is the mode of resistance associated with that drug. Overall, however, it was encouraging to see that there was a significant improvement in student learning in terms of core as well as applied pharmacological knowledge in all three RS (Figure 3).

Similar to the RS, the MCQ examinations provided a teaching and learning opportunity. In terms of teaching, there was an opportunity to ask students multiple targeted questions which is not possible in a traditional lecture. It also allowed for timely assessment of students' knowledge and understanding. From the student perspective the MCQs helped focus students' attention throughout the RS. They helped connect topics together which were specifically outlined by the students themselves and not the teacher. The MCQs provided opportunities for student self-assessment, a chance for the students to practice dealing with typical MCQs prior to the EOM MCQ examination and demonstrated to them the level of core and applied knowledge that was required in different pharmacological areas. Overall the MCQs enhanced student engagement and specifically guided student test taking logic.

Questionnaire

To investigate the effects of the various strategies used in this study student feedback was obtained using a written questionnaire covering various aspects of teaching, learning and assessment methods. Active participation was evident in all student disciplines (chemistry (95%), biochemistry (100%), physiology (57%), medicine (84%) and Erasmus (67%)). Of note least participation occurred in the physiology group which was previously observed in the RS (65.10 ± 1.6%). The first half of the questionnaire provided frequency rates for each statement in the survey. Frequencies were analysed per discipline (data not shown) as well as per total class student group (Table 4). Ninety three percent of the class "strongly agreed" and "agreed" with the first statement. In support of this statement one student commented (in the second half of the questionnaire) that "*I concentrated more in the lectures knowing that I had to complete a CAT at the end outlining areas I understood well and areas of misunderstanding*". The next statement "*CATs became monotonous after a while*" revealed the largest variation in student responses (Table 4). Interestingly, the medicine group (67%) provided the highest "strongly agreed" and "agreed" responses with this statement. In contrast all other disciplines demonstrated lower percentages (chemistry: 28%, biochemistry: 41%, physiology: 58% and Erasmus: 25%) (data not shown). The positive response from students for the next two statements (100% and 94% for statements 3 and 4 respectively) clearly indicated that the RS provided students with the opportunity to review lecture material which was directly focused on topics that students found difficult. Careful planning of the RS and associated MCQ examination questions proved fruitful

as 94% of respondents indicated that the RS were closely aligned with topics students found difficult while 92% indicated that the MCQ questions were at an appropriate level (statement 5, Table 4). Ninety percent of students welcomed active participation in their own assessment as a means to improving their overall leaning and understanding (statement 6, Table 4). Finally the majority of respondents (99%) agreed that the RS enhanced their learning outlining the positive response of students to the RS as a tool for learning (Table 4).

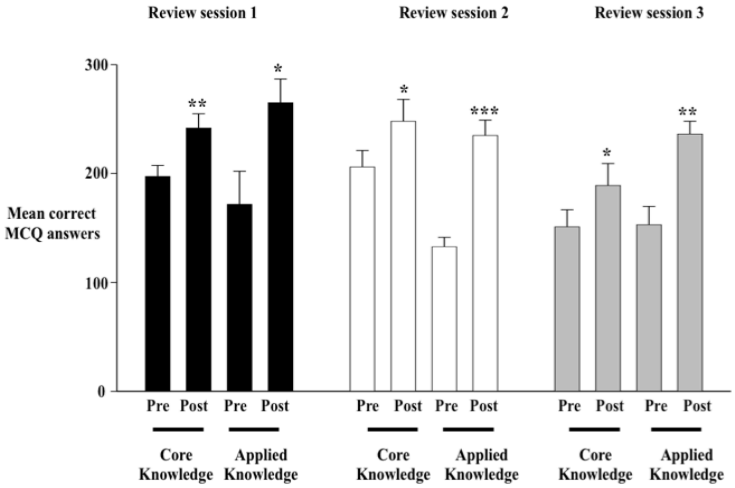


Figure 3. Students demonstrate improved learning in core as well as applied knowledge in pharmacology. *The mean correct MCQ examination results in pre- and post-review MCQs for the entire class in review sessions 1-3. Results are shown for core as well as applied knowledge in each of the review sessions. The results shown are mean ± S.E. of entire class performance. Significant (*, p<0.05; **, p<0.01; ***,p<0.001) changes from pre-review MCQ results.*

On a personal level, the second half of the questionnaire provided important student responses indicating the positive and negative aspects of the RS as a novel teaching, learning and assessment technique. Analysis of the qualitative student feedback allowed for eight distinct categories to be devised. The positive aspects of the RS can be summarised as follows; (i) they assisted student understanding (90%), (ii) they allowed for student self-assessment questions (82%), (vi) they were conducted in a positive relaxed atmosphere (23%), (vii) the topics covered were concise and student focused (86%) and finally (viii) they provided structure to the pharmacological topics covered in the module (31%).

Similarly, the negative student responses regarding the RS can be summarised into another eight distinct categories; (i) more time should be devoted to the RS (27%), (ii) the absence of appropriate student feedback in terms of providing correct MCQ answers as well as individual student performances (6%), (iii) the type of RS with the associated MCQ type examination as opposed to essay style or laboratory specific examinations (2%), (iv) monotony of the CATs (47%), (v) class size of the RS i.e the entire class instead of smaller individual groupings (4%), (vi) use of similar lecture slides in the RS vs didactic lectures (35%), (vii) large number of drug names (14%) and finally (viii) omission of topics identified as difficult in the RS (6%). It is important to note that while eight separate headings could be assigned to both the positive and negative student responses they were not equally weighted. While some of the negative responses can be readily addressed, others, however, are directly related to resource

issues which are more difficult to tackle and could continue to have considerable controlling or restrictive implications. However, it was apparent that the extent of positivity highlighted with respect to the use of RS as a non-traditional teaching, learning and assessment method clearly overwhelmingly dominated the overall findings of the student questionnaire.

Table 4

Student responses to review session questionnaire statements

N=72 (81% of the entire class)					
	SA	**A**	**N**	**D**	**SD**
The CATs were useful to outline difficult as well as easy topics	37	30	4	1	0
CATs became monotonous after a while	4	30	22	13	3
The review sessions were a good way of reviewing lecture material	68	4	0	0	0
The review sessions were targeted to topics students found difficult	57	11	4	0	0
The MCQs in the pre- and post-review sessions were at an appropriate level	33	33	5	1	0
Participation in your own assessment i.e. student CATs followed by review sessions helped your learning and increased your understanding	32	33	7	0	0
The review sessions in general enhanced your learning	53	18	1	0	0

Questionnaires were completed in-class in a relaxed and non-timed fashion. Students were asked to give their opinion on seven different statements by answering SA; strongly agree, A; agree, N; neutral; D; disagree, SD; strongly disagree.

Discussion

Pharmacology has long been recognized as a formal discipline in scientific medicine (Flexner, 1910). While there it holds its rightful place, it also seamlessly integrates into many interdisciplinary areas of basic and clinical sciences. In light of this fact it is not unusual to discover a number of distinct student groups from different disciplines in the one pharmacology class. For this particular study one of the key features that influenced the class choice was its multidisciplinary nature. While the module used in this study like all others in the discipline aims to teach the undergraduate student the principles of pharmacology along with the process of pharmacologic reasoning, the key goal of the intervention was to improve teaching and learning for multidisciplinary students. Chemistry students were taught to appreciate the structural diversity of pharmacological drugs, their chemical characteristics and applications. The biochemistry students were taught how drugs regulate various metabolic pathways and how important sequence data is to drug function. The physiology students were taught how drugs relate to physiological and pathophysiological processes in the human body and lastly, the medical students were taught the pharmacological management of selected major diseases. In contrast to the lectures the RS were student designed and targeted which encouraged more interactive student and teacher engagement. They were employed to continually engage students while keeping the information succinct and focusing on essential principles necessary for each discipline.

One of the important findings of this study is that, despite the differences in student disciplines and backgrounds they succeeded in using the RS to their advantage and improving their understanding and knowledge of both core and applied pharmacological concepts. Student performances significantly improved following participation in the RS as evident from analysis of pre- versus post-review MCQ examination results for each individual student. This enhancement was both discipline- as well as topic-independent as each of the RS readdressed different areas of specialisation in pharmacology. However, overall mean correct MCQ examination results did vary somewhat between the various groups in all three RS. This is not surprising given the diverse pharmacological topics and concepts covered and examined in this module and the associated diversity of the five distinct student groups. In addition the vertical and horizontal integration of the different curricula within each discipline adds another layer of diversity within individual groups. Overall vigilant preparation of appropriate MCQ examination questions as well as how topics were strategically re-addressed was extremely important. While this was somewhat challenging proper design and implementation of the RS was essential for them to be utilised as an effective teaching, learning and assessment tool.

Student feedback demonstrated that the RS were appreciated as a welcome change to traditional didactic lectures as outlined by the following two student responses "*I consider the RS and their associated MCQ examinations very useful compared with endless lectures to better understand pharmacology*" and "*I really enjoyed the RS, an unusual change from regular lectures*". They helped cement and reinforce essential topics in pharmacology in a relaxed and motivating environment conducive to learning. Student feedback provided important insight into the future design improvement of the RS. It is very encouraging that students suggested that the use of this technique would be beneficial in other modules both in related and unrelated disciplines. It was gratifying to observe the high level of attendance and engagement by students in all steps of the RS despite the absence of attainable marks which so often provides the only incentive for students' participation (Markham et al., 1998).

It is extremely important that recognition and appreciation of effective non-traditional teaching methods in pharmacology come from colleagues, students and institutions (Markham et al., 1998; Desai, 2009; Joshi & Trivedi, 2010). Teachers of pharmacology are aware of non-traditional methods in teaching the discipline but lack of student and collegial motivation, lack of recognition by institutes, shortage of resources and the overall time involved in these processes lead some teachers to conclude that perhaps these alternative methods of teaching are a waste of time (Markham et al., 1998; Desai, 2009; Joshi & Trivedi, 2010). More worryingly, a higher value can often times be accorded to conducting pharmacology research and consultancy activities rather than implementing and promoting effective teaching of pharmacology by focusing on teaching, learning and assessment. Thus, it is timely to not only seriously address the significance and value of researching teaching in pharmacology but to re-enforce its importance and equitable place in contributing to the effective functioning of the discipline.

In summation, we can conclude that RS are a successful and novel sole-instructor teaching and learning tool in pharmacology engaging large numbers of students from multidisciplinary backgrounds. This conclusion is based on comparing student performances, EOM student evaluations of the RS and student satisfaction with what they have learned and accomplished in the module. The success of our intervention is somewhat multifaceted. It lies in the active student participation in all elements of the RS which may be partially aligned to their novelty. The student-directed topics meant that the RS strategically focused on pharmacological elements directly related to each student discipline and thus readily pertinent. Overall the

teaching and learning was different to didactic lectures in terms of being focused, relevant, interactive and relaxed. Some challenges along the way, however, cannot be ignored. Some students thought the CATs became monotonous after a period of time. The many hours devoted to the various steps associated with the RS: (i) analyzing each individual student CAT from multiple lectures, (ii) the design of the review topics, (iii) the design of appropriate MCQ examination questions, (iv) analyzing each individual student pre- and post-review MCQ examination results in each of the three RS and (v) interpretation of the resultant data. There is the omission of topics from the RS proffered by individuals but which could not be addressed due to time constraints. Finally, as outlined previously a significant proportion of the class completed the various CATs, attended each of the three RS and completed the EOM questionnaire. It must be noted, however, that 100% student attendance was never achieved in any of the outlined activities. In addition there was variation in the numbers of students attending from the different disciplines.

Despite the limitations outlined above it is our hope that this study will be a starting point for continuous experimentation and optimization of the use of novel periodic RS in the teaching of pharmacology to undergraduate students. In fact armed with the knowledge of our research limitations we are presently implementing an amended version of the research outlined in this report. One way of overcoming the limitation of not reaching full student attendance and participation is to reward students and have it aligned with their overall continuous assessment (CA). Thus, for each RS attended and post-review MCQ completed students are currently awarded a mark that is carried forward as part of their module overall CA. As a result of this new intervention we have now achieved 100% student attendance and participation. Another change currently being imposed is the extension of the RS to an hour and a half to encompass more topics outlined by individual students. It is our desire that other limitations will continue to be ironed out with future implementation of our research activities.

Lastly, the flexibility of our RS model will permit prospective researchers to not only replicate our process but importantly tailor it to their own needs and specifications. Our current approach may be modified and subsequently maintained through varying the one minute paper format and/or timing (Simpson-Beck, 2011; Stead, 2012) as well as the number of CATs utilised, completing CATs per student discipline group as opposed to individual student feedback, reducing the number of RS if necessary and implementing a more automated system for analysis of student feedback in general where appropriate. Thus, awareness of the associated limitations of the existing RS model and some potential ways of overcoming them will assist further in the effective implementation of our research activities. In addition future research should not only further test but also extend our findings to go beyond the student disciplines we have outlined in this study. Importantly our research activities were positively linked with increased teacher job satisfaction and teachers' self-efficacy (data not shown). This is despite the increased workload associated with our research activities (which may be reduced in the future) and the possibility of lack of recognition for accomplishments (Greenglass & Burke, 2003; Klassen et al., 2010). However, the positive engagement by a high percentage of the students in all aspects of the RS, good teacher-pupil rapport and overall success in improving student learning were directly linked to achieving high levels of personal satisfaction from the process. These are important findings as overall job satisfaction can in turn influence students' motivation and achievement (Klassen et al., 2010; Skaalvik & Skaalvik, 2007). Our study clearly holds implications for improving teaching and student learning. Implementing our existing activities or an abridged adaptation in

future studies might provide more nuanced understandings of its effectiveness beyond the discipline of pharmacology.

Acknowledgements

The authors would like to acknowledge the support and assistance of the science, medicine, and Erasmus visiting students who participated in our research.

References

Achike, F. I., & Ogle, C. W. (2000). Information overload in the teaching of pharmacology. *Journal of Clinical Pharmacology, 40*(2), 177-183. doi: 10.1177/00912700022008838

Angelo, T. (1991a). Introduction and overview: From classroom assessment to classroom research. *New Directions for Teaching and Learning*, *46*, 7-15. doi: 10.1002/tl.37219914603

Angelo T. (1991b). Ten easy pieces: Assessing higher learning in four dimensions. *New Directions for Teaching and Learning*, *46*, 17-31. doi: 10.1002/tl.37219914604

Angelo T. (1998). Classroom assessment and research: An update on uses, approaches, and research findings. *New Directions for Teaching and Learning*, *75*, 1-2. doi: 10.1002/tl.7500

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques*, 2nd ed., San Francisco: Jossey-Bass.

Banerjee, I., Roy, B., Sathian, B., Pugazhandhi, B., Saha, A., & Banerjee, I. (2013). Teaching aids in pharmacology teaching and learning methodology: A study from a medical college in Nepal. *International Journal Interdisciplinary and Multidisciplinary Studies, 1*(1), 1-7.

Bligh, D. A. (2000). *What's the use of lectures?* San Francisco: Jossey-Bass.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W., & Krathwohl, D. (1956). Taxonomy of educational objectives. *The Cognative Domain*. New York: McKay.

Brewer, G., & Hiscock, D. (2001). Medical education and practice in the information age. *Postgraduate Medical Journal*, *77*(909), 425-427. doi: 10.1136/pmj.77.909.425

Brown, W. A., & Gamber, C. (2002). *Cost containment in higher education – Issues and recommendations, ASHE-ERIC Higher Education Report. 28*(5).

Byon, A. S. (2005). Classroom assessment tools and students' affective stances: KFL classroom settings. *Language and Education*, *19*(3), 173-93. doi: 10.1080/0950070508668673

Coleman, I. P. L., Foster, R. W., Hollingsworth, M., Morgan, R., Sewell, K., & Walker, J. (1995). Drug targets and transduction systems. *British Journal of Pharmacology, 115*, 159.

Cottell, P., & Harwood E. (1998). Do classroom assessment techniques (CATs) improve student learning? *New Directions for Teaching and Learning, 75*, 37-46. doi: 10.1002/tl.7504

Cronin Jones, L. L. (2003). Are lectures a thing of the past? *Journal of College Science Teaching*, *32*(7), 453-457.

Darling-Hammond, L. (Ed.). (2008). *Powerful learning: What we know about teaching for understanding*. San Francisco: Jossey-Bass.

Desai, M. (2009). Changing face of pharmacology practicals for medical undergraduates. *Indian Journal of Pharmacology, 41*(4), 151-152. doi: 10.4103/0253-7613.56062

Dewhurst, D. G., & Norris T. A. M. (2003). Helping teachers to embed e-learning materials into undergraduate pharmacology courses. *Bioscience Education*, *1*(1).

Ebert-May, D., Brewer, M., & Allred, S. (1997). Innovation in large lectures-teaching for active learning. *Bioscience*, *47*, 601-607. doi: 10.2307/1313166

Flexner, A. (1910). *Medical education in the United States and Canada. A report to the Carnegie Foundation for the advancement of teaching. Bulletin 4*. Boston: Updyke.

Fry, H., Ketteridge, S., & Marshall, S. (1999). *A handbook for teaching and learning in higher education: Enhancing academic practice*. Philadelphia, PA: Kogan Page Publishers.

Gardner, H. (1995). Reflections on multiple intelligences myths and messages. *Phi Delta Kappan, 77*(3), 200-209.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st Century*. New York: Basic Books.

Gardner, H., & Boix Mansilla, V. (1994). Teaching for understanding within and across the disciplines. *Educational Leadership*, *51*, 14-18.

Gregson, K., Romito, L. M., & Garetto, L. P. (2010). Students' attitudes toward integrating problem-based learning into a D.D.S. pharmacology curriculum. *Journal of Dental Education*, *74*(5), 489-498.

Greenglass, E. R., & Burke, R. J. (2003). *Teacher stress: Occupational stress in the service proffessions*. New York: Taylor & Francis.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S. M., & Wood, W. B. (2004). Scientific teaching. *Science*, *304*(5670), 521-522. doi: 10.1126/science.1096022

Joshi, A., & Trivedi M. (2010). Innovations in pharmacology teaching. *International Journal of Pharmaceutical and Biomedical Research*, *1*(2), 62-64.

