RESEARCH & PRACTICE IN ASSESSMENT

RPA

Editors

Executive Editor

Keston H. Fulcher
Assisstant Professor and Associate Assessment Specialist
James Madison University

Editor

Allen DuPont
Director of Assessment, Division of Undergraduate Affairs
North Carolina State University

Consulting Editors

Robin Anderson
Associate Director, Center for Assessment and Research Studies
James Madison University

Dorothy Doolittle
Professor of Psychology
Christopher Newport University

John T. Willse
Assistant Professor of Educational Research Methodology
University of North Carolina at Greensboro

Table of Contents

Comments from the Editor

Keston H. Fulcher
Associate Assessment Specialist/Assistant Professor
James Madison University

This issue features two articles related to assessment practice. The first by Sally Sledge and Pam Pringle not only summarizes studies investigating "honor" at Christopher Newport University, but also documents how the results were used to inform decision making. This work underscores the nexus between assessment and policy.

In the second article, Chris Orem and I explore operationalizing good assessment practice for academic degree programs. Specifically, James Madison University uses a behaviorally anchored rubric to convey the strengths and weaknesses of each program's assessment report. This assessment of the assessment is sometimes referred to as "meta-assessment."

Assessing Honor Code Effectiveness:
Results of a Multipronged Approach from a Five Year Study

Sally Sledge
Norfolk State University

Pam Pringle
Christopher Newport University

Author Note

Correspondence concerning this article should be addressed to either Sally Sledge, Brown Hall A-227, School of Business, Norfolk State University, 700 Park Avenue, Norfolk, VA 23504. Email: sasledge@nsu.edu or Pam Pringle, Room 401 BTC, Luter School of Business, Christopher Newport University, 1 University Place, Newport News, VA 23606. Email: ppringle@cnu.edu

Abstract

This paper describes an ongoing study at a small public university to assess student, faculty and alumni perceptions of academic integrity and business ethics. The phases of the research are detailed for application or replication by other institutions of higher education. The study involves anonymous surveys and academic integrity-based interventions to assess the opinions of ethics and the honor code among students and faculty. Alumni were surveyed regarding the impacts of the honor code on their behavior as students and as employees after graduation. The results show that interventions improved familiarity with the honor code and knowledge of academic integrity among students, faculty and alumni. Directions for continued assessments to include the aforementioned groups as well as administrators and staff are given.

Introduction

Era of Scandals and Fraud

With the accelerated number of fraud cases and unethical business practices rising in the corporate world over the last few decades, more pressure is being put on colleges and universities to include courses and training in ethics and integrity. As a consequence, many institutions of higher education are renewing their focus on promoting ethical decision making, often via academic integrity programs and honor codes (Gilbert, 2008). These codes can attract students and set the tone for the campus culture.

This paper details the case study of an ongoing assessment of the opinions and perceptions as well as the procedures and actions regarding academic integrity and the honor code at a small public university in the southeastern United States. First, current issues are discussed and the climate of ethics within academia is addressed. Then the rationale behind the study is given. Next the different assessment tools are explained. Multiple assessment measures are used in order to gain a holistic picture of the environment at the university. The results follow. Limitations and plans for additional assessments are given. Finally conclusions are drawn.

Problems in Academia

Many students acknowledge having honor codes in high school and even in middle school. These seem to be more common in private secondary schools, but their numbers are growing in public schools as well. Despite the codes, many high school students admit to cheating and plagiarism due to parental and teacher pressure to do well (Oleck, 2008). These students feel that if they do not make good grades, their chances to attend a good college are ruined. Added pressures from extra curricular activities, leadership roles, volunteerism, and sports combine to cause students to engage in unethical behaviors (Kisamore, Stone, & Jawahar, 2007). In fact, according to McCabe and Katz (n.d.) 74 % of high school juniors and seniors surveyed across 22 public schools identified at least one incidence of cheating in exams during the previous year and 59 % reported incidences of plagiarism. A survey of 30,000 high school students by the Josephson Institute of Los Angeles reported that 64 percent had cheated the previous year on a test (Gerold & Steinberg, 2009).

In college, technology has allowed cheaters to be quickly identified through websites such as

turnitin.com and essayrater.com. Some high profile cases have brought this issue to light, such as the 45 physics students at the University of Virginia who were dismissed as a consequence of the school's single sanction, student run honor code in 2002. Another case at Duke University in 2007 found 34 graduate business students cheated on a take-home exam. Punishments from the faculty and student judicial board ranged from receiving a failing grade on the exam to a failing grade in the course to a one year suspension to expulsion (Damast, 2007).

Despite the increased opportunities for cheating, some schools are making strides towards a culture of integrity, including Ohio Northern University, where all business students agree to uphold an honor code during a swearing in ceremony (Gilbert, 2008). The internationally recognized Thunderbird School of Global Management in Arizona has created a Professional Oath of Honor for its students, similar to the Hippocratic Oath that doctors use. In 2006, Penn State instituted a newly formed honor committee composed of students and faculty to promote ethical behavior on campus (Damast, 2007).