Klassen, R. M., Usher, E. L., & Bong, M. (2010). Teachers collective efficacy, job satisfaction and job stress in cross-cultural context. *The Journal of Experimental Education*, *78*, 464-486. doi: 10.1080/00220970903292975

Kwan, C. Y. (2002). Problem-based learning and teaching of medical pharmacology. *Naunyn-Schmiedebergs Archives in Pharmacology*, *366*, 10-17. doi: 10.1007/s00210-002-0561-y

Litke, R. A. (1995). Learning lessons from students: What they like most and least about large classes. *Journal on Excellence in College Teaching*, *6*(2), 113-129.

Lymn, J. S., & Mostyn, A. (2010). Audience response technology: Engaging and empowering non-medical prescribing students in pharmacology learning. *BMC Medical Education*, *10*, 73-83. doi: 10.1186/1472-6920-10-73

Markham, T., Jones, S. J., Hughes, I., & Sutcliffe, M. (1998). Survey of methods of teaching and learning in undergraduate pharmacology within UK higher education. *Trends in Pharmacological Science, 19*(7), 257-262. doi: 10.1016/S0165-6147(98)01221-8

McBane, S., & Mesaros, J. (2010). Teaching pharmacology in a physician assistant program. *Journal of Physician Assisted Education, 21*(3), 18-22.

McKeachie, W. J., Svinicki, M. D., & Hofer, B. K. (2006). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers*. Boston: Houghton Mifflin.

Morgan, J. H. C. (2003). New approaches to training general practitioners. *Postgraduate Medical Journal, 79*(930), 187-188. doi: 10.1136/pmj.79.930.187

Oppenheim, A. N. (1992). Questionnaire design, interviewing and attitude measurement. (2nd ed.) London: St. Martin's Press.

Rogers, K. M. A. (2012). A feasibility study on the development and integration of a teaching aid for pharmacology. *International Journal of Modern Education Forum, 1*(2), 53-61.

Rouseff-Baker, F., & Holm A. (2004). Engaging faculty and students in classroom assessment. *New Directions for Community* Colleges, *126*, 29-42. doi: 10.1002/cc.152

Routledge, P. A. (1999). The inter-face between clinical and laboratory pharmacology. *British Journal of Clinical Pharmacology, 47*(6), 611-612.

Sekhri, K. (2012). Teaching methodologies in pharmacology. A survey of students' perceptions and experiences. *Journal of Education and Ethics in Dentistry, 2*(1), 40-44. doi: 10.4103/0974-7761.115150

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*, 611-625. doi: 10.1037/0022-0663.99.3.611

Simpson-Beck, V. (2011). Assessing classroom assessment techniques. *Active Learning in Higher Education*, *12*(2), 125-132. doi: 10.1177/1469787411402482

Stead, D. R. (2012). A review of the one minute paper. *Active Learning in Higher Education, 6*(2), 118-131. doi: 10.1177/1469787405054237

Udovic, D., Morris, D., Dickman, A., Postlethwait, J., & Wetherwax, P. (2002). Workshop biology: Demonstrating the effectiveness of active learning in an introductory biology course. *BioScience, 52*(3), 272-281. doi: 10.1641/0006-3568(2002)052[0272:WBDTEO]2.0.CO;2

Wiecha, J. M. (2003). Collaborative online learning (COL): A new distance learning method. *Essential Drugs Monitor, 33*, 36.

Wellington, J. (2000). *Educational research: Contemporary issues and practical approaches*. London: Continuum.

Zgheib, N. K., Simaan, J. A., & Sabra, R. (2010). Using team-based learning to teach pharmacology to second year medical students improves student performance. *Medical Teaching, 32*(2), 130-135. doi: 10.3109/01421590903548521

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 57 - 69.
doi: 10.14434/josotl.v15i1.13000

Student motivation, intercultural competence and transnational higher education: Uzbekistan, a case study

EngKee Sia[1]

Abstract: This paper investigates some of the cross-cultural challenges faced by faculty members teaching transnational higher education in a foreign country. It employs the intercultural competence process model and attempts to provide some best practices that are already implemented in an international branch campus (IBC) in Uzbekistan. Hopefully, this sharing of practices will develop intercultural competence and better prepare transnational faculty members to be more efficient and effective in motivating students in transnational education programmes. Furthermore, apart from increased motivation for students, this important professional development initiative for faculty teaching staff may lead to improvement in learning outcomes over time.

Keywords: transnational teaching; professional development; intercultural competency; motivation

Introduction

With the advancement of worldwide network communications and technological innovations coupled with the strategic globalisation of higher education institutions (HEIs), the nature of international higher education is evolving. Students now have more choices in selecting foreign universities even within their home countries, either through attending classes at the international branch campuses (IBCs), collaborative public/private institutions, or via online platforms. This form of education is known as transnational higher education (THE) with the foreign universities referred to as transnational institutions, and the students, as transnational students. Indeed, the competition among universities in providing THE is increasingly intense. The consequence is that the nature of learning and teaching has also changed significantly. The availability of THE business opportunities attracts many new players including public and private, international and national, profit and not-for-profit organisations with varied alliances or partnerships that motivate innovative approaches to teaching and delivery.

Within the IBC perspective, foreign faculty members are usually sent from the home institution to teach students in the host country for a short period of time, which is known as transnational teaching (Smith, 2010). As such, block teaching by the fly-in fly-out (FIFO) faculty members is common. These transnational faculty members have a demanding schedule, since they must simultaneously manage their courses at the home institution while teaching intensive blocks of classes at the host country (McBurnie & Ziguras, 2007). The faculty members travel and teach on weekends, usually Fridays, Saturdays and Sundays, or for a full week made up of eight-hour teaching days. Within these 3 - 5 days of teaching, the students usually have to take annual leaves from their full-time job, concentrate and focus on absorbing the entire module contents taught by the flown in faculty members. In addition, the students also try to gather some

[1] International Management, Management Development Institute of Singapore, 501 Stirling Road, Singapore 148951, ekinsia@gmail.com, engkee_sia@mdis.edu.sg

focus areas for exam preparation as the flown in faculty member is usually the module leader who is responsible for setting the exam questions (Mok, 2012). However, transnational faculty members are not prepared by the home institutions to meet the challenge of the assignment (Leask, 2008) to teach culturally diverse students from the IBC, apart from formal intercultural competence training (Smith, 2010).

This paper attempts to address some of the cultural issues that transnational faculty members may encounter while teaching in the IBCs in Uzbekistan. It discusses some of the approaches that the foreign faculty members may adopt to teach cross-culturally at an IBC, thereby enhancing student motivation to learn in transnational education environment. By sharing best practices in the application of the intercultural competence process model (Deardorff, 2009), transnational faculty members may benefit by adopting the framework that focuses on three core elements - attitudes, knowledge and comprehension, and skills, so as to prepare them to teach in the international culturally diverse environment. Each components of the intercultural competence theoretical framework will be discussed with the support of real life issues occurring in an IBC of Uzbekistan. Some of the claims, suggestions and recommendations made for the IBC under study and presented in this paper are based on the author's ten years of teaching and managing experiences in the various THE provisions. The author has more than five years experience in managing the entire academic processes of the IBCs, one in Colombo, Sri Lanka and the other in Tashkent, Uzbekistan. Most of these claims and suggestions were raised and discussed during dialogue sessions with the transnational faculty members and local lecturers / tutors. Some good cross-cultural teaching methods were recommended and implemented with effective outcomes. The intercultural competence process model, together with the practical and real life issues of the IBC case study in support of the theoretical framework provide a new reference offering a comparative study for future research in the context of intercultural competence and transnational teaching.

Literature Review

The intercultural competence framework comprises attitudes, knowledge and skills (Deardorff, 2009). The essential attitudes include respect, openness, curiosity and discovery of other cultures. Openness and curiosity imply a willingness to risk and to move beyond one's comfort zone. In communicating respect to others, it is important to demonstrate that others are valued. The knowledge necessary consists of cultural self-awareness (meaning the ways in which one's culture has influenced one's identity and worldview), culture-specific knowledge, and deep cultural knowledge including understanding the world from others' perspectives. The skills are the ones that address the acquisition and processing of knowledge, i.e., observing, listening, evaluating, analysing, interpreting, and relating. These attitudes, knowledge and skills ideally lead to an internal outcome that consists of flexibility, adaptability, an ethnorelative perspective and empathy. At this point, individuals are able to see from others' perspectives and to respond to them according to the way in which the other person desires to be treated. Individuals may reach this outcome in varying degrees of success. Finally, the summation of the attitudes, knowledge and skills as well as the internal outcomes are demonstrated through the behaviour and communication of the individual, which become the visible external outcomes of intercultural competence experienced by others. These five overall elements can be visualized through the model of intercultural competence, as illustrated in Figure 1, thereby providing a framework to further guide efforts in developing intercultural competence.

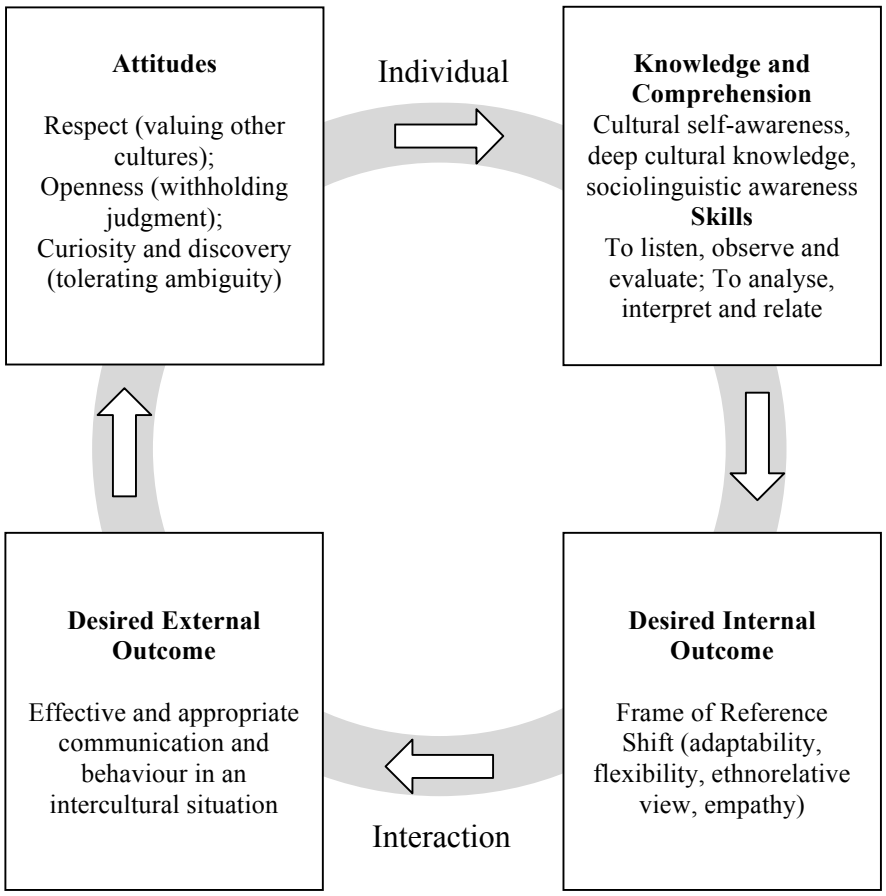


Figure 1: Model of Intercultural Competence (Source: Deardorff, 2009)

An educated and advanced Uzbekistan

In Uzbekistan, the percentage of population with higher education is 9.8 percent (Nessipbayeva & Dalayeva, 2013). This percentage is the lowest among the Central Asia republics. Higher education reforms in Uzbekistan started in 1997 with the adoption of the Education Act and the National Programme for Personnel Training (NPPT). The aim of these initiatives was to increase the percentage of the population with higher education qualifications and to train highly qualified specialists to the equivalent level in advanced or developed states. At present, there are 76 HEIs in Uzbekistan, including 11 joint higher education establishments (EACEA, 2012). These institutions are all legal entities and there are no non-government universities in Uzbekistan (EACEA, 2012).

When the Soviet Union collapsed in 1991, higher education reforms in Uzbekistan moved toward standardised university entrance tests as a criterion for admission. There was also a restructuring away from sectoral ministerial control, encouraged diversification of education provision as well as decentralisation of governance, salary, and tuition structures (Heyneman,

2011; Silova & Steiner-Khamsi, 2008). The implementation of the unified education policy is governed by the Ministry of Higher and Secondary Specialised Education (MHSSE), which is responsible for the development of the education sector as a whole, including the implementation of education reforms (EACEA, 2012). Generally, the changes were perceived as necessary to enhance the Soviet higher education system embedded in Uzbekistan, while upgrading the system to benchmark against international higher education requirements.

Russian influences on Uzbekistan education

Because Uzbekistan and Russia retain historical, social and economic relationships in the post-Soviet era, the influence on higher education still remains. However, the European Union (EU), especially the Bologna process, has increasingly played an important role in influencing the direction of higher education reforms in Central Asia. Due to the context of these overlapping international influences, the reconfiguration of post-Soviet higher education system in Uzbekistan may result in a hybrid of the East and West model of education system (Silova, 2011).

In an attempt to create its own model of a hybrid system, Uzbekistan has based its higher education reforms primarily on internal references to Soviet educational practices (Tomusk, 2008). Indeed, Russia has established three IBCs of its universities in Uzbekistan to respond to the demand for higher education in the Russian language, particularly the Moscow State University of Lomonosovt, Plekhanov Russian University of Economics and the Russian State Oil and Gas University of Gubkin, all in Tashkent. In this context, Russian based IBCs remain the first choice of HEIs for many Uzbek students, although an increasing number of students, especially those with English language abilities, choose to study in the Western based IBCs. These branch campuses undoubtedly influence higher education in the country and the nearby region by diversifying the available study options and increasing academic competition.

The Bologna Process

In spite of the deep-rooted Russian influence, the higher education reforms have increasingly diverted such influence due to the Bologna process, which has become a major reference point for Uzbekistan in adopting the European education system. Intensive cooperation between the EU and Uzbekistan began in 2007, when the European Education Initiative was launched, as part of the EU-Central Asia Strategy (Jones, 2011). By 2009, the initiative had prioritised higher education and emphasised links with the Bologna process. Since then, international cooperation among HEIs in Uzbekistan has been strong and there are a number of bilateral agreements with foreign universities from 45 countries throughout the world (EACEA, 2012) to foster greater exchange of academic experiences. Through Tempus and Erasmus Mundus External Cooperation Programmes, Uzbek universities have been collaborating with many universities in the EU, with MHSSE maintaining close contacts with the diplomatic missions located in Tashkent. Working relationships have also been established with international organisations, such as the Japan International Cooperation Agency (JICA), German Academic Exchange Programme (DAAD), the British Council, Korean International Cooperation Agency (KOICA), and UNESCO, in order to foster more fruitful international HEI cooperation (EACEA, 2012).

To advance the process of National Programme for Personnel Training (NPPT) in

Uzbekistan, international higher education cooperation was identified as the highest priority to achieving the objectives. The mechanisms for international cooperation can be in the form of

- international branch campuses (IBCs),
- academic collaboration projects involving foreign lecturers in the teaching at Uzbek universities,
- joint research work with foreign universities, and
- international conferences on current issues like world economics, business trends, science and technology innovations, as well as resources and energy saving.

Large-scale international cooperation, especially the establishment of IBCs and academic collaboration projects, has successfully enhanced the quality of teaching processes among HEIs in Uzbekistan. At present, there are four IBCs, with English as the medium of instruction, established in Tashkent, Uzbekistan, namely the UK Westminster International University in Tashkent (WIUT), Italian Turin Tashkent Polytechnic University (TTPU), South Korea Inha University in Tashkent (IUT), and the Management Development Institute of Singapore in Tashkent (MDIST). These universities were invited to establish their IBCs in Tashkent to assist in this meaningful initiative, share their best education management practices, and provide internationally recognised higher education academic programmes to the tertiary students in Uzbekistan.

Guidelines on educational standards

Guidelines on Educational Standards were developed by MHSSE to define the academic standards including the requirements for the students to attain undergraduate and postgraduate qualifications, for the academic staff to develop their capabilities in strengthening the assessment processes, as well as for the universities to build their capacities in attainting greater accreditation status. It also includes the provisions of collaborations with foreign HEIs in terms of academic teaching and contemporary research work on science and technology, as well as integration of education with commercial and industrial experiences.

Beside the Educational Standards, other priorities in education reform are for domestic universities to

- enhance the teaching of foreign languages, including English, so as to enable students to have better access to information resources,
- adopt new communication technologies and modern pedagogical methodologies in classroom teaching,
- implement educational computing networks for staff and students to access reference materials,
- upgrade the skills of local academic staff through sharing / training sessions with foreign universities faculties,
- develop relevant / updated teaching materials through demand surveys in commercial / industrial markets,
- implement quality assurance / management system, and
- optimise the resource facilities including the construction, renovation, refurbishment of academic buildings, scientific research laboratories, sports centres and student dormitories.

For IBC initiative, Uzbekistan has provided all the necessary conditions to enable a smoother transition to support this initiative, including higher education reform, guidelines on educational

standards, legal advice, qualified personnel, scientific-methodological research facilities, as well as financial arrangements..

Guidelines on intercultural competence

Guidelines on intercultural awareness for teaching in transnational institutions have been generated by the Global Alliance for Transnational Education (GATE) (Greenholz, 2000). Nonetheless, the studies and monitoring on transnational teaching in cross-cultural environments have been lacking (Gribble & Ziguras, 2003) despite that transnational faculty members are the primary facilitators of students' learning (Johnson, 2003).

Attitude

Developing intercultural understanding begins with the attitude of the faculty members (Crabtree & Sapp, 2004) to recognise and respect the value of other cultures (Deardorff, 2009). The ability to motivate oneself to be open to other cultures can strengthen intercultural adaptability (Spitzberg & Changnon, 2009), and suspending ethnocentric (Storti, 2009), assumptions (Dunn & Wallace, 2006) and judgments (Bennett et al., 2003) allows faculty members to be receptive to multiple perspectives.

Knowledge

Knowledge and comprehension is the second core element in the dynamic process of developing intercultural competence. Some societies still hold strong gender stereotypes that are deeply embedded in their cultures (Merriam, 2007). For example, it is common for male students in the IBC under study to approach and shake hands with male transnational faculty members however; this practice is prohibited to all female students. Furthermore, people in Uzbekistan get married in their early 20s, so it is common to see pregnant students in the lecture halls and classrooms. Therefore, certain adjustments may have to be made with respect to teaching spaces and to have more frequent breaks.