Yet we believe that many schools embrace a "head in the sand" mentality regarding honor and integrity. Some universities do not assess their environment and therefore do not know of the existing problems regarding cheating and unethical behavior. This case study can serve as a guide for those universities and colleges that are looking to implement programs and assessments of academic integrity.

Rationale

An article in the university student newspaper that detailed honor code violations prompted our interest in the fall of 2004. There were disparities in the number of violations per major across campus. Notably, nearly one-third of honor code violations occurred in the business school, yet it accounted for only 18% of the student body (Captain's Log, 2004). This disproportional cheating by business students is not unique to our institution. McCabe, a Rutgers business professor who has studied cheating among college students for 18 years, found that 56% of business students admitted to cheating at least once versus 47% in other areas of study. He notes a "bottom line mentality" where students are most concerned with "getting the job done" (Graves, 2008). To further explore this phenomenon we polled the faculty to get a baseline assessment of their perceptions and enforcement of the honor code. The initial research questions were as follows. 1. Are faculty knowledgeable about the Honor Code and its administration? 2. What are faculty perceptions of the Honor Code? 3. What educational resources would make the Honor Code stronger?

Methodology and Results

Given that this article summarizes a five-year research project and covers several mini-studies, we thought it more appropriate to combine the methodology and results sections and present them according to the chronology of the project. The sequence of steps across the project is as follows:

1. Initial, exploratory faculty assessment
2. Utilization of the Center for Academic Integrity's (CAI) national survey instrument for faculty and students
3. Research of ethical best practices in businesses and universities
4. Creation and implementation of interventions based on the data collected in steps two and three
5. Follow-up survey of faculty and students, post interventions
6. Evaluation of results
7. Involvement of students in the research process.
8. Expansion of assessment to alumni who had experienced the interventions as students
9. Broadening the research to include high school students and faculty and their perceptions of academic integrity

Throughout these steps we have purposefully shared our findings on campus as well as nationally through conference presentations and papers. Both students and faculty have been involved in these presentations. Our intent is to continue the assessments in partnership with CAI to further promote a culture of academic integrity within the university.

Initial Study Results from Faculty

The initial Fall 2004 faculty assessment included 30 faculty who responded to an in-house exploratory survey that incorporated the research questions. All faculty within the business school were invited to participate, and did so, and were informed of the goals of the survey. Anonymity was assured. The survey was piloted before it was given to faculty; no major problems with the survey were found. This initial assessment illustrated that faculty rarely promoted the honor code to students. Relevant findings are listed in Table 1.

Table 1
Initial Faculty Assessment Data – 2004

Topic	Level
Faculty discussion of the honor code in class	Low
Faculty inclusion of the honor code on syllabi	Low
Variation in faculty responses to honor code violations (includes no action)	Many
Faculty comfort level with their role in the honor code	Low
Faculty perception of student knowledge of the honor code	Low
Faculty perception of administration's support in enforcing the code	Medium
Faculty awareness of university procedures for honor code hearings	Medium
Faculty participation in honor code panels	Low

Faculty also indicated that students were not knowledgeable about the Honor Code, which prompted us to ask two follow-up research questions. 4. What are student perceptions of the Honor Code? and 5. What behaviors are students exhibiting relating to the Honor Code?

2005 Results from CAI Survey of Students and Faculty

We chose to use a nationally-used survey from the Center for Academic Integrity to gauge student and faculty perceptions of ethics on campus (www.academicintegrity.org). This allowed access to CAI's experts as well as comparative results from other universities for future benchmarking. This survey was comprehensive enough to include our research questions, which were posed to both faculty and students. McCabe administered the survey online. Participants were given the option to participate and their anonymity was assured. McCabe has used the survey for over a decade with 17,401 students and 3,752 faculty, and has gathered considerable validity and reliability evidence for its use in higher education (McCabe, Trevino and Butterfield, 2002). The first administration of the Center for Academic Integrity survey occurred in the Spring of 2005. Respondents were directed to a website over a one month period. The student online survey was made available to approximately 600 students in business classes. This included freshmen, sophomores, juniors and seniors in proportions representative of the school. The faculty survey was made available to all 200 faculty at the university. Three hundred thirty-eight students and 59 faculty completed the online survey. Some highlights of the 2005 data are shown in Table 2.

Table 2
Highlights – 2005 Students and Faculty Responses

Topic	%
Students informed about academic integrity policies on campus	98
Students reported learning the most about the honor code from:	
faculty	70
first year orientation	53
student handbooks	24
resident advisors	20
Had seen cheating on an exam never or once	
faculty	75
students	62
Students unlikely or very unlikely to report observed cheating	77
Students rating the honor code policies as high or very high	68
Faculty satisfied or very satisfied with the outcome of cheating cases they referred to judicial affairs	58
Faculty agree or strongly agree that assessments are effective in the evaluation of learning	89

From the findings, it was obvious that students and faculty were aware of the honor code, but the actual implementation of the honor code presented challenges. For example, 98% of students reported awareness of the academic integrity policies on campus yet more than three quarters of them indicated they would be unlikely or very unlikely to report observed cheating. Students also reported multiple

Honor Code education events during Freshman year but little thereafter. There was also wide disparity in the attention given to the honor code across different classes. Both faculty and students reported a lack of clear understanding of the judicial process and the ensuing consequences of Honor Code infractions. Due to these survey results, we researched best practices in both universities and businesses to develop and implement a set of interventions to promote understanding and improved adherence to the code.

First Set of Interventions

A first set of interventions took place in Fall 2005 and Spring 2006 and are listed below:

- Honor code information on all syllabi
- Student honor council established
- Honor code plaques in classrooms
- Honor code assignment in gateway and capstone courses
- Faculty presentation at university wide teaching and research conference
- Ethics Guest speakers on campus; event co-sponsored by judicial affairs
- Faculty awareness training
- Logo pencils handed out during exam week
- Student ethics research and presentations at university research conference
- Honor council website

Small focus groups were held with faculty and students to gauge the effectiveness of the actions. Suggestions from these meetings, such as making larger plaques that could be read from the back of the classroom, were utilized.

2007 Results from CAI Survey of Students and Faculty

In the Spring 2007, 326 students and 99 faculty participated in the second administration of the CAI survey instrument. This was again administered by McCabe of CAI using the same procedure as in Spring 2005. The findings from the 2007 survey are in Table 3.

Table 3
Highlights – 2007 Students and Faculty Survey Responses

Topic	%
Students said faculty discussed plagiarism often or very often	76
Students said faculty discussed source citing often or very often	78
Students said they never got help electronically on an exam	95
Faculty who had never ignored a case of cheating	62
Cheating punishments were more learning-centered in 2005:	
increase in faculty requiring a redo of the assignment	43
decrease in faculty giving an automatic F in the course	38
Faculty used safeguards in their courses to reduce cheating, such as changing exams, monitoring students during tests, and using the internet to confirm plagiarism	94
Faculty satisfied with the university process used to address violations of the honor code	72
Students agree or strongly agree that assessments helped them learn	73

Comparison of 2005 – 2007 Results

Following Rezaee, Elmore and Szendi (2001), t-test analyses were used to detect statistical differences among results in the 2005 and 2007 surveys. Many of the results from the 2007 survey were not statistically significantly different from those of 2005. However, some questions showed meaningful changes over the period. These results were likely due to the fact that many respondents were learning or relearning the honor code and the university processes related to a cademic integrity. Table 4 provides a sample of the student results. Questions 1, 2 and 5 show no statistical differences between the two surveys, while questions 3 and 4 do illustrate statistically significant differences in the findings. These changes can be attributed to campus-wide educational programming on academic integrity issues and increased exposure to the honor council during the study period.

Table 4
T-Test for Equality of Means: 2005 Versus 2007 Student Results

Topic	Year	*M*	*SD*	T-value	*df*	Sig. (2 tailed)
1. Students who state they have been informed about academic integrity policies on campus	2005	2.000	.398	.721	662	.760
	2007	2.000	.286			
2. Students' rating of the campus severity of penalties for cheating	2005	4.030	1.275	.843	662	.679
	2007	4.106	.998			
3. Students who learned from "other" sources about academic integrity policies on campus	2005	.448	.256	2.037	662	.008
	2007	1.276	.597			
4. Students who state frequency with which they worked with others when asked for individual work	2005	1.183	.776	2.674	662	.034
	2007	2.905	1.049			
5. Students report the likelihood that they would report an incident of cheating that they observed	2005	2.017	.987	1.370	662	.564
	2007	1.843	.855			

Question 1 Scale: 1=No; 2=Yes
Question 2 Scale: 1=Very Low; 2=Low; 3=Medium; 4=High; 5=Very High
Question 3 Scale: 1=Learned Little; 2=Learned Some; 3=Learned a Lot; 4=No Response
Question 4 Scale: 1=Never; 2=Once; 3=More than Once; 4=Not Applicable
Question 5 Scale: 1=Very Unlikely; 2=Unlikely; 3=Likely; 4=Very Likely

Table 5 provides a sample of the results from the faculty surveys. Questions 1, 4 and 5 show no statistical differences between the two surveys, while questions 2 and 3 do illustrate statistically significant differences in the findings. The changes can be explained in part by ethics-focused events on campus, faculty awareness training and other interventions as listed earlier.