Skills

Skills are the third core element in the process of developing intercultural competence. A core element of skills development is self-reflection, which comprises three levels that facilitate cultural transformation, namely content, process and premise reflections (Mezirow, 1998). Content reflection refers to the analysis of roles and the adjustment of relationship between faculty members and students (Gopal, 2011). For instance, in the IBC under study, transnational faculty members stay in the same hostel as the students. Beside the formal lectures and class tutorials, students are able to meet teaching staff in the hostel study room where they mentor students during their free time. Despite staying in the same hostel and providing voluntary coaching, foreign faculty members are required to discipline themselves and to keep a professional distance from the students in order to prevent further development of closer social relationships with them.

As can be seen from the content reflection above, self-reflection involves a subjective understanding of reality that has the value of enabling transnational faculty members to think

about the effect of their actions on the students (Gray, 2007). Process reflection involves the analysis of the situation and the adjustment of actions between faculty members and students (Mezirow, 1998). In Uzbekistan, transnational faculty members may find students sitting at the back of the class not paying attention to teaching. They are either sleeping, talking with fellow classmates, messaging on their phones, playing computer games on their laptops or making creative drawings on their textbooks or study materials. This is a common classroom culture where students are not attentive in class and transnational faculty members may find it difficult to negotiate cross-cultural adjustment (Smith, 2010). They cannot ask the students to leave the classroom as lectures and tutorials require compulsory student attendance. Some transnational faculty members are able to stimulate the students' right brain by encouraging them to draw mindmaps on the topics that they have just learnt. Some foreign faculty members have suggested removing compulsory class attendance; however, if this system is to be implemented, it is likely that more than fifty percent of students will not attend class.

Premise reflection pertains to the analysis of perception and the adjustment of perspectives between faculty members and students (Mezirow, 1998). Many transnational faculty members perceive that students in the IBC are not working hard and usually study two days before the exams. However, a casual chat with students will reveal that they are actually working full-time and studying part-time in order to pay their school fees and accommodation as most of the students travel to university from outside of Tashkent. As soon as the perception is clarified, transnational faculty members would be more willing to provide mentoring to these students during their free time. Perhaps work and lack of sleep accounts for the reason why students are not attentive in class. Furthermore, if non-compulsory student attendance is implemented, more than fifty percent of the students will not attend class, as they need to work in order to pay for their school fees.

Reflexivity is the ability to constantly reflect on the significance of experience (Greenholz, 2000) and improve one's interaction with others. When we are being reflexive, we are aware of the ways in which our interpretations and actions are influenced by others; we become conscious of the rules that guide our context, and are able to explore other contexts and rules for interpreting an action in a situation (Littlejohn & Domenici, 2007). These three points can be illustrated by taking the example of examination process in the IBC under study. During an examination, the guidelines state that students are not permitted to leave the exam hall fifteen minutes before the end of the exam. Nonetheless, when transnational invigilators notice that the amount of talking and cheating is the greatest during the last fifteen minutes and are unable to prevent the talking, a change needs to be implemented. One approach is to modify the exam guidelines and allow those students who have submitted exam scripts in the last fifteen minutes to leave the exam hall, in order to minimise talking and cheating. There are rules and regulations that guide invigilators in the examination process, however, due to the customs that the students have been brought up since their younger school days, these guidelines may have to be modified to suit the local culture. Also due to their full-time jobs, which affect their preparation for the exam, students may resort to talking and cheating during an exam.

Effective communication skills are another key component in developing intercultural competence; such skills will lead to meaningful dialogue and eliminate miscommunication (Hannigan, 1990). Dialogue and feedback sessions are organised between the management and teaching staff every semester in the IBC to share teaching experiences and approaches to handling challenging students. Transnational faculty members are able to voice their concerns during the sharing session, so that the management is able to take timely corrective action, such

as the formation of the disciplinary committee to take actions against students who have difficulty engaging with staff. When transnational faculty members confronted students not paying attention in class, this has resulted in that students behaving impolitely towards them. Instances such as these can be avoided when foreign faculty members understand the students' background through open dialogue and communicating effectively with them.

Transnational faculty members also need to develop active listening skills that require suspending their own judgment and asking clarifying questions (Littlejohn & Domenici, 2007). Never insist on the management changing the culture of the transnational institution; instead offer some constructive suggestions and ideas to modify specific parts of the operational processes, such as removing the compulsory attendance for lectures, arranging all lectures in the morning and tutorials in the afternoon, if you wish to improve student engagement in class. Indeed, it takes time to be able to change the deep-rooted culture in an education institution and the priority of change has to come from the management systems.

Adding the necessary cultural knowledge and skills are to ensure that transnational faculty members can be more effective and appropriate in their intercultural interactions. Transnational faculty members can never become completely inter-culturally competent, but the most important is in the development process, i.e., how they acquire the necessary knowledge, skills, and attitudes. Therefore, critical reflection becomes a powerful tool in the process of intercultural competence development. In addition, intercultural competence must be intentionally addressed through programmes, orientations, experiences, and courses, which are essential to be a global-ready transnational faculty member. The framework/model of intercultural competence discussed with support from the case study helps to guide our efforts in ensuring a more comprehensive and integrated approach in enhancing transnational teaching thereby improving student learning outcomes.

Transnational teaching and student motivation

Attribution theory suggests that successful learning outcomes are commonly attributed to the quality of the teacher, instead of the motivation and effort of the learner (Weiner, 1974). This theory might be more applicable to teacher-centred learning in the pre-tertiary education where students are not ready to apply their knowledge that they have learnt. In the tertiary education however, academic staff act as facilitators to enable student-centred learning and encourage independent research and study. Therefore, student self-motivation to acquire academic knowledge and practical skills is a key determinant of their academic attainment, and it is the task of the academic staff to play a contributing role in stimulating student motivation.

Many students (and parents) expect that they will be taught by foreign faculty members from the home institution when they enrol in a transnational academic programme. In the IBC under study, some of the foreign faculty members are contracted from countries like India, Malaysia, the Philippines, Iran, etc. In addition, some lecturers/tutors are engaged locally often on a part-time basis, which may be an effective method of reducing costs, but less effective method of achieving employee commitment and higher levels of teaching engagement (Wilkins, 2010). Nonetheless, local academic staff often have a richer understanding of student needs and as a result, are able to manage and control class discipline. With a good mix of foreign faculty members and local lecturers/tutors, at the end of the module, students will be able to evaluate their lecturers/tutors based on the ability to stimulate their motivation and engagement,

Students of the IBC under study are mainly Uzbek with only a handful of foreign

students from nearby regions such as Ukraine, Afghanistan, Kyrgyzstan, Tajikistan, Kazakhstan, Turkmenistan and South Korea. As school teaching in the Uzbek language predominates (over 8,800 schools in 2006/7) with a small number of school (about 760 schools in 2006/7) instruction in Russian and other languages (e.g., Russian-English) (UNESCO-IBE, 2011), this might make student learning more challenging as all the modules taught in the IBC are in English. As English is the medium of instruction in the IBC, students' ability to listen and comprehend in class, speak and clarify doubts with faculty members, read beyond study packs, and write fluently in coursework and exams usually has a significant impact on their overall academic performance. Therefore, students who have attended a Russian-English academic lyceum or vocational college with a minimum IELTS of 5.5 can only be admitted to the IBC Foundation Year programme.

In addition, culture, historical traditions, and the teaching methodologies in Uzbekistan pre-tertiary education generally focus on teacher-centred in preference to student-centred learning. This leads to Uzbek students to expect to be passive recipients of information and to rely on summarised study packs instead of reading from the textbooks and other reference sources. They are not able to adopt an independent approach to learning and problem solving, especially applying critical analysis to essay writing and coursework preparation. Uzbek students expect lecturers/tutors to state the facts for them to memorise and regurgitate. Students who perform poorly in their academic study often suffer from the inability to adapt to the independent learning style in the international programme (Rahal & Palfreyman, 2009). Foreign faculty members face the challenge of adjusting their teaching methods to satisfy both the expectations of their institutions and the preferences of the students. It will be the students who usually have to adjust to the new styles of learning and lecturers/tutors will often need to offer additional guidance and support to students not familiar with student-centred learning methods.

In order to motivate students learning, transnational faculty members ought to understand that in different countries, students may have their own preferred learning styles (Mahrous & Ahmed, 2010). It is necessary therefore for foreign faculty members to suspend their assumptions about the teaching methods they have used elsewhere and to review the suitability of modifying their teaching methodologies for local contexts (Smith, 2009). In Uzbekistan, the first language is Uzbek, second language is Russian, third language is Tajik, and English is an optional language for the majority of students not only in the IBC, but also in other state universities in the country. Students may not be able to fully understand lectures, especially the subjects are full of technical jargon, therefore, transnational faculty members are advised to take note to allow time for them to digest in class.

Conclusion

In order to improve student learning outcomes, the home institution or IBC is advised to prepare transnational faculty members by providing advice and guidance on pedagogical issues or country-specific issues and differences before they are deployed abroad (Dunn & Wallace, 2006; McBurnie & Ziguras, 2007). This professional development would enhance their transnational teaching experience by developing their understanding on local culture and traditions, such as religious customs and family relationships and expectations, and avoid offending students. In addition, they would appreciate the personal and socio-economic factors affecting student performance. In order to adapt to local culture, they are advised to learn to respect students, being resilient to classroom culture, listen to their problems and provide advice, being friendly and having a sense of humour, being dedicated and knowledgeable in the subject

of teaching, and being patient and fair (Saafin, 2008).

Higher education reform in Uzbekistan is seen to be necessary to upgrade the embedded Soviet education system to benchmark against the international standards. With the establishment of three Russian- and four Western-based IBCs in Tashkent, the reconfiguration of post-Soviet higher education system in Uzbekistan has resulted in a hybrid of the East and West model of education system. Guidelines on education standards and intercultural competence of transnational teaching are available to all IBCs to enable a smoother transition to support the reform process.

The selection and recruitment of foreign faculty members in the IBC under study is usually based on their international teaching exposure as well as religious and cultural similarities with Uzbekistan. In general, students in the IBC have high regard for both the foreign faculty members and local lecturers/tutors. Professional development in intercultural competence is essential for faculty members working in transnational contexts. Cultural diversity in the global market place has shone the spotlight on intercultural competence as a very important skill for teaching staff (Spitzberg & Changnon, 2009). Beside acquiring the right attitudes, knowledge and comprehension, and skills, the ability to adapt to other cultures, navigate one's emotions, learn intercultural sensitivity, and manage conflict are also key aspects of developing intercultural competence.

Since intercultural competence is not a naturally occurring phenomenon, it has to be intentionally addressed at the institutional level, i.e., either through pre-departure briefing by the home HEIs or dialogue sessions with the management of IBC. In utilising such a framework in orientation briefings and management dialogue sessions, the efforts toward developing intercultural competence in transnational faculty members can be included in a more comprehensive, integrated approach instead of through random, ad-hoc approaches that often occur at institutions. It is also important to assess these efforts so as to improve the process of developing intercultural competence among transnational faculty members and to also provide meaningful feedback to them that could help them on their intercultural journey.

References

Bennett, J. M., Bennett, M. J., & Allen, W. (2003). Developing intercultural competence in the language classroom. In D. L. Lange, & R. M. Paige (Eds.), *Culture at the Core: Perspectives on Culture in Second Language Learning* (pp. 237 - 270). Greenwich, CN: Information Age Publishing.

Crabtree, R. D., & Sapp, D. A. (2004). Your culture, my classroom, whose pedagogy? Negotiating effective teaching and learning in Brazil. *Journal of Studies in International Education, 8*, 105. doi: 10.1177/1028315303260826

Deardorff, D. K. (2009). *The Sage Handbook of Intercultural Competence*. Thousand Oakes, CA: Sage Publication, Inc.

Dunn, L., & Wallace, M. (2006). Australian academics and transnational teaching: An exploratory study of their preparedness and experiences. *Higher Education Research & Development, 25*(4), 357 - 369. doi: 10.1080/07294360600947343

Education, Audiovisual and Culture Executive Agency (EACEA) (2012). *Higher Education in Uzbekistan*. Retrieved from http://eacea.ec.europa.eu/tempus /participating_countries/overview/Uzbekistan.pdf

Gopal, A. (2011). Internationalization of higher education: Preparing faculty to teach cross-culturally. *International Journal of Teaching and Learning in Higher Education*, *23*(3), 373 - 381.

Gray, D. E. (2007). Facilitating management learning: Developing critical reflection through reflective tools. *Management Learning*, *38*(5), 495 - 517. doi: 10.1177/1350507607083204

Greenholz, J. (2000). Assessing cross-cultural competence in transnational education: The intercultural development inventory. *Higher Education in Europe, 25*(3). 411 - 416. doi: 10.1080/713669273

Gribble, K., & Ziguras, C. (2003). Learning to teach offshore: Pre-departure training for lecturers in transnational programs. *Higher Education Research & Development, 22*(2), 205 - 216. doi: 10.1080/07294360304115

Hannigan, T. (1990). Traits, attitudes, and skills that are related to intercultural effectiveness and their implications for cross-cultural training: A review of the literature. *International Journal of Intercultural Relations, 14*, 89 - 111. doi: 10.1016/0147-1767(90)90049-3

Heyneman, S. P. (2011). A comment on the changes in higher education in the post-Soviet Union. In I. Silova (Ed.), *Globalization on the Margins: Education and post-Socialist Transformations in Central Asia* (pp. 27 - 40). Charlotte, NC: Information Age Publishing.

Johnson, L. S. (2003). The diversity imperative: Building a culturally responsive school ethos. *Intercultural Education, 14*(1), 17 - 30. doi: 10.1080/1467598032000044629

Jones, P. (2011). Education as foreign policy: The European Union in Central Asia. In I. Silova (Ed.), *Globalization on the Margins: Education and post-Socialist Transformations in Central Asia* (pp. 63 - 94). Charlotte, NC: Information Age Publishing.

Leask, B. (2008). Teaching for learning in the transnational classroom. In L. Dunn, & M. Wallace (Eds.), *Teaching in Transnational Higher Education: Enhancing Learning for Offshore International Students* (pp. 120 - 132). NY: Taylor and Francis.

Littlejohn, S. W., & Domenici, K. (2007). *Communication, Conflict, and the Management of Difference*. Long Grove, IL: Waveland.

Mahrous, A. A., & Ahmed, A. A. (2010). A cross-cultural investigation of students' perceptions of the effectiveness of pedagogical tools. *Journal of Studies in International Education*, *14*(3), 289 - 306. doi: 10.1177/1028315309334738

McBurnie, G., & Ziguras, C. (2007). *Transnational Education: Issues and Trends in Offshore*

Higher Education. London: Routledge.

Mezirow, J. (1998). On critical reflection. *Adult Education Quarterly, 48*, 185 - 198. doi: 10.1177/074171369804800305

Merriam, S. B. (2007). *Non-western Perspectives on Learning and Knowing: Perspectives from Around the World*. Malabar, FL: Krieger Publishing Company.

Mok, K. H. (2012). Rise of transnational higher education in Asia: Student mobility and studying experiences in Singapore and Malaysia, *Higher Education Policy*, *25*, 225 - 241. doi: 10.1057/hep.2012.6

Nessipbayeva, O., & Dalayeva, T. (2013). Developmental perspectives of higher education in the post-Soviet countries (for the cases of Kazakhstan, Uzbekistan, Kyrgyzstan, Tajikistan and Turkmenistan). *Procedia - Social and Behavioral Sciences*, *89*, 391 - 396. doi: 10.1016/j.sbspro.2013.08.865

Rahal, T., & Palfreyman, D. (2009). Assessing learning styles of students at Zayed University. *Learning and Teaching in Higher Education: Gulf Perspectives*, *6*(2). Retrieved from http://www.zu.ac.ae/lthe/lthe06_02_01_rahal.htm.

Saafin, S. (2008). Arab tertiary students' perceptions of effective teachers. *Learning and Teaching in Higher Education: Gulf Perspectives*, *5*(2). Retrieved from http://www.zu.ac.ae/lthe/lthe05_02_02_saafin.htm.

Silova, I. (2011). Higher education reforms and global geopolitics: Shifting cores and peripheries in Russia, the Baltics, and Central Asia. *Russian Analytical Digest*, *97*, 9 - 12.

Silova, I., & Steiner-Khamsi, G. (Eds.) (2008). *How NGOs React: Globalization and Education Reform in the Caucasus, Central Asia, and Mongolia*. Bloomfield, CT: Kumarian Press. Retrieved from http://www.tc.columbia.edu/faculty/steinerkhamsi/_publications/Gitas%20Professional%20Files/Chapters%20in%20edited%20volumes/Unwrapping2008.pdf

Smith, K. (2010). Transnational teaching experiences: An under-explored territory for transformative professional development. *International Journal for Academic Development, 14*(2), 111 - 122. doi: 10.1080/13601440902969975

Smith, L. (2009). Sinking in the sand? Academic work in an offshore campus of an Australian university. *Higher Education Research & Development*, *28*(5), 467 - 479. doi: 10.1080/07294360903154118

Spitzberg, B. H., & Changnon, G. (2009). Conceptualizing intercultural competence. In D. K. Deardorff (Ed.), *The Sage Handbook of Intercultural Competence* (pp. 2 - 52). Thousand Oakes, CA: Sage Publications, Inc.

Storti, C. (2009). Passing it on: Intercultural competence in the training arena. In D. K. Deardorff (Ed.), *The Sage Handbook of Intercultural Competence* (pp. 272 - 286). Thousand Oakes, CA: Sage Publications, Inc.

Tomusk, V. (2008). The Bologna process and the enlightenment project: A bird's eye view from the bottom of the slope. *European Education: Issues and Studies*, *2*(3), 9 - 28. doi: 10.2753/EUE1056-4934400201

UNESCO-IBE (2010/11), World Data on Education, 7th Ed. Retrieved from http://www.ibe.unesco.org/fileadmin/user_upload/Publications/WDE/2010/pdf-versions/Kyrgyzstan.pdf

Weiner, B. (Ed) (1974). *Achievement, Motivation and Attribution Theory*. Morristown, NJ: General Learning Press.