Table 5
T-Test for Equality of Means: 2005 Versus 2007 Faculty Results

Topic	Year	*M*	*SD*	T-value	*df*	Sig. (2 tailed)
1. Faculty sentiment that cheating is a serious problem on campus	2005	3.16	1.287	1.047	156	.840
	2007	3.07	1.195			
2. Faculty who report being very satisfied with the outcome of referring a case of suspected cheating	2005	1.08	.683	2.845	156	.029
	2007	2.12	1.720			
3. Faculty who provide information on the syllabus about cheating or plagiarism	2005	1.06	.420	2.358	156	.031
	2007	1.53	.719			
4. Faculty who agree that assessments are effective in evaluating learning	2005	4.37	1.185	.993	156	.704
	2007	4.25	1.167			
5. Faculty who felt student understanding of academic integrity policy was high	2005	3.59	1.570	.650	156	.832
	2007	3.91	1.423			

Question 1 Scale: 1=Disagree Strongly; 2=Disagree; 3=Not Sure; 4=Agree; 5=Agree Strongly
Question 2 Scale: 1=Very Satisfied; 2=Satisfied; 3=Neutral; 4=Unsatisfactory; 5=Very Unsatisfactory
Question 3 Scale: 1=No; 2=Yes
Question 4 Scale: 1=Disagree Strongly; 2=Disagree; 3=Not Sure; 4=Agree; 5=Agree Strongly
Question 5 Scale: 1=Very Low; 2=Low; 3=Medium; 4=High; 5=Very High

Interestingly, in 2007 we did notice a more positive tone in the comments from both faculty and students. This led us to our second set of interventions. These actions were designed based on the results of the 2007 survey.

Second Set of Interventions

The second set of interventions took place in the 2007 - 2008 academic year. They included the following:

- In class teaching segments on plagiarism, ethics and integrity
- 2007 CAI International Conference Host
- Development of a one day bi-annual integrity conference for high school students and faculty advsiors
- T-shirts for Honor Council
- Award for student demonstrating high levels of integrity
- Student ethics research and presentations at international conference, with faculty guidance
- Student blog and interactive media activated on honor council website
- Fun events, such as pie a professor and ethics based movies
- Ethics based extra credit opportunities for students

These interventions were implemented with the assistance of the faculty and the student honor council. For example, students raised money for T-shirts, promoted the events and volunteered at the CAI conference. Faculty shared teaching strategies for the classroom.

Student Research and Results

Members of the student honor council were encouraged to create their own survey and administer it in their classes. Faculty oversaw the research project. The survey was pretested on students outside of the classes with good results. The student survey was given in Fall 2007 and Spring 2008 resulting in 275 usable responses. The findings are located in Table 6. Based on the data and student open comments on the surveys the student research team concluded that students may be the best resource to teach their peers about the honor code. This information will be used by the honor council to create new educational programs.

Table 6
Highlights from Student Research: 2007-2008

Topic	%
Students who believe the honor code is enforced fairly	48
Students who do not know the range of sanctions that can occur	42
Students who would report a fellow student for cheating	8
Students who say the honor system is discussed in class and on the syllabus	65
Students who have violated the honor code and not been caught	40
Students who believe that failure on the assignment was a reasonable sanction for a violation of the honor code	88

We will encourage the honor council to periodically assess the climate of ethical behavior in similar projects. The students presented their results at multiple conferences, including CAI, Virginia Assessment Group and their university student research conference.

Alumni Study and Results

The alumni survey took place in the Fall of 2008 via the use of the online survey system surveymonkey. The purpose of the survey was to gain input from alums regarding honor activities on campus for the goal of continuous improvement. A sixth research question was incorporated into this phase of the study: 6. What impact does the Honor Code have on alumni after graduation? Several ethics experts were consulted to assure content validity and multiple items were created for each construct to achieve acceptable levels of reliability. The survey was successfully pretested on a small sample of alums for content comprehension and ease of use. All business alums within the past 2 years with current contact information were sampled, as these students would be familiar with the interventions. Respondents were assured anonymity and they were contacted via email. Twenty-six alums participated for a 40% response rate. Some of the pertinent findings from the alumni study are in Table 7.

Table 7
Highlights – Alumni Survey 2008

Topic	%
No experience with the honor code as an undergraduate	73
Not members of the student honor council	77
Aware of the honor code as undergraduates	81
Said the presence of an honor code did not impact their decision to attend a given college	69
Said their degree was more valuable due to the honor code at the university	44
Said the honor code influenced their actions at the university	65
Believed the honor code and academic integrity at university prepared them mostly or a great deal for the challenges of the workplace	55

Results from In-House Alumni Survey

Alumni provided advice regarding honor and ethics activities for undergraduates as listed below. These data will be used to inform faculty regarding teaching ethics.

- Use company specific cases
- Don't harp on cheaters; Focus on the positive too
- Use real world examples such as computer security issues and difficult coworkers
- Stress ethics evenly over the college years
- Include ethics discussions in general education courses
- Discuss consequences for professors, not just students
- Practice decision making. Experience is the best kind of training.

Table 8
Student, Faculty, and Alumni Comments

Source	Comments
Students 2005	"I believe students understand the honor code and abide by it overall. I feel more secure in the fact that if I left my wallet or my door unlocked or something that it would stay there longer than at a public place or other school. Students understand the consequences through professors, and other students and stories of incidents."
Faculty 2005	"I am afraid cheating is more widespread than we know. I would also surmise that most of the cheating is unintentional and is not malicious. Students are trying to get the work done, and haven't had the training or instruction in what is considered okay and what is wrong in the university/academic setting."
Students 2007	"Taking this survey really opened my eyes to many different things – including some of the things I do actually are more serious than I think of them at the time I do them. It was a major reality check and it made me want to stop everything and change what I was doing the few times I did it. I really did learn something from answering the questions about myself and then their severity. Thank you."
Faculty 2007	"Foster a community of honor by trusting students' word. Do not ask for doctor's notes, etc. Publicize to students the codes of academic integrity that we follow in our disciplines so that students know we all are part of a community of honor. Students are bound by the [university] honor code, faculty uphold the [university] honor code and are bound by their own codes within their disciplines."
Alum 2008	"The more involved you can get with the research, presentation, and teaching of ethics, the more you will learn – both on your own and from others. The CAI conference is a great way to learn a lot about ethics from a variety of students and scholars!" "Keep doing what you do."

Discussion

Student, Faculty and Alumni Comments

In order to gain additional insights into opinions about academic integrity, open ended questions were included in the CAI and alumni surveys. Table 8 contains some of the comments that reinforce the need for honor code programming at the undergraduate level. We believe that anonymous and open-ended assessment techniques led to honest responses that would not have surfaced in non-anonymous, selected-response-type questionnaires.

Plans for Additional Assessments and Limitations

After the second set of interventions is completed, another CAI survey will be given to faculty and students in 2010 to assess progress. Also a separate survey using the online instrument surveymonkey will be created for the administration and staff. These groups will be included for a more complete picture of the culture of academic integrity on the campus. Other assessments planned include focus groups, online chats and freshmen entry and senior exit discussions. These will provide more opportunity for in depth answers.

One limitation of the study is that it focused exclusively on the school of business. Our plan is to extend this research to other areas of the university in the future. Another limitation is the use of self-report measures. This might influence some students and faculty against fully participating because they could fear that negative comments about the honor code would prompt retaliation from administrators, supervisors or faculty. We try to overcome this by offering multiple mechanisms for participation, including over the internet, on paper and via suggestion box.

Conclusion

We believe that a multipronged approach to the development and assessment of the culture of academic integrity is the best way to gauge the pulse of campus ethics. By partnering with the Center for Academic Integrity, we were able to use a survey instrument with considerable reliability and validity evidence for students and faculty, and also compare our results to national averages. By creating our own survey instrument for alumni, we were able to ask questions about specific campus interventions. Holding focus groups and entry and exit interviews in the future, will hopefully allow us to gain richer information than can be expressed via surveys. Furthermore, by polling the administration and staff in the future, we hope to learn the issues and concerns that they feel are important with regards to academic integrity. The use of grouped assessments to create a holistic system of assessment is illustrated by Flateby (2009). An advantage of this technique is the ability to change the methods of assessment as the project progresses.

Another area of assessment related to the current study is the high school academic integrity conference that we have sponsored and will continue to sponsor in the future. We believe it is important to gain data from the high school population and learn their knowledge and commitment to ethics in the classroom. We also intend to keep in touch with the business community to continue to bring in ethics speakers to illustrate the realities of unethical behavior in the workplace.

Through continual assessment we intend to gauge the movement in attitudes and behaviors among the university community with regard to integrity. Multiple mechanism of data collection, will allow us to best educate students and faculty about the Honor Code and ethical decision making. We believe this program of assessment will also be useful to other institutions in promoting a culture of academic integrity. Accordingly, a long term goal of the study is to create a guidebook based on this research and make it available to other universities. As suggested by McCabe and Pavela (2005), encouraging student involvement in the development of community standards of conduct is the best way to develop buy in and acceptance of honor codes and ethical behavior on college campuses. Including alumni, administrators and faculty in the process can result in honor codes that create a culture of integrity that works for all.

References

Cheating Varies Across Campus (Fall, 2004). *Captain's Log.*

Center for Academic Integrity. (n.d)Retrieved June 20, 2009 from www.academicintegrity.org

Damast, A. (2007, April 30). Duke MBAs fail ethics test. *B-School News.* Retrieved May 15, 2009 from www.businessweek.com

Flateby, T. (2009). Developments and changes resulting from writing and thinking assessment. *Research & Practice in Assessment*, *3*(1).

Gerold, R. & Steinberg, D.J. (2009, May 31). Honesty, such a lonely word. Cheating is rampant in schools from the lowest levels on up. *Newsday*, Long Island, N.Y. p.A40. Retrieved on August 4, 2009 from Proquest, document id 1735985951.