Wilkins, S. (2010). Higher education in the United Arab Emirates: An analysis of the outcomes of significant increases in supply and competition. *Journal of Higher Education Policy and Management*, *32*(4), 389 - 400. doi: 10.1080/1360080X.2010.491112

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 70 - 87.
doi: 10.14434/josotl.v15i1.12954

The relationship between cumulative credits and student learning outcomes: A cross-sectional assessment of information literacy and communication skills

Teresa Thonney[1] and Joe C. Montgomery[2]

Abstract: This article relates the efforts of faculty at one community college to define standards for achievement of two SLOs (critical thinking and effective communication) and to gather and analyze evidence of how well students meet those standards. Faculty from 13 disciplines assessed writing samples from 265 students. We found that, in general, students with more credits outperformed those with fewer credits. However, many students at every level demonstrated poor information literacy skills, indicating an area for curriculum improvement.

Keywords: critical thinking, information literacy, student learning outcomes, writing skills assessment

Introduction

Today, more than ever, accreditation commissions are asking colleges and universities to "provide substantive data-driven evidence of quality and effectiveness" (Baker, 2002, p. 5), and regional accreditors are finding an increasing number of institutions "out of compliance" with standards that require them to assess student achievement and improve programs based on their assessment findings (Nunley, Bers, & Manning, 2011).

At Columbia Basin College, a community college in Washington State, our accrediting body has called for a cycle of outcomes-driven assessment. In fact, our funding is increasingly based on demonstrating that students are achieving the institution's student learning outcomes (SLOs). Those outcomes include:

- Apply information tools and resources
- Develop cultural awareness
- Think critically
- Communicate effectively
- Reason quantitatively and symbolically

These core competencies, listed among the student learning outcomes of most community colleges, cut across disciplines and are addressed in many courses; yet they are not mastered in any single course. They are instead developed over time.

According to the American Association for Higher Education 1999 Assessment Forum, assessment involves "setting appropriate criteria and high standards for learning quality; systematically gathering, analyzing, and interpreting evidence to determine how well performance matches those expectations and standards; and using the resulting information to document, explain, and improve performance" (cited in Ohlemacher & Davis, 2012, p. 11). In this article, we describe the efforts of faculty at Columbia Basin College to define standards for

[1] Columbia Basin College, TThonney@columbiabasin.edu
[2] Columbia Basin College, jmontgomery@columbiabasin.edu

student achievement of two SLOs—critical thinking and effective communication—and to gather and analyze evidence of how well our students meet those standards.

Critical Thinking and Information Literacy

Communication and critical thinking skills are, of course, essential for success in college and careers. In our digital age, one especially important kind of critical thinking is information literacy, specifically the ability to critically analyze and assess source information for reliability and relevance. It's a skill that is undeveloped in most first-year students, according to the findings of Project Information Literacy, an ongoing series of student surveys and interviews. In one of the project's recent surveys, for example, almost 50% of the 8,353 respondents expressed uncertainty about how to evaluate their research efforts (Head & Eisenberg, 2010).

Researchers involved with The Citation Project, another ongoing national study, are also evaluating how students select and use sources. They have found—by reviewing student papers from colleges of all types—that few students share faculty's definition of reliable information. In fact, say the project directors, many of the sources students cite are "stunningly cheesy" (Jamieson & Howard, 2011). Local studies routinely confirm that many undergraduates uncritically accept sources and source information (e.g., Daniels, 2010; Twait, 2005; Choinski & Emanual, 2006; Swoger, 2011). Choinski, Mark, and Murphey (2003) found that even after taking an information resources class, many students are unable to recognize signs of a questionable website.

Given the importance of information literacy and given the evidence indicating it is a skill first-year students need to develop, our college's Teaching and Learning Committee designed a study to assess students' information literacy and written communication skills. These questions guided our research: 1) Would students recognize signs of bias or question the reliability of information in sources? and 2) Would students demonstrate what we considered to be college-level writing skills? Finally, in light of recent research indicating that little growth occurs in students' thinking and writing skills during the first two years of college (e.g., Arum & Roksa, 2011), we asked a third question: Do students in their second year of course work demonstrate higher critical thinking and written communication skills than students just beginning college? In other words, is there a correlation between the number of credits earned and students' information literacy and written communication skills?

Measuring Student Learning Outcomes

There are many tools for measuring critical thinking and communication outcomes at the institutional level, each with benefits and drawbacks. **Student satisfaction or experience surveys**, such as the National Survey of Student Engagement (NSSE), provide insight into students' satisfaction and attitudes (Zoellner, Samson, & Hines, 2008), as well as self-perceived gains in cognitive ability and writing skill (Cheng, 2001); but they may not be accurate measures of learning or skills (Coupe, 1993; Maughan, 2001; Maguire, Evans, & Dyas, 2001; Thompson, Pilgrim, & Oliver, 2005). In fact, the literature on self-perception indicates that novices tend to over-estimate their skill level (Schilling & Applegate, 2012; Kruger & Dunning, 1999).

Portfolios, common assignments, or other **course work** can provide more objective evidence of student achievement level. At the Community College of Baltimore County, for example, teams of faculty develop Common Graded Assignments. Students in designated

courses complete the assignment during a given term, and then normed faculty score random samples of the assignments using a six-item rubric that corresponds with the college's general education outcomes. (See https://ccbcmd.edu/loa/great.html.) Faculty at Miami Dade College have also designed writing tasks that allow them to measure several learning outcomes by scoring one artifact. Each term, willing faculty assign a common writing project, and teams of faculty use rubrics to score a random sample of completed projects from students who are about to graduate (Nunley, Bers, & Manning, 2011).

Common assignments, scored by normed faculty, provide a snapshot of the achievement level of a college's graduates, but the results do not necessarily tell faculty anything about how much students have learned or gained from their coursework. **Pre- and post-testing** is needed to demonstrate learning or skill development in students. A pre-test administered at the beginning of a course or program can gauge a student's initial skill or knowledge level, which can later be compared to the student's performance at the end of the course or program (Caspers & Bernhisel, 2005; Lombardo & Miree, 2003). However, post-tests administered immediately after a single course may reflect short-term retention rather than real learning (Cmor, Chan, & Kong, 2010); and pre- and post-tests at entrance and graduation are a challenge to administer at the institutional level (Nunley, Bers, & Manning, 2011).

Standardized aptitude tests are another direct measurement of student learning outcomes. Cisneros (2009), for instance, measured improvement in critical thinking skills by having graduate pharmacy students take the California Critical Thinking Skills Test (CCTST) at the beginning of the school year and the California Critical Thinking Disposition Inventory (CCTDI) at the end of the same school year. The CAAP Critical Thinking Test is another nationally normed exam for measuring students' ability to analyze and evaluate. These multiple-choice instruments have been validated and they are easy to administer, but they provide no opportunity to measure students' written communication skills.

The Collegiate Learning Assessment (CLA), on the other hand, is an open-ended test of analytic reasoning, critical thinking, problem solving, and writing skills. Its goal is "to provide a summative assessment of the *value-added* by the school's instructional and other programs (taken as a whole) with respect to certain important learning outcomes" (Klein, Benjamin, Shavelson, & Bolus, 2007, p. 418, original emphasis). Students analyze case studies and use information from sources provided to them to justify their decisions in real-world documents, such as memos or policy recommendations. Students don't necessarily take the CLA more than once, as a pre- and post-test, but schools can assess how their students perform in relation to students with the same mean SAT (or ACT) scores at other similar schools. Administering a standardized test like the CLA is, however, expensive (Swing & Coogan, 2010), and, like most community colleges, our institution does not require SAT or ACT scores, making a comparison of our students' CLA scores to those of students at comparable schools difficult.

To assess students' critical thinking and writing skills, we needed a means of assessment that, like the CLA, was open-ended, and that, like the common assignments used at the Community College of Baltimore County and Miami Dade College, would allow us to measure more than one outcome. We decided to have a sample of students write essays in response to a common prompt.

Methods

The Writing Prompt

In order to assess our students' ability to critically evaluate information from sources, we presented an argument and asked a sample of students to respond to that argument using data provided to them. We explicitly directed students to identify weaknesses in the evidence provided. Figure 1 includes the first page of the packet students received.

The Assignment: In a recent book, Mark Bauerlein, an English professor, argues that "the digital age stupefies young Americans and jeopardizes our future" by producing hyper-networked kids who rarely read books and who know more about the latest pop idol than they know about history, politics, economics, or culture. In fact, Bauerlein has dubbed today's youth "The Dumbest Generation."

Read and review the attached documents (Documents A-E), and then answer the following question:

Is the so-called intellectual decline in America as serious as critics like Bauerlein believe?

Respond in a well-organized, multi-paragraph essay (2 pages minimum, double-spaced). Include specific details from **all** of the attached documents in your essay. *In addition, explain the limitations of the attached data*. (For example, what additional information would help you write a more complete response to the question or help you assess the reliability of the attached data?)

Your essay will be evaluated according to how well you do the following:

- Respond appropriately to the assigned question (in bold type, above)
- Interpret the evidence provided
- Recognize the limitations or flaws of the evidence provided
- Respond in an organized and focused essay
- Use correct grammar, word choice, and spelling

Figure 1. The Writing Prompt

By directing students to "explain the limitations of the attached data," highlighting this requirement in italics, and providing examples of questions they might ask about the data, we hoped to get an accurate picture of students' ability to assess evidence.

Five documents followed the prompt. Document A was a summary of Bauerlein's argument, written by Sharon Begley for *Newsweek Magazine*. The following sentences introduced Begley's two paragraph summary:

> The following passage is taken from an article titled "The Dumbest Generation? Don't Be Dumb," by Sharon Begley. It was originally published in *Newsweek Magazine.* In these paragraphs, Begley summarizes some of the evidence *Mark Bauerlein, a professor at Emory University, includes in his book*. (original emphasis)

A reader could not determine from Begley's summary whether or not Begley agrees with Bauerlein's argument, although her article title suggests she takes issue with Bauerlein's position.

Document B was a line graph showing U.S. high school graduation rates between 1870 and 2005. The graph indicates that the graduation rate rose steadily until 1970, when it peaked at

75% and then leveled off at roughly 70%. Document C was also a line graph, showing that college graduation rates have risen steadily since 1900. Document D included two tables, showing scores for the National Association of Educational Progress (NAEP) test for Math and Reading for two years (1975 and 2008), for three ethnicities (white, black, and Hispanic), and for three age groups (9-year-olds, 13-year-olds, and 17-year-olds). For all age groups and ethnicities, average math and reading scores have risen since 1975, although the rate of improvement has varied by ethnicity. The following citation appeared at the bottom of Documents B, C, and D:

> Source: Murray, Charles. *Coming Apart*. Random House: New York, 2012.

The final document (E) was introduced in this way:

> From *The Oprah Winfrey Show* titled *Waiting for "Superman": The Movie That Could Revolutionize Schools.* Published September 20, 2012. The screen shot below is taken from http://www.oprah.com/relationships/Shocking-Education-Statistics

The screen shot, titled "Shocking Education Statistics," with the subtitle "Get the truth about America's school system," included several figures. One figure claims that "approximately 7,000 kids drop out of school **every** day" (original emphasis). Other figures contradict information found in Documents B, C, and D. For example, one chart suggests that students' reading and math scores remained unchanged between 1971 and 2010 (no student ages or actual scores were identified). No source is identified for any of the statistics in Document E, and ads for McDonald's and birth control appear in the margins of the screen shot.

A number of red flags make the information in Document E at best limited and at worst suspect. The reported reading and math scores, for example, are dubious if for no other reason than they have stayed exactly the same for 40 years; and without knowing the dropout rates for earlier decades, the 7000-dropouts-a-day figure is of no use for responding to Bauerlein's argument.

Of course, we hoped that students would critically read the other documents as well, as some students did. For example, many students recognized that in order to respond to Bauerlein's argument, they must consider graduation rates and test scores in historical context. On the other hand, no student thought to ask who Charles Murray is, despite the fact that several figures were attributed to him. We do not assume that this was because students are familiar with the scholar. Some students missed other rhetorical clues, wrongly attributing claims in Begley's summary (of Bauerlein), for instance, to Begley. In addition, the provided documents addressed only one marker of intelligence (formal education), something only a few students noted in their essays.

Study Procedure

We recruited faculty volunteers willing to have their students write the essay. At the beginning of the class hour, each instructor read a script provided to them summarizing the instructions and grading criteria. Students received the writing instructions (Figure 1), followed by the collection of five sources (documents A-E) and paper on which to write their essays

In all, 265 students wrote essays in response to the writing prompt during a 60-minute class session. The students were enrolled in either a 100 or 200-level business, composition, history, nursing, psychology, or sociology course (n = 176) or a 1-credit first-year introduction (FYI) course (n = 89). The FYI course is required for all students and is taken before they earn any college credit. The sample was not random or representative, but it did include students who

had not yet taken any courses, students in their first year of course work, and students in their second year of course work.

Students enrolled in the FYI course wrote their essays during spring break of 2013; students enrolled in a college-level course wrote their essays during the first three weeks of spring quarter 2013. To ensure that we measured summative skill level rather than short-term learning acquired in a specific course, none of the instructors discussed the writing topic with students beforehand. In addition, with the exception of composition, the focus of the courses was not on written communication or information literacy.

The faculty volunteers told students that the writing was "required," although it was left to the instructor to determine the weight of the assignment. In this way, we hoped to avoid the bias inherent in volunteer samples. However, making the essay a course requirement was likely more effective in the college-level courses than in the FYI course, where students receive a grade of pass or fail, a point we return to later.

Scoring Procedure

A team of 23 faculty members, representing 13 disciplines, met to score the essays using a rubric designed and previously revised by members of the college's Teaching and Learning Committee. Used for decades to assess student writing, rubrics have also been used to evaluate other SLOs, including information literacy skills (Choinski, Mark, & Murphey, 2003; Knight, 2006; Oakleaf, 2009). Using a rubric allowed us to quantify and compare assessments of various graders.

The rubric included different performance levels for five criteria:

- responds appropriately to the prompt;
- interprets the evidence correctly;
- recognizes limitations or flaws in the evidence;
- responds in an organized and coherent essay; and
- uses correct grammar, word choice, and spelling.

The first three items in the rubric addressed the college's critical thinking SLO; the last two items addressed communication.

Before rating essays individually, faculty scored sample essays and discussed their scores, repeating the process until sufficiently "normed." Each criterion was rated on a four-point scale with "1" being the highest performance and "4" being the lowest. (See Figure 2.) Giving each criterion a separate score allowed us to measure several outcomes from a single writing sample.

Two scorers read each essay. When there was discrepancy of more than one point on any criterion, a third person read the essay. For example, if one rater gave an essay a score of "4" for "recognizes limitations or flaws in the evidence" and another rater gave the same element a score of "2," a third rater read the essay, even if scores for the other rubric criteria were identical. In all, 40 of 265 essays (15%) received a third reading because scores awarded for one or more criteria by the first two readers differed by more than one point.

We averaged the readers' scores to create a single rating for each element. We classified a criterion receiving a rating between 1.0 and 3.0 to be at "college level," meaning that specific criterion within the paper would warrant a grade of "A," "B," "C," or "D" if the paper were submitted in a college-level course. We classified ratings above 3.0 as below college level, meaning the specific criterion within the paper would warrant a grade of "F." Because of the

different rubric criteria, some aspects of a paper could be deemed to be at college level (i.e., warranting a grade of A-D); other aspects could be deemed to be below college level (i.e., warranting a grade of F).

Inter-Rater Reliability

An earlier version of the rubric, used in a smaller pilot study in 2012, included six rating levels for each criterion, but that pilot resulted in low inter-rater reliability, attributed to both insufficient norming and too many gradations of evaluation. Raters using the revised rubric (Figure 2), however, had high levels of agreement.

Element	**Superior** (Always displays this element) 1	**Skilled** (Displays this element most of the time) 2	**Minimal Competence** (Displays this element occasionally or somewhat) 3	**Inadequate/Below College Level** (Displays this element rarely or not at all) 4
Responds appropriately to the prompt *Addresses the assigned question and all sections are related to the assigned question (i.e., are on topic).*				
Interprets the evidence correctly *Interprets evidence correctly **and** connects the evidence to the assigned purpose.*				
Recognizes limitations or flaws in the evidence *Recognizes flaws in the evidence **or** recognizes how the evidence is insufficient to address the problem.*				
Responds in an organized and coherent essay *Ideas and paragraphs are focused, sufficiently developed, logically connected, **and** related to the assigned purpose.*				
Uses correct grammar, word choice, and spelling				

Figure 2. Rubric Used to Assess Critical Thinking and Written Communication SLOs

We evaluated inter-rater reliability using the intra-class correlation coefficient (ICC) provided by the SPSS RELIABILITY procedure (SPSS, Version 19). High levels of the ICC indicate that raters gave similar ratings; low levels of agreement indicate that raters gave varying ratings. ICC values exceeding 0.7 were classified as "High" agreement, 0.5 to 0.7 as "Moderate" agreement, and 0.3 to 0.5 as "Fair" agreement. Values below 0.3 were labeled "Low" reliability.

As shown in Table 1, there were high levels of inter-rater agreement for four of the five elements: "responds appropriately to the prompt" (0.78); "interprets the evidence correctly"

(0.81); "recognizes limitations or flaws in the evidence" (0.85); and "responds in an organized and coherent essay" (0.80). The remaining element, "uses correct grammar, word choice, and spelling," showed moderate to fair levels of agreement (0.59). By definition, papers read by three raters had lower inter-rater reliability than papers read by two raters. These values ranged from a high of 0.72 for "recognizes limitations or flaws in the evidence" to a low of 0.36 for "uses correct grammar, word choice, and spelling."

Table 1

Inter-Rater Reliability

Number of Raters	Number of Cases	Q1. Responds appropriately to the prompt		Q2. Interprets the evidence correctly		Q3. Recognizes limitations or flaws in the evidence		Q4. Responds in an organized, coherent essay		Q5. Uses correct grammar, word choice, and spelling	
		ICC	Rater agreement	ICC	Rater agreement	ICC	Rater agreement	ICC	Rater agreement	ICC	Rater agreement
2 raters	225	0.78	High	0.81	High	0.85	High	0.8	High	0.59	Moderate
3 raters	40	0.47	Fair	0.61	Moderate	0.72	High	0.6	Moderate	0.36	Fair

Results

Overall Findings

Of the five rubric elements, as shown in Table 2, "uses effective grammar, word choice, and spelling" had the highest percentage of papers rated at "passing" level performance (85%), followed closely by "responds appropriately to the prompt" (82%). "Interprets the evidence correctly" and "responds in an organized and coherent essay" had 70% at passing level. "Recognizes limitations of the evidence" had the lowest rate of college-level performance (54%). While many students could correctly interpret the data and write a coherent paper, few challenged the data or questioned its sufficiency to answer the question in the prompt.