Gilbert, M. (2008, October 23). Ohio justice presides as Ohio Northern students vow to uphold business creed. *McClatchy-Tribune Business News.*

Graves, L. (2008, October 3). Which types of students cheat most? Retrieved July 1, 2009 from www.usnews.com

Kisamore, J., Stone, T., & Jawahar, I. (2007). Academic integrity: The relationship between individual and situational factors on misconduct contemplations. *Journal of Business Ethics, 75*, 381-394.

McCabe, D., & Katz, D. (n.d) Curbing cheating. *New Jersey Education Association*. Retrieved August 4, 2009 from www.njea.org/page.aspx?a=3955

McCabe. D., & Pavela, G. (2005). New honor codes for a new generation. Retrieved July 27, 2009 from www.insidehighered.com

McCabe, D., Trevino, L. & Butterfield, K. (2002). Honor codes and other contextual influences on academic integrity: A replication and extension of modified honor code settings. *Research in Higher Education*, *43*(3), 357-378.

Oleck, J. (2008, March 10). Most high school students admit to cheating. *School Library Journal.*

Rezaee, Z., Elmore, R., & Szendi, J. (2001). Ethical behavior in higher educational institutions: The role of the code of conduct. *Journal of Business Ethics, 30*, 171-183.

Evolving from Quantity to Quality:
A New Yardstick for Assessment

Keston H. Fulcher and Chris D. Orem
James Madison University

Author Note

Keston H. Fulcher, Center for Assessment and Research Studies, James Madison University; Chris D. Orem, Center for Assessment and Research Studies, James Madison University Correspondence concerning this article should be addressed to Keston H. Fulcher, Center for Assessment and Research Studies, James Madison University, Harrisonburg, VA 22807.
E-mail: fulchekh@jmu.edu

Abstract

Higher education experts tout learning outcomes assessment as a vehicle for program improvement. To this end the authors share a rubric designed explicitly to evaluate the quality of assessment and how it leads to program improvement. The rubric contains six general assessment areas, which are further broken down into 14 elements. Embedded within the article are links to the full rubric, an example of an exemplary assessment report, and a how-to guide for conducting and reporting quality assessment.

Introduction

As assessment practice in higher education evolves so too do the questions institutions and accreditors pose about assessment. Until recently the questions focused on participation and could be answered with statements like, "Ninety-seven percent of our academic degree programs submitted assessment reports in the current academic year." Although certainly important and an indicator of compliance, this information reveals little regarding the *quality* of assessment. If, as we believe, assessment's primary purpose is to guide programs toward improvement, then quality must be considered. Examples of legitimate questions include: Are objectives stated appropriately? Is there a clear link between the objectives and the methodology? Is the methodology sound? Is the interpretation of the program's strengths and weaknesses justified by the results? Do the program's plans for improvement logically fit with the results and interpretation? However, conveying information about quality is more challenging than conveying information about quantity.

Nonetheless, like Suskie (2009), we believe evaluating the integrity of assessment is a worthwhile endeavor. To this end James Madison University has developed a rubric to provide constructive feedback on the quality of assessment that can be used diagnostically at the academic program level and higher organizational levels. In this article we highlight the (a) focus of this rubric, (b) the assessment elements that are evaluated, (c) possible uses of resulting information, and (d) further considerations.

Focus of Rubric

To clarify our conceptual position, consider a scenario where a provost is reading two year-end assessment reports. Reviewing these documents, she discovers that the first program's report includes exceptionally positive results. On closer inspection, however, the results are based exclusively on indirect measures, course experiences are not mapped to learning outcomes, and information regarding the veracity of the assessment instruments or data collection design is absent. Further, the program provides no record of using results for improvement.

The second program's assessment report differs drastically from the first. It does not boast the same glowing results, but it clearly walks the reader through its assessment process. Specifically, the second program provides a convincing argument that the results are trustworthy and directly answer questions related to its objectives. Furthermore, the report clearly outlines how these results will be used to make improvements to both the program and the assessment process. If you were the provost, with which program would you be most satisfied?

This hypothetical scenario illustrates two contrasting perspectives when evaluating assessment reports. One approach concentrates primarily on the results; the other focuses on the trustworthiness of the results and how a program responds to its findings. From our perspective, we hope that administrators and faculty embrace the second. If assessment's primary role is for program improvement then assess-

ment should be evaluated on the quality of information it provides and the logic of the decisions that are derived from it.

Elements of the Rubric

From this perspective, James Madison University created a rubric that guides evaluative feedback on assessment. It is most directly applicable for academic degree programs. You can examine this rubric by going here: http://www.jmu.edu/assessment/JMUAssess/APT_Help_Package_4_15_2010.pdf The link also leads to several other related documents including a hypothetical exemplary report and a how-to-guide for conducting assessment. The interested reader will find that these documents provide much more detail than this article.

The rubric consists of six general areas that are further broken down into 14 elements (see Figure 1). The selection of elements was based upon several common models of assessment including Erwin's (1991) and Suskie's (2009). Although other rubrics have been developed for this purpose (e.g., Christopher Newport University: http://assessment.cnu.edu/docs/uaec_review_form.pdf), this rubric most clearly articulates the expectations for sound methodology, the area where many assessments break down.

Objectives	Learning Experiences	Methodology	Results	Results Shared	Program Improvement
Clarity & Specificity	Learning Experiences	Relationship Objs & Measures	Presentation	Results Shared	Program Improvement
Orientation		Types of Measures	History		Assessment Improvement
		Specification	Interpretation		
		Data Collection			
		Additional Validity Evidencer			

Figure 1. Organization of Rubric

Each of these elements is evaluated on a four point scale where 1 = Absent; 2 = Needs Improvement; 3 = Meets Expectations; and 4 = Exemplar. For each element the rubric provides a behavioral description associated with each level of performance. See Figure 2 for examples the verbs describing the desired actions of the students, and the content and skills to be exhibited – leads to the highest scores.

Learning Experiences

The rubric's second area targets the degree to which a program's courses/learning experiences are mapped to its objectives. Exemplary scores represent programs that have matched all of their objectives to curricular and sometimes co-curricular learning experiences. Note that a good curriculum map itself is not evidence of student learning. Rather, it represents where students should theoretically gain knowledge and skills.

Methodology

The rubric's third area covers methodology, the critical process that occurs between objectives and results. We find this is the area where faculty feel least comfortable and need the most feedback. Therefore, this section is divided granularly into five elements. The first element gauges the relationship between the measures (such as tests, essays, portfolios) used by a program and its objectives. Programs that score well not only provide a list of their measures, but they describe in detail *why* the measure is a good fit for assessing one or more objectives. To this end, faculty subject experts can specify exactly what component of a test corresponds to the objective. For example, a biology program could indicate that an entire rubric on oral communication corresponds to how it specified its objective on oral communication, which included eye contact, a good hook, clear organization, etc. Similarly, for a multiple choice test, the faculty would need to specify which items correspond to which objective(s). The main idea here is that faculty should choose a test or rubric that represents the skills and content outlined by one or more objectives.

3A. Data collection & research design integrity			
Absent	Needs Improvement	Meets Expectations	Exemplary
No information is provided about data collection process or data not collected.	Limited information is provided about data collection such as who and how many took the assessment, but not enough to judge the veracity of the process (e.g., thirty-five seniors took the test).	Enough information is provided to understand the data collection process, such as a description of the sample, testing protocol, testing conditions, and student motivation. Nevertheless, several methodological flaws are evident such as unrepresentative sampling, inappropriate testing conditions, one rater for ratings, or mismatch with specification of desired results.	The data collection process is clearly explained and is appropriate to the specification of desired results (e.g., representative sampling, adequate motivation, two or more trained raters for performance assessment, pre-post design to measure gain, cutoff defended for performance vs. a criterion)
6A. Improvement of programs regarding student learning and development			
No mention of any improvements.	Examples of improvements documented but the link between them and the assessment findings is not clear.	Examples of improvements (or plans to improve) documented and directly related to findings of assessment. However, the improvements lack specificity.	Examples of improvements (or plans to improve) documented and directly related to findings of assessment. These improvements are very specific (e.g., approximate dates of implementation and where in curriculum they will occur).

Figure 2. Examples of Behavioral Anchors Associated with Two Elements of the Rubric

The type of measure being used is also reviewed. Compared to essays, portfolios, or multiple choice tests, surveys are considered indirect and less objective. Correspondingly, the rubric rewards programs for using direct measures associated with each of its objectives. Note that it is good practice to include indirect measures but only as supplements to the direct measures.

Programs are also evaluated on whether they specify desired results for their objectives. The purpose of this element is to provide context for assessment results. Too often faculty will look at their assessment results and have little context for interpretation. If, at the outset (i.e., a priori), they indicate what results would indicate success, then the findings become more interpretable. Exactly what these results should look like depends on the type(s) of questions asked. What percentage of students meets a standard? How do students compare to similar programs across the country? To what degree did students change regarding their skills and knowledge? How does this cohort compare to the previous cohort? The rubric rewards specificity and rationale. As opposed to - "We intend for this cohort to perform better than last year's students." - a statement like this is much more powerful: "For the current cohort, our desired result is an average score of 83 on the exit exam. This score would connote a ½ standard deviation improvement from the previous year. We chose this moderate level of improvement because the current cohort is the first to undergo a modified curriculum where core content was emphasized more heavily." Articulating the desired results in such a fashion not only makes the results more interpretable but will likely entice faculty to engage with the findings; results are always more interesting when they address a question.