Table 2

Percentage of students scoring at college level for each criterion

Rubric Criteria	**Percent of students who scored at college level**
Uses effective grammar, word choice, spelling	85%
Responds appropriately to the problem	82%
Interprets the evidence correctly	70%
Responds in an organized and coherent essay	70%
Recognizes limitations of evidence	54%

Demographics, Ability, and Performance Variables

Additional analyses were performed to identify demographic or ability variables that might be associated with, and predictive of, rubric ratings. For example, one might anticipate that students with a higher cumulative grade point average, students with a greater number of credits, or students who performed at college level on the college's reading, writing, or math placement test (COMPASS) would earn higher ratings on the rubric elements. Consequently, multiple regression analyses were performed using the five rubric elements as the dependent variables and age, cumulative GPA, cumulative credits, and COMPASS outcomes as independent variables. Multiple regression analyses have the benefits of showing the total predictability of the dependent variable and the unique significance of each independent variable in predicting the dependent variable. The analysis holds constant the remaining independent variables while assessing the contribution of a single independent variable.

Results of the multiple regression analyses are shown in Tables 3 - 7. Table 3 shows the results for prediction of the first rubric element ("responds appropriately to the prompt") from age, cumulative GPA, and placement at college- or pre-college levels for writing, reading, and mathematics. The overall R-squared value of 0.15 was significant ($p < .01$), although the only significant predictor was cumulative credits (beta = -0.261, $p < .0001$). That is, those with more credits received significantly higher ratings on this criterion.

Table 3

Multiple regression: Predicting the element "responds appropriately to the prompt"

Responds Appropriately to the Prompt Regression Table					
	Unstandardized Coefficients		Standardized Coefficients		
Variables	B	Std. Error	Beta	T	Sig.
Age	.013	.007	.115	1.710	.089
Cumulative GPA	-.025	.050	-.033	-.504	.615
Cumulative Credits	-.003	.001	-.261	-3.824	.000
College-level Writing	-.221	.116	-.143	-1.908	.058
College-level Reading	-.202	.118	-.125	-1.719	.087
College-level Math	.018	.131	.009	.136	.892

Notes: $R^2 = .15$ (ps <.01)

Table 4 shows a similar analysis for the second rubric element ("interprets evidence correctly") as the dependent variable. The R-squared value was again 0.15 ($p < .01$) and cumulative credits was again a significant predictor (beta = -0.228, $p < .001$). However, placement in writing was also a significant predictor (beta = -0.167, $p < 0.026$), with those placing at college-level in writing scoring significantly higher for this rubric element.

Table 5 shows the multiple regression results for predicting the element "recognizes limitations of evidence." The R-squared value for this analysis was 0.22 ($p < 0.01$), and there were three significant predictors. The strongest predictor was cumulative credits (beta = - 0.302,

p < .0001), followed by placement in writing (beta = -0.183, p < .011) and placement in reading (beta = -0.271, p < .032). Those with a higher number of accumulated credits and college-level placement in writing and reading received significantly higher ratings.

Table 4

Multiple regression: Predicting the element "interprets evidence correctly"

Interprets the Evidence Correctly Regression Table					
	Unstandardized Coefficients		Standardized Coefficients		
Variables	B	Std. Error	Beta	t	Sig.
Age	.008	.007	.077	1.153	.250
Cumulative GPA	-.048	.050	-.062	-.955	.340
Cumulative Credits	-.003	.001	-.228	-3.344	.001
College-level Writing	-.259	.116	-.167	-2.239	.026
College-level Reading	-.123	.117	-.076	-1.046	.297
College-level Math	-.119	.131	-.060	-.908	.365

Notes: R^2 = .15 (ps <.01)

Table 5

Multiple regression: Predicting the element "recognizes limitations of evidence"

Recognizes Limitations of Evidence Regression Table					
	Unstandardized Coefficients		Standardized Coefficients		
Variables	B	Std. Error	Beta	t	Sig.
Age	.009	.008	.071	1.109	.268
Cumulative GPA	.016	.054	.019	.308	.759
Cumulative Credits	-.004	.001	-.302	-4.629	.000
College-level Writing	-.318	.124	-.183	-2.570	.011
College-level Reading	-.271	.125	-.150	-2.161	.032
College-level Math	-.177	.140	-.080	-1.264	.207

Notes: R^2 = .22 (ps <.01)

Table 6 shows the multiple regression analysis for the fourth rubric element, "responds in an organized and coherent essay." The R-squared value of 0.24 was significant (p < .01) and cumulative credits was the strongest predictor (beta = -0.296, p < .0001), followed by placement levels in writing (beta = -0.187, p < .009) and reading (beta = -0.172). Those with more credits and higher placement scores in writing and reading received significantly higher ratings.

Table 7 provides the results of the final multiple regression analysis, with the rubric element "uses effective grammar, word choice, and spelling" as the dependent variable. The R-squared value of 0.26 was significant ($p < .01$). Cumulative credits was the strongest predictor (beta = - 0.293, $p < .0001$), followed by placement levels in writing (beta = -0.242, $p < .001$) and reading (beta = -0.175, $p < .011$). Again, those with a higher number of college credits and college-level placement in writing and reading received significantly higher ratings.

Table 6

Multiple regression: Predicting the element "responds in an organized and coherent essay"

Responds in an Organized and Coherent Essay Regression Table					
	Unstandardized Coefficients		Standardized Coefficients		
Variables	B	Std. Error	Beta	t	Sig.
Age	.001	.007	.014	.212	.832
Cumulative GPA	-.052	.047	-.069	-1.120	.264
Cumulative Credits	-.004	.001	-.296	-4.580	.000
College-level Writing	-.284	.107	-.187	-2.647	.009
College-level Reading	-.272	.109	-.172	-2.496	.013
College-level Math	-.021	.122	-.011	-.174	.862

Notes: $R^2 = .24$ (ps <.01)

Table 7

Multiple regression: Predicting the element "uses effective grammar, word choice, and spelling"

Uses Effective Grammar, Word Choice, and Spelling Regression Table					
	Unstandardized Coefficients		Standardized Coefficients		
Variables	B	Std. Error	Beta	t	Sig.
Age	-.010	.006	-.102	-1.631	.104
Cumulative GPA	-.013	.042	-.018	-.305	.761
Cumulative Credits	-.003	.001	-.293	-4.601	.000
College-level Writing	-.335	.096	-.242	-3.472	.001
College-level Reading	-.252	.098	-.175	-2.574	.011
College-level Math	.084	.109	.048	.769	.443

Notes: $R^2 = .26$ (ps <.01)

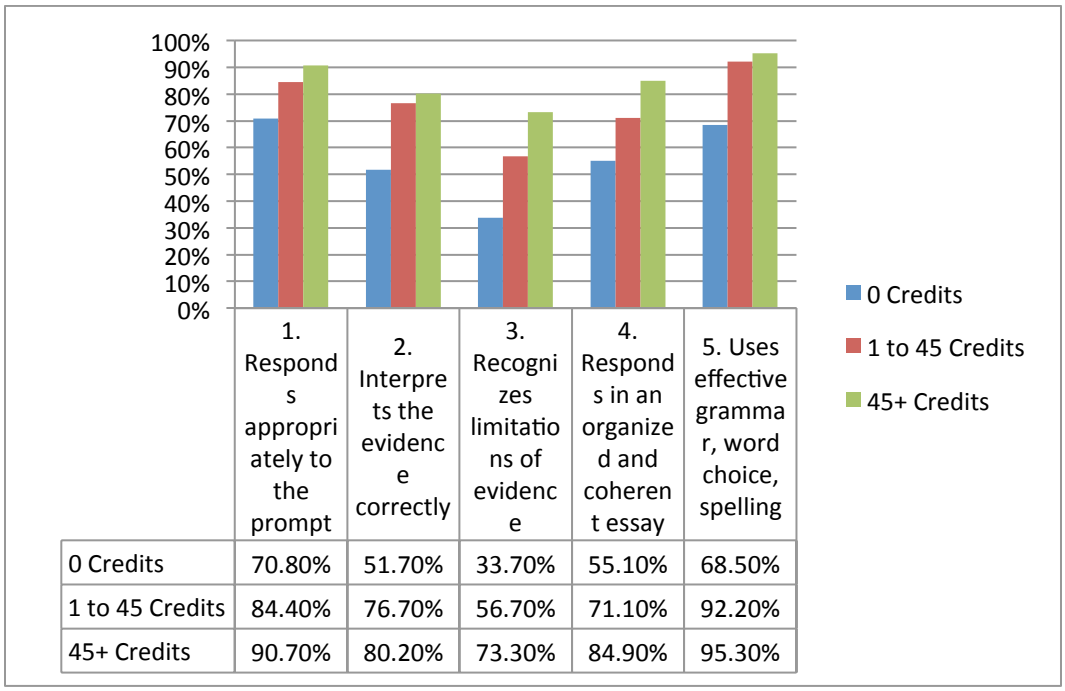
For all five rubric elements, "cumulative credits obtained" was the most highly significant predictor of ratings. For each analysis, higher levels of credits were associated with

higher ratings on the elements, with the significance levels at $p < .001$ or higher. The standardized regression coefficient, indicating the relative predictor importance, was considerably higher for cumulative credits than for any other predictor. COMPASS writing placement was a significant predictor for four of the elements, and COMPASS reading placement was a significant predictor for three of the elements. Note that three predictors did not achieve statistical significance for any of the five analyses: age, cumulative GPA, and COMPASS math placement. That is, being older, having higher overall grades, and having higher math placement level were not associated with higher ratings on the rubric.

These analyses suggest the importance of accumulating college credits to receiving higher ratings on the critical thinking and communications rubrics. The analyses also downplay the likelihood that the results can be explained by maturation effects, ability effects, or student attrition (as discussed further below).

Analysis by Credit Levels

We performed a follow up analysis to obtain a better subjective sense of the importance of credit levels on rubric ratings. Subjects were categorized as having 0 college-level credits (entering students, $n = 89$), having 1-44 college-level quarter credits (first-year students, $n = 90$), or having 45+ college-level quarter credits (second year, if full time, $n = 86$). Students with 45+ credits outperformed the other two student groups on all five elements. (See Figure 3.) For the element showing the weakest performance, "recognizes limitations of evidence," over 70% of those with 45+ quarter credits achieved college-level performance, compared to 34% of entering students. For all five elements, students with 45+ quarter credits were rated at college-level performance by at least a 20 percentage-point margin over entering students.



	1. Responds appropriately to the prompt	2. Interprets the evidence correctly	3. Recognizes limitations of evidence	4. Responds in an organized and coherent essay	5. Uses effective grammar, word choice, spelling
0 Credits	70.80%	51.70%	33.70%	55.10%	68.50%
1 to 45 Credits	84.40%	76.70%	56.70%	71.10%	92.20%
45+ Credits	90.70%	80.20%	73.30%	84.90%	95.30%

Figure 3. Percent of Students Performing at College Level by Number of Quarter Credits Earned

Discussion of Results

When faculty see inferior sources cited in student papers, they often assume that students have used the first sources they encountered in order to complete an assignment with minimal effort. In other words, inferior evidence is thought to be a symptom of lazy research. But our results suggest the explanation is not that simple. Students in this study were provided with sources and asked to identify the limitations of the evidence. Yet despite being provided with some "stunningly cheesy" sources (e.g., the *Oprah Winfrey Show* website), 46% of the students failed to recognize any limitations. Like the students Choinski, Mark, and Murphey (2003) observed, many of the students in the present study did not recognize indicators of an unreliable website. This finding is of concern, given that students rely on the Web more than the library when conducting research (Head, 2013).

Many of today's digital-savvy students enter college aware that the Web is rife with misinformation (Manuel, 2005); so why did so many in our sample fail to recognize the evidence in the *Oprah Winfrey Show* site as flawed or incomplete? One possible explanation is that while students may "know" that they should not believe everything they read online, they may not know how to determine what not to believe. That is, students may cite inferior sources because they believe they are citing quality sources (Grimes & Boening, 2001). Certainly, in the case of the *Oprah Winfrey Show* website, the ethos of Winfrey herself may have influenced students. Students might assume that information from her show's website would be credible. Another possible reason so many students failed to critically assess the *Oprah Winfrey Show* website, in particular, is the medium of presentation (print vs. web). We provided students with a screen shot in hopes that the ads in the margins would signal to students that the site was not an academic one. However, seeing a screen shot reproduced on a page is not the same as encountering the website online. Presentation affects how viewers rate the reliability of Internet sources (Wathen & Burkell, 2002). Did students assume that sources distributed by their professors would be credible, despite being directed to critique the sources? Would students have been more skeptical of the information if they had discovered it for themselves online? Results from the Citation Project suggest the answer to the latter question is no, and, indeed, many of the students who cited information from the *Oprah Winfrey Show* website made no mention of the suspect nature of the information.

Many others students, perhaps recognizing that the data they received conflicted, opted to write one-sided arguments, picking and choosing the evidence that helped make their case but ignoring the rest. This is not surprising, as it is common for novice writers to draw from only those sources who agree with each other or with whom they agree (Penrose & Geisler, 1994). In fact, Jonathan Baron's (1995) research into "myside" bias in student arguments indicates that many students regard a one-sided argument to be *more* convincing than one that recognizes other viewpoints. Further research is needed to determine the most effective methods for training students to consider multiple viewpoints in their writing.

Closer analysis of our rating data revealed another interesting finding: For all five rubric criteria, students with 0 credits had the lowest percentage performing at college level, those with 1-44 quarter credits had a higher percentage at college level, and those with 45 or more quarter credits had the highest percentage at college level. This finding was somewhat unexpected, given what other research has shown about the rate of student development. Arum and Roksa (2011), for example, found that for 45% of the 2,322 four-year college or university students they studied there were "no statistically significant gains in critical thinking, complex reasoning, and

writing skills" during their first two years of college (p. 36). Similarly, when Cisneros (2009) measured the critical thinking skills of graduate pharmacy students at the start and end of a given school year, he found no significant improvement in students' total scores. Miller (2004) documented critical thinking gains in pharmacy students between *admission* and *graduation*, but he also found no significant difference *from year to year*, suggesting that measurable and significant gains in thinking are cumulative, occurring over a time span longer than one year.

Ultimately, the study procedures do not allow us to determine why students with the most credits earned the highest critical thinking and communication ratings. Some of the difference between student groups might be attributed to attrition. Completion rates at community colleges are routinely around 50% (Nunley, Bers, & Manning, 2011). It seems plausible that those who enter college with the lowest critical thinking and writing skills would be among those most likely to drop out, and those who enter college with higher than average critical thinking and writing skills would be more likely to persist. In addition, as already noted, students with 0 credits (enrolled in FYI) had the least motivation to do their best. However, it bears repeating that the multiple regression analyses served to control for age, GPA, and ability level (through placement test results), that these three factors had *no* significant relationship to the ratings, and that cumulative credits showed strong, significant relationships to the ratings for all five rubric criteria. These findings suggest that improvements in critical thinking and communication skills were not due simply to maturation effects, academic success, or attrition of students with lesser ability levels. Even after removing FYI student results from the data, significant differences remained between students with 45 or more quarter credits and students with 1-44 credits. It is difficult to discount the finding that across all rubric items students with more credits generally displayed better critical thinking and writing skills than did students with fewer credits.

Without both formative and summative assessment, measuring gains in any individual student's abilities isn't possible. Unfortunately, at community colleges, assessing at entrance and graduation would capture only a small percentage of students, thanks to an ever-shifting student population (Nunley, Bers, & Manning, 2011). In addition, most community college students work, raise families, have attended other colleges, and/or temporarily leave and then return, making it difficult for any college to "assert the student attained the [general education] knowledge and skills as a result of the courses and activities in which they engaged at the college" (Nunley, Bers, & Manning, 2011, p. 15). Further, the likelihood of obtaining pre- and/or post-test scores for a *non-college* attending control group is virtually nil. In short, as much as we would like to see the feature film version of each student's development, we must settle for the occasional snapshot. Nonetheless, the snapshot of our students led us to two important conclusions: Students with the most accumulated credits demonstrate the highest critical thinking and writing skills, yet many students at every credit level fail to recognize the difference between quality and inferior source information.

The college's research team will continue to assess students' critical thinking and communication skills, as well as other general education outcomes. It will be important in these further studies to introduce a more theoretical basis into the research. At this point, it is unclear why so many students (70%) could correctly interpret the data yet so few (54%) could evaluate the quality of data. According to Bloom's Taxonomy (1956) and the Revised Taxonomy (Anderson, Krathwohl, & Bloom 2001), challenging assumptions and questioning data reflect higher level thinking skills. In addition, the willpower literature (e.g., Baumeister & Tierney, 2011) suggests that cognitive resources are limited and that thinking and reasoning skills decline as these resources are expended. Perhaps the effort required to read and interpret the data, and to

write an analytic essay, depleted the cognitive resources needed for questioning and challenging the data. At any rate, future research should incorporate more of a theory base and develop *a priori*, theory-based hypotheses for testing.

Conclusion

For faculty, the reasons for measuring student achievement at the institutional level go beyond fulfilling the requirements of administrators and accrediting bodies. Without institutional-level research, faculty cannot know whether or not student learning lasts beyond the span of a single course. Significant gains in thinking and writing skills, in particular, don't occur in 12-15 weeks, making it difficult for faculty to determine whether or not "the environments they create are having their intended effects on student outcomes" (Astin & Antonio, 2012, p. 141).

This article describes how faculty at our community college assessed students' critical thinking and written communication skills. It is the first step in a cycle of assessment that we hope leads to data-driven curriculum changes. In our case, we found evidence that students are not getting sufficient practice in evaluating source information. The Teaching and Learning Committee has shared these findings with all college faculty, and we hope that information literacy will now be more frequently addressed across the curriculum—in assignment prompts, in course requirements, and in course lectures. As Angelo (1999) notes, "Assessment should be first and foremost about improving student learning and secondarily about determining accountability for the quality of learning produced. In short: Though accountability matters, learning still matters most."

References

Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Angelo, T. A. (1999). Doing assessment as if learning matters most. *AHHE Bulletin*, *51*, 1-10. Retrieved from http://assessment.uconn.edu/docs/resources/ARTICLES_and_REPORTS/Thomas_Angelo_Doing_Assessment_As_If_Learning_Matters_Most.pdf

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago: University of Chicago Press.

Astin, A. W., & Antonio, A. L. (2012). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education* (2nd ed.). Blue Ridge Summit, PA: Rowman & Littlefield Publishers.

Baker, R. L. (2002). Evaluating quality and effectiveness: Regional accreditation principles and practices. *Journal of Academic Librarianship*, *28*(1), 3-7. doi: 10.1016/S0099-1333(01)00279-8

Baron, J. (1995). "Myside bias in thinking about abortion." *Thinking and Reasoning, 1*, 221–235.

Baumeister, R. F., & Tierney, J. (2011). *Willpower*. New York: Penguin Books. doi: 10.1080/13546789508256909

Bloom, B. S. (Ed.) (1956). *Taxonomy of educational objectives, handbook I*: *Cognitive domain*. New York: Longman.

Caspers, J., & Bernhisel, S. (2005). What do freshmen really know about research? Assess before you teach. *Research Strategies*, *20*(4), 458-468. doi: 10.1016/j.resstr.2006.12.016

Cheng, D. X. (2001). Assessing student collegiate experience: Where do we begin? *Assessment & Evaluation in Higher Education*, *26*(6), 525-538. doi: 10.1080/02602930120093869

Choinski, E., & Emanuel, M. (2006). The one-minute paper and the one-hour class: Outcomes assessment for one-shot library instruction. *Reference Services Review*, *34*(1), 148-155. doi: 10.1108/00907320610648824

Choinski, E., Mark, A. E., & Murphey, M. (2003). Assessment with rubrics: An efficient and objective means of assessing student outcomes in an information resources class. *Portal: Libraries & the Academy, 3*(4), 563-575. doi: 10.1353/pla.2003.0078

Cisneros, R. M. (2009). Assessment of critical thinking in pharmacy students. *American Journal of Pharmaceutical Education*, *73*(4), 1-7. doi: 10.5688/aj730466

Cmor, D., Chan, A., & Kong, T. (2010). Course-integrated learning outcomes for library database searching: Three assessment points on the path to evidence. *Evidence Based Library and Information Practice, 5*(1), 64–81.

Coupe, J. (1993). Undergraduate library skills: Two surveys at Johns Hopkins University. *Research Strategies*, *11*(4), 188-201.

Daniels, E. (2010). Using a targeted rubric to deepen direct assessment of college students' abilities to evaluate the credibility of sources. *College & Undergraduate Libraries*, *17*(1), 31-43. doi: 10.1080/10691310903584767

Grimes, D. J., & Boening, C. H. (2001). Worries about the Web: A look at student use of Web resources. *College & Research Libraries*, *62*, 11–23. doi: 10.5860/crl.62.1.11

Head, A. J. (2013). *Learning the ropes: How freshmen conduct course research once they enter college.* Project Information Literacy Research Report. Retrieved from http://projectinfolit.org/images/pdfs/pil_2013_freshmenstudy_fullreport.pdf

Head, A. J., & Eisenberg, M. B. (2010). *Truth be told: How college students evaluate and use information in the digital age*. Project Information Literacy. Retrieved from http://projectinfolit.org/images/pdfs/pil_fall2010_survey_fullreport1.pdf

Jamieson, S., & Howard, R. M. (2011). Unraveling the citation trail. Project Information Literacy smart talk, no. 8. Retrieved from http://projectinfolit.org/st/howard-jamieson.asp

Klein, S., Benjamin, R., Shavelson, R., & Bolus, R. (2007). The collegiate learning assessment: Facts and fantasies. *Evaluation Review*, *31*(5), 415-439.

Knight, L. A. (2006). Using rubrics to assess information literacy. *Reference Services Review*, *34*(1), 43-55.

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. doi: 10.1037/0022-3514.77.6.1121

Lombardo, S. M., & Miree, C. E. (2003). Caught in the web: The impact of library instruction on business students' perceptions and use of print and online resources. *College & Research Libraries*, *64*, 6–22. doi: 10.5860/crl.64.1.6

Maguire, S., Evans, S. E., & Dyas, L. (2001). Approaches to learning: A study of first-year geography undergraduates. *Journal of Geography in Higher Education*, *25*(1), 95–107. doi: 10.1080/03098260125539

Manuel, K. (2005). What do first-year students know about information research? And what can we teach them? In H. Thompson (Ed.), *Currents and convergence: Navigating the rivers of change, proceedings of the ACRL 12th National Conference* (401–417). Chicago: Association of College and Research Libraries.

Maughan, P. D. (2001). Assessing information literacy among undergraduates: A discussion of the literature and the university of California-Berkeley assessment experience. *College & Research Libraries*, *62*(1), 71-85. doi: 10.5860/crl.62.1.71

Miller, D. (2004). An assessment of critical thinking: Can pharmacy students evaluate clinical studies like experts? *American Journal of Pharmaceutical Education*, *68*(1), 1-6. Retrievedfrom http://archive.ajpe.org/aj6801/aj680105/aj680105.pdf

Nunley, C., Bers, T., & Manning, T. (2011). *Learning outcomes assessment in community colleges*. National Institute for Learning Outcomes Assessment (NILOA). Retrieved from http://www.learningoutcomeassessment.org/documents/communitycollege.pdf

Oakleaf, M. (2009). Using rubrics to assess information literacy: An examination of methodology and interrater reliability. *Journal of the American Society for Information Science and Technology, 60*(5), 969-983. doi: 10.1002/asi.21030

Ohlemacher, J., & Davis, A. (2012). Community college strategies. *Assessment Update, 24*(3), 11-13.

Penrose, A. M., & Geisler, C. (1994). Reading and writing without authority. *College Composition and Communication, 45*(5), 505-520. doi: 10.2307/358762

Schilling, K., & Applegate, R. (2012). Best methods for evaluating educational impact: A comparison of the efficacy of commonly used measures of library instruction. *Journal of the Medical Library Association*, *100*(4), 258–268. doi: 10.3163/1536-5050.100.4.007

Swing, R. L., & Coogan, C. S. (2010). *Valuing assessment: Cost-benefit considerations* (NILOA Occasional Paper No. 5). Urbana: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment. Retrieved from http://learningoutcomeassessment.org/documents/SwingCoogan_001.pdf

Swoger, B. J. M. (2011). Closing the assessment loop using pre- and post-assessment. *Reference Services Review*, *39*(2), 244-259. doi: 10.1108/00907321111135475

Thompson, G., Pilgrim, A., & Oliver, K. (2005). Self-assessment and reflective learning for first-year university geography students: A simple guide or simply misguided? *Journal of Geography in Higher Education*, *29*(3), 403-420. doi: 10.1080/03098260500290959

Twait, M. (2005). Undergraduate students' source selection criteria: A qualitative study. *Journal of Academic Librarianship, 31*(6), 567-573. doi: 10.1016/j.acalib.2005.08.008

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134-144. doi: 10.1002/asi.10016

Zoellner, K., Samson, S., & Hines, S. (2008). Continuing assessment of library instruction to undergraduates: A general education course survey research project. *College and Research Libraries, 69*(4), 370–383. doi: 10.5860/crl.69.4.370

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 88 - 105.
doi: 10.14434/josotl.v15i1.12866

Development of teaching expertise viewed through the Dreyfus model of skill acquisition

Lucinda J. Lyon[1]

Abstract: This study was designed to explore development of skill acquisition in dental education, utilizing the Dreyfus and Dreyfus continuum. By identifying what skill progression may be recognized in the expert dental educator and what experiences appear to influence this growth, the knowledge gained may inform more efficient, effective faculty support, development and life-long learning. Employing a qualitative approach, individual interviews were conducted with experienced educators and analyzed. Open coding of responses revealed that skill acquisition necessary to good teaching, as expressed by these experienced educators, reflects common themes and a learning curve similar to that noted by Dreyfus and Dreyfus and other investigators. In addition, to supporting knowledge and technique development, dental faculty described working to share with students a wide range of non-cognitive competencies such as professionalism, communication, and an ethic of care and service. Findings increase understanding of teaching skill acquisition in dental education and may help provide support for health care faculty who desire to become excellent educators.

Keywords: skill acquisition, teaching methods, novice to expert, expertise, reflection

Introduction and Purpose

Former Cornell University President Frank Rhodes spoke to the inspirational nature of teaching when he powerfully described:

> Because of its profound impact upon both the individual student and society, teaching can never be just a job, however demanding; not even a career, no matter how professional. To the best professors, teaching is a moral vocation. It is moral because it seeks to develop not only comprehension, but also commitment; it influences and shapes not only the intellect, but also the will; it involves the cultivation of not only the mind, but also the heart … it is a vocation because it is a calling and not simply a job.
>
> Great teaching still has the power to inspire, to encourage critical but open outlook, a breadth of interest and a generosity of sprit. And students are challenged and inspired not by these qualities in the abstract, but by their embodiment in the professor (Rhodes, 2001, p. 67).

Today's most talented dental educators hope to make this impact on students. In pursuit of this goal, they embrace peer-reviewed standards, innovative teaching methods, topically integrated content, competency-based learning, authentic outcomes assessment, state-of-the-art simulation technology, and a growing commitment to interprofessional practice, and life-long

[1] University of the Pacific, Arthur A. Dugoni School of Dentistry, clyon@pacific.edu

learning (American Dental Education Association Commission on Change and Innovation in Dental Education, 2009; Hendricson, et al., 2007). Dental faculty are called upon to shape reflective practitioners, recognizable by their ability to solve well-formed problems by applying evidence based knowledge and technique, with the associated artistry and practical adaptability that make it truly valuable (Schon, 1983).

This unprecedented potential, however, is accompanied by well-documented challenges. Increasing graduate debt, as well as income disparity between private practice and a career in education challenges institutions as they compete to recruit and retain high quality faculty (Livingston et al., 2004; Haden et al., 2000). For these reasons, among others, a major source of new dental education faculty come from the practicing community, looking to education as a rewarding second career (Hand, 2006). Although these veterans bring the strength of contextual ease, most lack pedagogical knowledge. This growing paradigm tests the notion that the quality of dental students' educational experience is critically dependent on an adequate number of committed faculty, possessing both content and teaching methodology expertise (Hand, 2006; Berliner, 1988; Benner, 2001; American Association of Dental Schools, 1999).

Given the prevailing conclusion that performers in a variety of domains need approximately ten years of intense involvement before they reach peak performance (Ericsson et al., 1993; Ericsson, 2008) the challenge of faculty development, especially that of the second career dental educator, becomes evident. Dental schools must recognize and support development of expertise in this valuable human resource, as efforts to recruit and retain a strong faculty are intensified. The success of our students, the care of our patients, and the health of our dental education system depends on effective implementation of the above issues.

It is surmised that having a general theory about the development of expertise, and data about the ways novices versus experts perform pedagogical tasks, may contribute to policy considerations benefiting faculty development (Berliner, 1988). Professors Hubert Dreyfus and Stuart Dreyfus describe such a path to competence as a theoretical continuum of skill acquisition; a learning process moving through distinct stages from novice to expert, the highest level of skill demonstrated by fluid performance based on previous situations without obvious thought (Dreyfus & Dreyfus, 1986).

The purpose of this study was to explore development of teaching expertise utilizing the Dreyfus and Dreyfus construct of a continuum of skill acquisition (Dreyfus & Dreyfus, 1986). A product of philosophical deliberation and phenomenological research, the Dreyfus model was adapted by Benner (2001) to explore skill acquisition in nursing (Benner). For this reason, among others, it appeared a logical construct to use to explore development of educators in another health care profession. By identifying what skill progression may be recognized in the expert dental educator; and 2.) What experiences appear to influence this growth, knowledge gained may help schools more efficiently and effectively support faculty development and life-long learning. This paper begins with a theoretical framework of skill acquisition, provides and analyses qualitative data gathered through interviews of experienced dental educators, and explores skill progression revealed in this group of subjects. It offers both support of the literature, with regard to skill progression from novice to expert teaching, and findings that add to the literature. Lastly, questions for additional related research are suggested.

Theoretical Framework

Professional dental education shares many traditional expectations of its professoriate as do most professions. Expressed by Rhodes, these include "the conviction that teaching is a moral vocation, that scholarship is a public trust, that service is a societal obligation and that an independent and open community is the essential means to both learning and discovery" (Rhodes, 1999, p. 37). Boyer proposes that the work of the professoriate might be thought of as having four distinct, intersecting functions: the scholarship of *discovery*; *integration*; *application*; and, finally, the scholarship of *teaching*, inferring that activities become impactful only when understood by others (Boyer, 1990). To that end, growing attention is being paid to the scholarship of teaching, which Smith defines as entailing a "public account of some or all of the full act of teaching – vision, design, enactment, outcomes, and analysis – open to peer review and benefit" (Smith, 2001, p. 69-70).

Faculty in the health care professions are called upon to grow practitioners who are less reliant on the opinion of others and more ready to utilize critical thinking and problem-solving skills to evaluate evidence and arrive at an independent decision (Strohschein et al., 2002; Roth, 2007). These same acquired skills are equally important to strong pedagogical abilities (Haden et al., 2006; Trotman et al., 2007).

Skill Acquisition

Researchers Hubert and Stuart Dreyfus proposed a theoretical model of skill acquisition reflecting a "progression *from* analytic behavior of a detached subject, consciously decomposing his environment into recognizable elements, and following abstract rules, *to* involved skilled behavior based on an accumulation of concrete experiences and the unconscious recognition of new situations as similar to whole remembered ones" (Dreyfus & Dreyfus, 1986, p. 35). The five distinct stages of this learning process are described in Table 1.

Table 1

Five Stages of Skill Acquisition – Dreyfus and Dreyfus

Skill Level	Components	Decision	Commitment
Novice	Context-free	Analytical	Detached
Advanced Beginner	Context-free and situational	Analytical	Detached
Competent	Context-free and situational	Analytical	Detached understanding and deciding; Involved in outcome
Proficient	Context-free and situational	Analytical	Involved understanding; Detached deciding
Expert	Context-free and situational	Intuitive	Involved

Adapted from Dreyfus and Dreyfus (1986, p.50)

Novice: The novice stage is characterized by recognition of relevant discrete facts, features, and rules for decision-making that are so clearly and objectively defined as to be

virtually *context-free* (Dreyfus & Dreyfus, 1986). Lacking a coherent sense of the overall task, the novice treats each situation as new and looks for appropriate rules to follow. Because the learner is without an experiential base, a forecast of expected outcomes must be provided, which do not rely on prior experience (Benner, 2004). Description of how textbook examples and actual application of this information relate is necessary and important (Benner, 2004).

A teacher at this stage of development may use rules and guidelines for designing a syllabus, delivering a lecture, leading a discussion, and presenting feedback (Smith, 2001). The novice teacher would likewise judge their own performance based on their adherence to learned action steps (Flyvbjerg, 2001).

Advanced Beginner: Progression to advanced beginner occurs only with significant practical experiences, the learner begins to perceive similarity of concrete situations with prior examples of the same experience. Actions can now be based on both the new situation and the earlier context-free components (Dreyfus & Dreyfus, 1986).

No amount of verbal description provided at the advanced beginner level of learning is more productive than practical experience (Dreyfus & Dreyfus, 1986). However, the educator can assist the student by formulating principles that dictate a particular action, thus creating a set of guidelines (Benner, 2001). Advanced beginners are keenly attuned to feedback and intently focused on the example of colleagues and mentors (Benner, 2014). The teacher at the advanced beginner level is newly able to recognize a poor classroom climate, uninspiring lecture, or confused students, then relies on rules to remedy the situation (Smith, 2001).

Competent Performer: The competent performer places a growing amount of situational experience and context-free rules into a problem-solving format by adopting a hierarchical process of decision making. (Dreyfus & Dreyfus, 1986). Experts operate based on knowledge of several thousand such cataloged concrete cases in their area of expertise (Flyvbjerg, 2001). The speed with which competence increases is dependent upon the volume and complexities of learning experiences encountered (Benner, 2004).

While the competent performer may lack the efficiency of the more experienced performer, they are developing a growing ability to manage contingencies (Benner, 2001). Context becomes more critical. The learner develops a sense of when using rules is appropriate and when these rules provide diminishing returns and may be disregarded (Flyvbjerg, 2001).

The competent teacher chooses a goal or purpose to give focus to all information available and then works deliberately to simplify and solve the problem, anticipating or forecasting possible ramifications (Smith, 2001; Benner, 2004). Organizing curricula in ways that encourage conceptual understanding becomes beneficial to the student (Bransford et al., 2000). The new educator must begin to cultivate their ability to engage and interact with students, placing equal importance on developing knowledge, skills, and values (Benner, 2001).

Unlike the novice and advanced beginner, who react according to externally developed rules and, thus, feel little responsibility for outcomes, the competent performer becomes vested in outcomes resulting from his actions (Dreyfus & Dreyfus, 1986). Reflection and outcome evaluation becomes important at higher levels of the learning process (Schon, 1983; Smith, 2001; Flyvbjerg, 2001).

Proficiency: Someone at the proficient level "while intuitively organizing and understanding his task, will find himself thinking analytically about what to do" (Dreyfus & Dreyfus, 1986, p. 29). Situations are recognized simultaneously, judged to be similar or dissimilar to previous experiences, and acted upon in accordance with what actions have

achieved successful outcomes in the past (Flyvbjerg, 2001). The situation guides the practitioner's response (Benner, 2004).

Case-based learning is of particular benefit at the proficient stage, particularly if the learner is asked to present specific examples and experiences that lead them to their chosen path or conclusion. Reflection on both good and less than desirable outcomes can stimulate effective learning (Benner, 2001). The proficient teacher intuitively identifies a problem then consciously analyses options to remedy it (Smith, 2001). "Learning to teach is an ongoing process of observing, reflecting and experimenting" (Pinsky et al., 1998, p. 215).

Expertise: The expert is fully engaged in fluid, efficient performance, responsive to context, based on previous situations, without obvious thought. "When things are proceeding normally, experts don't solve problems and don't make decisions; they do what normally works" (Dreyfus & Dreyfus, 1986, p. 31; Chambers, 2012)). Responses are reactive rather than studied and premeditated. The expert performer engages in the situation, perceiving subtle changes and relying on a "creative search and cue sensitivity." Related to health care, based on genuine caring for the patient and ownership of outcomes, "It's not a question of choosing either science or practical wisdom, rather how to relate the two" (Benner, 2001, p. viii). In reflection, the expert focuses more on their intuition rather than the calculations they've made (Flyvbjerg, 2001).

Advances in neurophysics technology indicates that expert practitioners actually have "integrated neural networks that facilitate instantaneous retrieval of chains of knowledge relevant to task performance" which is very different from those of the novice (Hendricson, 2006, p. 5). Though this intuitive responsiveness may be an asset in many ways, an inability to break actions down into clearly verbalized, discrete elements can be an impediment to teaching novice students (Flyvbjerg, 2001). When experts are able to deconstruct and describe the significance of their actions, the student may more easily identify and learn from information embedded in their practice (Benner, 2001).

Deliberate Practice and Reflection

Ericsson (2008) describes an additional perspective, that expert performance may be traced to active engagement in deliberate practice, the process of continually stretching oneself with increasing goals and making time for serious self-assessment and reflection on personal performance (Ericsson).

Schon (1983) described, in complimentary terms, that professionals use a form of tacit experiential knowledge, which he terms *knowing-in-action* (Schon). He asserts that the most important competence, which aids the acquisition and continuous enhancement of all other competencies is *reflection*, the process of evaluating and learning from experience. He describes this activity taking place as *reflection-in-action* at the time one is engaged in an activity; and *reflection-about-action* after an activity (Schon, 1983). "Every attempt to produce an instruction is an experiment that tests both the coach's reflection on his own knowing-in-action and his understanding of the student's difficulty" (Schon, 1983, p. 104). Teaching "requires more than knowledge of theories and technical skills; it also requires analysis and reflective critique" (Smith, 2001, p. 76).

Methods

This investigation sought to explore the nature of expert educator's skill and work patterns and to define how they are developed and manifested from the point of view of the educators themselves, and those who have recognized their abilities and nominated them for participation in this study. Grounded theory methodology was employed, utilizing qualitative data gathered through independent subject interviews. The study was approved with Exempt Status by the Institutional Review Board at the University of the Pacific.

Qualitative research involves "the studied use and collection of a variety of empirical materials – case study; personal experience; introspection; life story; interview; artifacts; cultural texts and productions; observation, historical, interactional, and visual texts – that describe routine and problematic moments and meanings in individual's lives" (Denzin & Lincoln, 2003, p.5). "The qualitative researcher reviews social phenomena holistically [which] explains why [such] research studies appear as broad, panoramic, views rather than microanalysis" (Creswell, 2003, p.182). Wiersma and Jurs describe data analysis in qualitative research as "a process of categorization, description, and synthesis. Data reduction is necessary for the description and interpretations of the phenomenon under study" (Wiersma & Jurs, 2005, p. 207).

In a type of qualitative approach that is participatory the researcher seeks to examine an issue by collecting stories from individuals who are interviewed at length to determine their personal experiences (Creswell, 2003). Individual interviews were utilized in this study to explore perceptions, investigate in depth events, activities, processes, and one or more individuals (Creswell).

Grounded theory methodology allowed the data gathered in this research to be analizeded for identification of common themes. Struass and Corbin describe qualitative analysis as involving the "nonmathematical process of interpretation, carried out for the purpose of discovering concepts and relationships in raw data and then reorganizing these into theoretical explanatory scheme" (Strauss & Corbin, 1998, p.11). Theories or understandings emerge from the data gathered rather than data being to collected to confirm a pre-exisitng theory, as is more likely the case in guatatative studies (Charmaz, 2006, p. 5-6).

A purposeful sample was drawn from Academic Deans and experienced full-time dental school faculty in California. The literature (Ericsson, 1993) asserted that at least 10 years of practice was necessary to develop expertise so, for the purpose of this study, subjects studied had a minimum of ten years of teaching experience. This was defined as ten years or more of full-time commitment to graduate level teaching in a professional program, in this case dentistry.

Academic Deans, from the five dental schools in California who had students enrolled at the time of this study, were introduced to the subject and purpose of this research. They were invited to take part in the study in two ways 1.) to participate as the subject of an independent interview and 2.) to nominate faculty, with a minimum of ten years of teaching experience, whom they identified as expert dental educators, for interview. Subjects eventually included three academic deans and seven experienced dental faculty, nominated by their academic deans. The specific areas of teaching focus of these educators varied, including didactic and clinical instruction or both (Table 2). In addition to their current administrative responsibilities, the three deans included in this study had deep teaching experience and continued to present content to students.

Table 2

Subject Demographic Data

Name Code No.	Prof.	Assoc. Prof.	Discipline BS Basic Science CS Clinical Science	Dentist	Dental Specialty	PhD	Other graduate degrees MBA, EdD, etc.	Gender	Age	Ethnicity C Caucasian O Other
1	x		BS	x		x		M	50-59	C
2	x		BS, CS	x			x	M	50-59	C
3	x		BS			x	x	F	50-59	C
4	x		BS			x		F	60+	C
5		x	CS	x	x			M	50-59	C
6		x	CS				x	F	40-49	C
7	x		CS	x	x			M	60+	C
8	x		CS	x			x	F	60+	C
9		x	CS	x			x	M	40-49	O
10	x		CS	x		x		F	50-59	O

Detailed qualitative data was gathered via in-depth, open-ended interviews based on a standard set of questions. (See Appendix) Interviews, each approximately sixty minutes in length, were conducted by the primary investigator, either in person or by telephone. Although interview questions were slightly different for deans and faculty, responses were so similar that aggregating responses appeared more effective. Responses appeared to yield an in-depth description of the phenomenon being examined and enough variation in responses to illustrate a well-rounded picture. These were numbered to protect subject identity, transcribed verbatim, coded to identify and categorize segments of data, and further analyzed for identification of common themes.

Findings

Findings in the form of topics, dimensions, and categories describe skill progression perceived by these dental educators and provides a picture of experiences that appear to influence this growth process. Categories of *Basic Knowledge, Functional Skills, Personal and Behavioral Qualities*, and *Reflection* represent the main areas of observation or perception offered by subjects. The topics, which further emerged from the qualitative data gathered, serve to ground the categories noted (Table 3).

The *categories* were subsequently examined using the Dreyfus model to establish which were significant for different levels of novice to expert performance (Dreyfus & Dreyfus, 1986).

Table 3

Overview of Expert Faculty Codes

Category	Topic
A. Basic knowledge	
Dimensions	
A1. Content Specific	*a. Higher order content knowledge* *b. Current with developments in their field* *c. Habit of life-long learning*
A2. Application	*a. Loyalty to the learning process* *b. Current with methodology developments* *c. Ability to conceptualize, apply theory* *d. Ability to synthesize, transfer knowledge* *e. Aware of institutional curricular content and goals* *f. Personal growth stimulated by constructive tension* *g. Awareness of role models or mentors*
B. Functional Skills	
Dimensions	
B1. Occupation Specific	*a. Skilled in a variety of methodologies* *b. Clear goals and expectations* *c. Plans adequately* *d. Assesses outcomes – formative/summative* *e. Encourages student self-assessment*
B2. Process	*a. Considers the learner* *b. Draws on personal experiences* *c. Learning is the goal – rather than teaching* *d. Brings context to the subject matter* *e. Limits content – increases application* *f. Connects the new or unknown to the known* *g. Fosters critical thinking* *h. Helps student think like an expert* *i. Creates active learning opportunities* *j. Challenges students in a non-threatening way* *k. Innovates*
C. Personal and Behavioral Qualities	
Dimensions	
C1. Vocational	*a. Enthusiasm for the subject and the learning process* *b. Highly Motivated – enjoys challenge* *c. High standards* *d. Commitment to personal excellence and growth* *e. Willingness to take chances* *f. Flexible in action*
C2. Intraprofessional	*a. Values collegiality* *b. Stimulated by peers* *c. Strong citizen of the educational community*
D. Reflection	*a. Reflects upon personal performance* *b. Reflects upon process* *d. Reflects out loud with students*

Basic Knowledge

Dreyfus and Dreyfus (1980) define the first stage of learning, noted here by the term *Basic*, as context-free features which the beginner can recognize without benefit of experience" (p. 7). Acquisition of both basic content knowledge and the ability to apply it was recognized as the initial challenge faced by subjects. Interviewees recollected that their earliest teaching efforts revolved around mechanical presentation of unadorned content, almost to the exclusion of context. One subject confessed to:

> ". . . taking refuge in the facts" but noting that "providing that sort of [purely content] lecture was simple, but it was very unsatisfying. It quickly became boring, not only for the students, but for me". [1]

Acquisition of basic knowledge, and determination of its relevance within a growing sense of context, are among the earliest novice and advanced beginner experiences in the Dreyfus model, as is personal experience via trial and error. These seasoned educators almost universally described a beginning progression in concordance with this continuum.

As they brought greater context to the subject matter, faculty reported becoming better able to stimulate application of foundational knowledge. One recited a favorite quote from a 19th century mathematician to underscore the importance of context over memorization of discrete facts:

> A science is built of facts, just as a house is of bricks. But, a pile of facts is no more science than a pile of bricks is a house [1].

As they gained command of discipline knowledge and learned to prioritize information, faculty concentrated on teaching methodology. They began to approach presentation in a more holistic way, becoming increasingly aware of situational elements, including students' concurrent learning in other subject areas, influencing timely integration of their own content.

> If you hook students in with a clinical scenario, they pay attention to the facts. Whereas, if you put the facts first, they're kind of going, O.K., what's the point of all this *stuff*? [2]

From the perspective of recruiting professionals to academia, although these educators brought practical expertise in a given area, they did not commonly have formal education in basic teaching skills. This simultaneous differential where an individual may be at once expert with regard to certain problems and less skilled in other areas is recognized in the Dreyfus learning continuum. It is important to note that, at this early career stage, implicit expectations and formal, or informal, support of peers was deemed important to development.

Functional Skills

Dreyfus and Dreyfus (1980) propose that "competence comes only after considerable experience after coping with situations in which the student notes or the instructor points out recurrent, meaningful component patterns" (p. 8). For the purpose this study, this has been termed *Functional* knowledge. With initial exposure to teaching, higher order content knowledge developing, and growing utilization of varied teaching methodologies, awareness of the dynamics that students bring to the learning experience emerged. In the spirit of a decision-making and prioritization process, described by Dreyfus as typical of the competent performer, faculty reported the critical nature of recognizing *what* content was appropriate to the learner's

level of understanding and then paring information down to its most important concepts; *Less is more* being the overarching sentiment.

Linking new concepts to known concepts was deemed critical, as was helping students differentiate pertinent from non-pertinent information. Describing their own personal growth being stimulated by constructive tension, there was apparent agreement that the learning environment, and the teacher, should challenge students in a non-threatening way. The ability to personalize teaching and learning approaches was believed especially important in situations where the student was having difficulty grasping a concept. These teachers challenged themselves to bring new ideas and concepts to their teaching, to innovate.

In the sense that the Dreyfus competent beginner struggles with the choice of plan and subsequently feels responsible for the outcome, assessment of learning outcomes assumed growing importance as these educators became more experienced. Although they conveyed pragmatic plans for the learning experiences, subjects' expression of outcomes alternated between highly structured and more broad and theoretical. All agreed that assessment was critical not only to confirming student learning, but also to developing strong teaching skills, informing progressive refinements as their careers progressed.

> It's my opinion that learning is internal and individual, that each person comes to it in their own way. They have to have some commitment and involvement. So, the outcomes I look for are changed behavior in clinic. I look for changes to approaches in certain problems. [6]
>
> I have a great deal of contact with students, especially in the lab. I listen for their conversations, check their reactions, and look at their faces. I depend on that a lot. Students appreciate the opportunity to demonstrate what they know. [1]

Finally, helping learners self-assess was a recurring theme. With experience, subjects described a paradigm shift in their thinking about the purpose of their work. Student learning, rather than their own teaching, became their goal. While they understood more naturally the outcomes they hoped to achieve, they continued to reason analytically about how to obtain these.

With progression, subjects became better able to read and guide the learning experience, making adjustments smoothly and intuitively. They came to view students, and the learning process, in a more holistic fashion and felt responsible for, and deeply vested in, the success of both. As they assumed the level of proficient performers, although their actions became increasingly intuitive, they continued to reflect analytically on how to improve the process further, as the Dreyfus model might predict.

Personal and Behavioral Qualities

Personal and behavioral qualities appeared to compliment functional skills. Enthusiasm for subject matter, the learning process, and the challenges of staying abreast of both was displayed frequently by subjects.

> Dentistry is so much more complex. 3D imaging will change the face of my specialty, as well as dentistry. The microscope has already [done that]. Technology has changed my specialty dramatically over the last ten years. So, it requires you to continually look for efficiencies in the way you deliver [content] because there's so much more to deliver. You can't get lazy. [7]

Experts displayed a commitment to personal excellence and growth. They held themselves and their students to a high standard, technically, morally, and ethically, and shared

the rigors of their own personal development with students. These faculty seemed to take as a given that becoming professionally accomplished required not only high expectations, but a willingness to experiment, to be flexible in action, and take chances. One subject alluded to his increasingly intuitive ease with a more dynamic learning environment and his willingness to let class sessions unfold organically:

> I plan less; my lectures have become much more organic. I just begin my story and see where it takes us. I have less need to know all of the answers. [1].

The group alluded to the importance of their community of learners. They appeared to interact intimately with peers, and to value the power of role models, mentoring, and camaraderie. With palpable affection and allegiance, several subjects spoke to this community:

> I love the university atmosphere, there's no doubt. That kind of higher level thinking, when you can find it, is very stimulating. I gravitate towards individuals who constantly challenge your paradigms. [7]

Interviewees also held a high level of interest and concern for education beyond their personal classrooms, expressing a number of philosophical concerns, including patients' ability to receive services, academia's responsibility to provide care today, and to graduate professionals who will shape the future practice environment. Dreyfus and Dreyfus (1986, p. 30) reason that expert performers "see themselves as involved participants in a world of opportunities, threats, strengths, weaknesses, hopes, and fears." This concept appears to be confirmed by these senior educators whose professional views have become broad and holistic.

Reflection

Reflection appeared important to the growth of these individuals. Whether regarding personal performance, or the process of teaching, the cohort carefully considered their teaching and how to improve it. By design or example, most subjects incorporated expectations for student self-assessment and reflection in their learning plan. Some reflected out loud to help students learn by example. Dreyfus and Dreyfus (1986) reasoned that even within the context of a fluid, rational erformance, in order for continued learning to occur, some portion of the mind must remain detached and observational. It may be argued that without intentional reflection and improvement of process and outcome, skill acquisition may plateau.

An overview of the codes recognized — *Basic Knowledge, Functional Skills, Personal and Behavioral Qualities,* and *Reflection* — were viewed with the Dreyfus model of novice to expert development (Dreyfus & Dreyfus, 1986). All five stages of the Dreyfus continuum from novice to expert were recognized (Table 4).

While outcomes find significant confirmation in the literature, study limitations include the following: relatively small sample size, limited number of schools, and subject self-reporting. Arguably the culture in which subjects teach may have an effect upon their development.

Future studies might include larger sample size, variety of schools and direct observation to confirm findings with actual behavior. Interviewing for self-reported perceptions is a direct approach for data gathering and provides important first step in identifying additional relevant issues.

Table 4

Expertise in Dental Education: Codes Interpreted Relative to the Dreyfus and Dreyfus Model of Skill Acquisition

Novice: The novice recognizes clearly, objectively defined context-independent elements which are processed using externally provided rules and procedures to determine action. Difficulty forecasting outcomes based on lack of experience.

- Reliant on predetermined curriculum, course structure, and independent content elements. Information delivered without significant context
  (Codes: A1-a, A1-b)

Advanced Beginner: The advanced beginner starts to recognize situational elements through experiential learning. External prompts remain important; however the advanced beginner is broadening understanding through experience which, at this stage is vastly more important than verbal description. Beginning to learn from mistakes.

- Reliant upon highly organized presentation of material, teaching experiences contribute to learning, observes experts in action
  (Codes: A2-b, A2-f, A2-g, C1-a)
- Recognizes the elements the learner brings to the environment
  (Codes: B2-a, B2-b)

Competent Performer: The competent performer develops the ability to prioritize overwhelming volumes of information and procedures in order to simplify and improve rational decision making. At this stage, the competent performer perceives responsibility and is vested in the outcomes of his performance.

- Gaining ability to conceptualize, synthesize, apply knowledge
  (Codes: A2-a, A2-c, A2-d)
- Methodologies are becoming internalized; information is prioritized; a plan is created
  (Codes: B1-a, B1-b, B1-c, B1-d, B2-e, B2-f)
- Assessments are structured analytically, outcomes become of increasing value; reflects on progress
  (Codes:B1-e, Da, Db)

Proficient Performer: The proficient performer intuitively recognizes, organizes, and understand his task. The proficient performer continues to analyze options and make decisions about how to proceed based on reason

- Sees curricula more holistically, able to integrate subject matter
  (Codes: A2-e, B2-i, Dd)
- Realizes the ability to make a difference. Learning becomes the goal rather than teaching.
  (Codes: B2-c, C1-b,)

Expert: The expert performer intuitively recognizes his task and fluidly reacts. They react fluidly and unconsciously. Rather than making decisions, they do what works based on experience and understanding.

- Teaching becomes instinctive, presentations more spontaneous, assessments more intuitive. Creates constructive anxiety in the classroom.
  (Codes: B2-g, B2-j, B2-k, C1-f)
- Methods include more experimentation. Understands that outcomes are not guaranteed
  (Codes: C1-d, C1-e)
- Understands the importance of being a role model, inspiring the learner. Concerned for patient as well as learner – communication with both is fluid and intuitive. Values ethic of patient care.
  (Codes: B2-h, C1-c C2-c, Dc)

Conclusion

The goal of this study was to explore skill progression recognizable in the expert dental educator and to learn more about the experiences supporting progression from novice to expert teaching.

Findings that Find Confirmation in the Literature

Results of this study indicate that novice dental educators experience many of the challenges described by Dreyfus and Dreyfus and others. Command of clinical expertise, which many new or second career faculty possess, is not the same skill as that of an educator. Simple accumulation of experience was the primary benefit to subjects' earliest teaching, followed by a growing recognition of context, environment, and learners' response. Novice faculty had simplistic ideas of how their content fits into the whole and were dependent on the example of mentors and role models (Schon, 1983; Haden et al., 2000; Berliner, 1988; Benner, 2001; Boyer, 1990; Smith, 2001; Haden et al., 2006; Trotman, 2007; Bransford et al., 2000; Pinsky et al., 1998; Hovland, 2002; Richlin, 2001; Forrest, 2006; Johnson & Ridley, 2004; Schenkein & Best, 2001; Wright & Carrese, 2002).

As teaching skills progressed, faculty became increasingly aware of outcomes and their importance to the learning process. They experimented with assessment methods including guiding learner self-assessment. The process of consistent, formal outcomes assessment was viewed as critical to faculty development, as confirmed by a number of researchers (Dreyfus & Dreyfus, 1986; Ericsson, 2008; Smith, 2001; Flyvbjerg, 2001; Pinsky et al., 1998; Hovland, 2002). Development of problem solving and critical thinking skills helped move both these developing faculty and their students from dependent to more self-directed life-long learning and action (Schon, 1983; Strohschein et al., 2002; Roth, 2007; Haden et al., 2006; Benner, 2004; Bransford et al., 2004; Hendricson et al., 2006). Many faculty also referenced constructive tension, within a humanistic atmosphere, as productive to personal growth and satisfaction in the academy, concepts supported in the literature (Rhodes, 1999; Haden et al., 2006; Pinsky et al., 1998). They voiced that personal growth and pivotal progression often occurred when they were forced to meet new challenges under stressful circumstances.

Finally, virtually all subjects displayed *reflection-in-action* and *reflection-on-action* described by Schon (1993). In the way that Ericsson (2008) reasoned that expert performers avoid arrested development by consistently striving for higher levels of achievement through deliberate practice, these faculty worked to continually improve as educators.

Findings that add to previous studies

Some unique challenges present themselves in the course of educating dental and other professional healthcare students. In addition to technique development, teachers convey non-cognitive, interpersonal competencies such as professionalism, communication, and an ethic of care and service. The novice educator may benefit from teaching strategies displayed and described by subjects:

a) Stress context early in the learning process – excellent patient care is the goal
b) Expect students to apply foundational knowledge to clinical care, to make evidence based decisions

c) Value and reinforce critical thinking, organizing information, and problem solving
d) Provide accessible skill-appropriate learning opportunities for the student
e) Engage students as actively as possible
f) Challenge students in a humanistic way
g) Assess formative and summative learning outcomes, both formally and informally
h) Help students recognize skills embedded in expert practice and connect these demonstrations with their own personal delivery of clinical care
i) Share personal experiences to connect explicit information with practical wisdom
j) Experiment with teaching methodologies to continually improve learning outcomes
k) Hold high expectations of students' capacity for learning and skill development

In addition to technical expertise, experienced faculty displayed personal and behavioral qualities, which enhanced their teaching and work satisfaction:

a) Enthusiasm for the learning process
b) Extraordinary levels of motivation
c) High standards and a commitment to personal excellence and growth
d) Passion for the academe, a genuine commitment to their peers and the educational community.

Dreyfus and Dreyfus asserted that a novice with inherent ability, given the opportunity to acquire a critical amount of experience, may become an expert. The experts studied additionally had highly-valued relationships with students and peers, and a positive learning environment. Professional challenge, stimulation, and ever-changing opportunities to develop mastery were equally motivating. Findings support supposition about qualities that attract an individual to education: professional challenge, a stimulating environment, endless variety of task, and opportunity for growth (Bertolami, 2007; Trower, 2007).

The collective reflections gained through interviews with expert dental educators provide an emerging profile of teaching expertise in dental education consistent with the Dreyfus skill acquisition continuum. The data gleaned opens conversation and prompts the following questions:

1.) How might development opportunities be better targeted to the educator's level of skill acquisition?
2.) How might faculty be better trained to utilize and model critical thinking?
3.) Would added focus on student learning assessment help educators advance more effectively?
4.) How might guided reflection reinforce positive qualities that inspire, challenge, and satisfy the individual educator?
5.) Do study findings have implications for faculty recruitment and retention strategies?
6.) Are there comparisons that should be explored between Novice to Expert stages of development and the traditional promotion and tenure model, timeline, and emphasis?

Future research on these questions may build upon the findings revealed.

Subjects described personal growth, skill acquisition, attributes, assessments, and aspirations integral to development of teaching expertise in dental education. The Dreyfus model appeared descriptive of the learning sequence of the dental educator. Outcomes of this study may prove helpful to those actively involved in recruiting, challenging, developing, supporting, and enjoying life-long collegial relationships with dental and other healthcare educators at a myriad of points along their professional learning continuums.

Appendix

OPEN-ENDED STANDARD INTERVIEW QUESTIONS

Interview Questions - Academic Dean Participants:

1. What attracted you to dental education?
2. What were your own milestones as an educator?
3. How do you recognize potential in new dental educators?
4. What faculty skills are important to student development?
5. How do you identity faculty with expertise?
6. Are there recognizable student learning milestones that denote faculty expertise?
7. What do you believe are the greatest challenges for dental educators?
8. Are there new challenges today in training students to become safe beginners?
9. What advise would you give a beginning educator?

Interview Questions – Dental Faculty Participants:

1. What inspired you to become a dental educator?
2. What are the core strategies you use to be an effective teacher of content?
3. What strategies do you use to be an effective teacher of students?
4. How do you know learning is going on?
5. Has your teaching changed over time? How?
6. What baseline skills, attitudes, or behaviors of expertise can you recognize in new students?
7. What advise would you give a beginning educator?
8. Is there anything else you would like to tell me?

References

American Association of Dental Schools. (1999). *Future of dental school faculty: Report of the president's task force*. Washington DC: American Association of Dental Schools.

American Dental Association, Commission on Dental Accreditation. Accreditation. (n.d.). Retrieved from: http://www.ada.org/117.aspx

Benner, P. (2001). *From novice to expert: Excellence and power in clinical nursing practice*. Upper Saddle River, New Jersey: Prentice Hall.

Benner, P. (2004). Using the Dreyfus model of skill acquisition to describe and interpret skill acquisition and clinical judgment in nursing practice and education. *Bulletin of Science, Technology and Society,* 23(3), 188-199. doi: 10.1177/0270467604265061

Berliner, D. (1988). *The development of expertise in pedagogy*. Washington, DC: American Association of Colleges for Teacher Education Publications.

Bertolami, C. (2007). Creating the dental school faculty of the future: A guide for the perplexed. *Journal of Dental Education, 71*(10), 1267-1280.

Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Chambers, D. W. (2012). Dental education's involvement with dentists' learning in practice: Data and theory. *Journal of Dental Education, 76,* 107-117.

Charmaz, K. (2001). *Constructing grounded theory: A practical guide through qualitative analysis*. London: Sage Publications.

Creswell, J. (2003). *Research design, qualitative, quantitative, and mixed methods approaches,* 2nd edition. Thousand Oaks: Sage Publications.

Denzin, N., & Lincoln, Y. (Eds.). (2003). *The landscape of qualitative research, theories and issues*. Thousand Oaks: Sage Publications

Dreyfus, H., & Dreyfus, S. (1980). *A five-stage model of mental activities involved in directed skill acquisition*. (Supported by the U.S. Air Force, Office of Scientific research (AFSC) under contract F49620-C-0063 with the University of California) Berkeley.

Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Blackwell Publishers.

Ericcson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363-406. doi: 10.1037/0033-295X.100.3.363

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine*, *15*, 988-994. doi: 10.1111/j.1553-2712.2008.00227.x

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511810503

Forrest J. (2006). Treatment plan for integrating evidence-based decision making into dental education. *Journal of Evidence Based Dental Practice, 6(*1), 72-78. doi: 10.1016/j.jebdp.2005.12.022

Haden, K., Andrieu, S., Chadwick, D., Chmar, J., Cole, J., George, M., et al. (2006). The dental education environment. *Journal of Dental Education, 70*(12), 12165-12170.

Haden, K., Beemsterboer, P. L., Weaver, R. G, Valachovic, R. W. (2000). Dental school faculty shortages increase: An update on future dental school faculty. *Journal of Dental Education*, *64*(9), 657-673.

Hand, J. (2006). Identification of competencies for effective dental faculty. *Journal of Dental Education*, *70(*9), 937-947.

Hendricson, W., Anderson, E., Andrieu, S., Chadwick, D., Cole, J., George, M., et al. (2007). Does faculty development enhance teaching effectiveness? *Journal of Dental Education*, *71*(12), 1513-11533.

Hendricson, W., Andrieu, S., Chadwick, D., Chmar, J., Cole, J., George, M., et al. (2006). Educational strategies associated with development of problem-solving, critical thinking, and self-directed learning. *Journal of Dental Education*, *70*(9), 925-936.

Hovland, E. (2002). Introduction: Dental school's relations with organized dentistry and accreditation: The Gies report reconsidered. *Journal of the American College of Dentists*, *69*(2), 13-17.

Johnson, W., & Ridley, C. (2004). *Elements of mentoring*. New York. Palgrave Macmillan.

Livingston, H., Dellinger, T., Hyde, J., & Holder, R. (2004). The aging and diminishing dental faculty. *Journal of Dental Education*, *68*(3), 346-354.

Pinsky, L. E., Monson, D., & Irby, D. M. (1998). How excellent teachers are made: Reflecting on success to improve teaching. *Advances in Health Sciences Education*, *3*, 207-215. doi: 10.1023/A:1009749323352

Rhodes, F. H. T. (1999). *Thoughts on the American university at the dawn of the third millennium*. Retrieved from: http://theuniversityfaculty.cornell.edu/forums/pdfs/Rhodes.pdf

Rhodes, F. H. T. (2001). *The creation of the future: The role of the American university*. Ithaca, NY: Cornell University Press.

Richlin, L. (2001). Scholarly teaching and the scholarship of teaching. *New Directions for Teaching and Learning, 86,* 57-68. doi: 10.1002/tl.16

Roth, K. (2007). Dental education: A leadership challenge for dental educators and practitioners. *Journal of Dental Education*, *71*(8), 983-987.

Schenkein, H., & Best, A. (2001). Factors considered by new faculty in their decision to choose careers in academic dentistry. *Journal of Dental Education*, *65*(9), 832-840.

Schon, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Smith, R. (2001). Expertise and the scholarship of teaching. *New Directions for Teaching and Learning*, *86*, 69-78. doi: 10.1002/tl.17

Strohschein, J., Hagler, P., & May, L. (2002). Assessing the need for change in clinical education practices. *Physical Therapy, 82*(2), 160-172.

Struass, A, L., & Corbin, J. M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks, CA: Sage Publications.

Trotman, C., Haden, K., & Hendricson, W. (2007). Does the dental school work environment promote successful academic careers? *Journal of Dental Education, 71*(6), 713-725.

Trower, C. (2007). Making academic dentistry more attractive to new teacher-scholars. *Journal of Dental Education, 71*(5), 601-605.

Wiersma, W., & Jurs, S. (2005). *Research methods in education: An introduction*, 8th Edition. Boston: Pearson A and B.

Wright, S. M., & Carrese, J. A. (2002). Excellence in role modeling: Insight and perspectives from the pros. *Canadian Medical Association Journal*, *167*(6), 638-643.

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 106 - 107.
doi: 10.14434/josotl.v15i1.13318

Book Review

Motivating and Retaining Online Students: Research-Based Strategies That Work

Peggy Ann Everett[1]

Citation: Lehman, R. M., & Conceição, S. O. (2014). Motivating and retaining online students: research-based strategies that work. San Francisco: Jossey-Bass, A Wiley brand. ISBN: 978-1-118-53170-9

Publisher's Description: Finally, the first research-based book of sound strategies and best practices to help instructors motivate students to complete their online courses.

Although studies support the effectiveness of learning online, students often fail to complete online courses. Some studies have found that as many as 50–70% drop out of their online courses or programs. Retention is not only a growing expectation and imperative, but it is also as opportunity for faculty members to take the lead in innovating, researching, and implementing new strategies while demonstrating their effectiveness.

Designed for instructors and instructional designers, *Motivating and Retaining Online Students* is filled with empirical research from the authors' study of motivation and retention strategies that can reduce online learner dropout. Focusing on the most important issues instructors face, such as course design; student engagement and motivation; and institutional, instructional, and informal student support strategies, the book provides effective online strategies that help minimize student dropout, increase student retention, and support student learning.

While helping to improve the overall retention rates for educational institutions, the strategies outlined in the book also allow for student diversity and individual learner differences. Lehman and Conceição's proven model gives instructors an effective approach to help students persist in online courses and succeed as learners.

The demand for online education keeps growing. It was once seen as a lesser form of education, but there are some who want to offer online classes that are on par with face-to-face classes. Just how to do that well is the question. Teaching online courses frustrates many facilitators whether just beginning or seasoned veterans of the field. Communicating dialogue and taking students through learning exercise in the virtual world leaves something left to be desired for many.

Lehman and Conceição gathered extensive data to compile their third book together. Between the two, the authors have decades of experience in the field and offer what comes across as genuinely helpful, well researched explanations and suggestions in taking a student-centered approach to teaching online courses. They detail why students are successful (or not so

[1] Texas State University, 601 University Drive, San Marcos, TX 78666

successful) in online learning. The first of five chapters explain why online education is a growing field and gives new ways of looking at learning as well as the changing characteristics of the typical student. They clarify that many online students are nontraditional and have other obligations like work and family (Lehman & Conceição, 2014, p. 2). Many students do not do well, or feel like they will not do well, in an online learning environment. The authors explain this tendency towards self-fulfilling prophesy, concerns for student retention in an online setting, and reasons students leave classes as well. Juxtaposing these ideas, the authors then explain why some students succeed in online classes and what characteristics these students exhibit. The second chapter offers strategies for designing courses. This chapter focuses on motivation and finding out about learner needs. When the facilitator builds motivating learning activities into the course, students are able to reward themselves in completing one activity, and inertia propels them to subsequent learning activities. Chapter three examines motivating students. It focuses on self-motivation. One particular point the authors make in this chapter stood out to me. They explain that self-efficacy is a key to student success in online learning environments and offer ways in which instructors may incorporate the strategies into their course designs. The authors suggest a scavenger hunt during orientation to familiarize students with the online experience as well as the learning management system (Lehman & Conceição, 2014, pp. 40-41). This is something that can be easily incorporated into courses regardless of the course content. The fourth chapter details supporting students through Human Resources, institutional support, and self-care. The fifth chapter is more of a summary chapter and attempts to pull all the strategies given together.

In addition to the Glossary, the book includes two appendices. The first supplemental text is a survey for students while the second is for the instructor. The student survey is overtly lengthy and students, especially undergraduate students, may be intimidated by the length. This could cause some students anxiety, or instructors may receive partial data. If a learning needs and resource assessment is to be given in a class, instructors should probably tailor it to the class rather than relying heavily on the one included in this text. The second supplemental text, the instructor's survey, is much shorter and probably more helpful. This reinforces the implementation of the strategies given in the text into the course design.

Lehman and Conceição take a straightforward approach. The text is written in accessible language, the charts are clearly explained, and the suggestions in course design are helpful. They explain the problem, how we should address that problem, and give solid evidence on how their approaches work. The research is well-founded, and the strategies are sound. In a second edition, Lehman and Conceição may want to consider addressing issues or strategies conderning cheating or plagiarism in online courses. This seems to be a constant concern for online instructors. They could include information on the accessibility of internet/computers to students, especially low socioeconomic status students who may not have access to either in the home or within a reasonable distance. Some students may take an online course expecting to complete it on a smart phone, commuting to campus, or going to a public library for resources. This may lead to concerns about deadlines or even basic course information. Lastly, an additional supplemental text, like a basic timeline, outline, or calendar for six-, eight-, and sixteen-week courses may help students and instructors immensely. Organization remains key to doing well in online learning. Overall, the text contains useful suggestions for improvement in the field of online learning. The last thing busy instructors need is a long, drawn out text. Nothing in this book is superfluous, and it can be read in a relatively short timeframe, a good investment for a busy profession.

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

Journal of the Scholarship of Teaching and Learning, Vol. 15, No. 1, February 2015, pp. 109 – 111.

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.

General Guidelines for the Manuscript

Submissions should be double-spaced. The final manuscript should be prepared in 12-point, Times New Roman, and single-spaced. All margins should be 1 inch. Justify lines; that is, use the word-processing feature that adjusts spacing between words to make all lines the same length (flush with the margins). Do not divide words at the end of a line, and do not use the hyphenation function to break words at the ends of lines. 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. Please use the footnote function of your word processing program; there are a variety of instructions available online for each program. The abstract should be indented 0.5" left and right from the margins, and should be in italics.

Indent the first line of every paragraph and the first line of every footnote; all first line indentations should be 0.5". 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 keywords.

Pages should be unnumbered since they will be entered by the JoSoTL 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 author's 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.

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.

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.

Section and Sub-Section Headings

Major Sections

Major section headings should be centered and bold-faced (i.e., Section and Sub-Section Headings as seen above). Major section headings should have one-line space before and after. The first paragraph(s) of the article do not require a major heading.

Sub-Sections

Sub-section headings should also be flush-left and bold-faced. 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.

Sub-subsections. Sub-Subsections of your manuscript should be formatted like this.

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, with data entries single-spaced and numbered. Column labels should be half-line spacing above data. Please use the table functionality in your word-processing program rather than adding an image of a table from MS Excel, SPSS, etc. This allows for more flexibility in laying out the final print version.

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. The Editorial staff may adjust layout to allow optimal use of space.

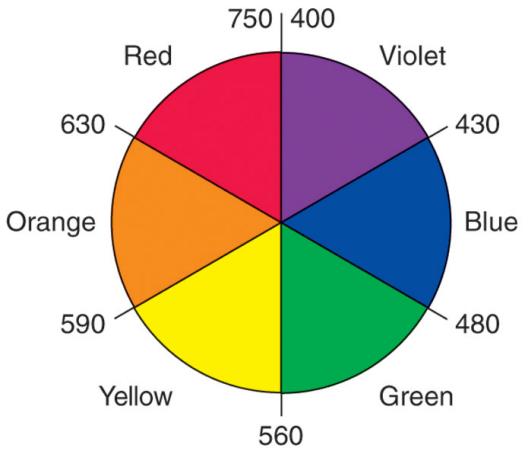


Figure 1. Color wheel with wavelengths indicated in millimicrons. Opposite colors are complementary.

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. If your submission has only one appendix, this section should be labeled '*Appendix.*' More than one appendix will change the section label to '*Appendices.*' Each appendix should have a title; if you are including items from your class or research, please alter them to include a title. Appendices should be alpha-order (Appendix A, Appendix B, etc.) These labels and titles should be at the top of the page, left justified, italicized.

Appendix 1. The Title of the Appendix.

The content of your appendix will appear here.

References

Coffin, D. A. (1993). Using the competitive edge. *Journal of Economic Education*, *24*(1), 62-69.

Garcia, J., & 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. doi: 10.1177/152700250200300103

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