The next element under methodology is data collection. The most common problem we see in this area is insufficient information. At a minimum, an evaluator would need to know which students are targeted (i.e., population of interest, which should be specified in the objective), how students were sampled, the conditions under which students took the assessment, and their effort level. In addition, for a performance assessment, one would need to know about the raters and how they were trained. As an example, a program may report that "40 out of 41 seniors took the assessment during a set day in their senior seminar class in the spring semester; the test was proctored by faculty members, and was a graduation requirement. Consequently, proctors observed that students gave a good effort."

The fifth and final element under methodology refers to additional validity evidence. One may note that all six of the rubric's areas relate to validation of results and interpretations, or as Benson (1998) puts it, "…the process by which scores take on meaning" (p. 10). This element focuses on a particular part of validation: the psychometric properties of data like reliability. Note, we realize that some practitioners may be unfamiliar with these concepts. Nevertheless, they are necessary conditions of trustworthy results. We therefore strongly encourage faculty to consult with their institution's assessment consultants. Reliability estimates like coefficient alpha, inter-rater reliability, and other measures of consistency are all appropriate to report. The highest ratings are awarded to those programs whose assessment data have decent reliability and additional validity evidence. For example, if students who take more general education courses in mathematics score higher on a quantitative reasoning test, then such a result lends validity evidence to the test scores . Of all 14 elements on the rubric, this is likely the most difficult. Onlythe most mature programs who have worked with assessment experts (internal or external to the program) will receive exemplar marks.

Results

The fourth area of the rubric corresponds to assessment results, which is broken down into three elements: (a) presence of results—to what extent do they correspond to objectives? (b) history of results—in order to demonstrate trends, do programs report more than one year of data for some or all of their objectives? (c) interpretation of the results—does a program make reasonable inferences about the scores based on the methodology used? It is important to reiterate that the rubric does not directly evaluate whether or not desired results are achieved, but instead evaluates whether programs address the veracity of the results and how the program interprets and responds to them. In other words, a program can fall short of reaching their desired results, but still receive a high score. They can do so by providing a logical interpretation of the findings and reasons it believes the results fell short of expectations.

Sharing Results

Area five covers the ways in which a program disseminates its results to various stakeholders. Programs that do not share their results, or only provide data to a limited number of faculty members will score lower than ones that make their scores widely available to both internal and external audiences. The idea here is that assessment should be a collaborative enterprise among all faculty within a program and, ideally, external stakeholders such as an advisory board. Conversely, an assessment report viewed only by the eyes of the author rarely has bearing on a program.

Using Results

Making thoughtful programmatic changes to improve student learning is the very impetus of assessment, and it is the focus of the rubric's sixth area. The best assessments guide stakeholders in decision making, whether it be curricular, co-curricular, pedagogical, budgetary, etc. One may note that to make sound data-driven decisions, one needs to trust the assessment results first. Thus the emphasis on good objectives, methodology and the reporting of results noted in previous areas of the rubric. Exemplary assessment reports follow a clear logic from the assessment results to improvements that have been (or will be) implemented; as always, the more detailed the better.

In addition to evaluating the presence of results-driven improvements, the rubric also reviews whether programs address shortcomings to the assessment process itself. This element emphasizes that assessment is an ongoing process. As already stated, trustworthy results are a pre-requisite to using results for improvement. Therefore, by improving one's assessment, the likelihood that good decisions will be made about the program also increases. Recognizing that programs with strong assessment practices may not need to make drastic improvements to their assessment process, those who receive exemplary marks on the majority of the first five areas automatically receive a high score on this final element.

Using Information Obtained from Rubric

This article describes a process for evaluating assessment of academic programs via a rubric with six areas and fourteen elements. As with all assessments, it is essential to consider how the results will be used. We recommend two uses: (a) as a vehicle that provides diagnostic feedback about individual program's assessment and (b) as a mechanism to convey the quality of assessment across programs (i.e., at the department, college, and university levels). Regarding individual feedback, it informs faculty within programs about the strengths and weaknesses of their assessment. For example, perhaps a program's objectives are well articulated but concerns about methodology (e.g., absence of data collection procedures) cast doubt about the meaningfulness of the results. Consequently, in the next year the program can focus its efforts on improving the data collection process.

Additionally, feedback from the rubric can be diagnostic at the larger university level. The scores can be aggregated to identify strengths and weaknesses in the assessment process across programs, departments, and colleges. This information provides a university insight into how it can most efficiently support programs by creating or adapting services to bolster common needs. For example, the Office of Assessment could host a workshop on articulating desired results. Additionally, aggregated scores from across the university provide a gauge of where an institution stands regarding overall quality of academic program assessment. This information is easily interpreted by stakeholders and accrediting bodies. In essence, this aggregated data could be used to answer the quality-of-assessment questions at the macro level posed at the beginning of this article.

Further Considerations and Conclusions

While the primary focus of this article is on the rubric itself, there are several other important questions to consider when instituting an evaluation system of assessment reports. Will the assessment reports be collected electronically, or will they be turned in via hardcopy to a central location? Is there a common format required to make the reports easier to read, or are programs granted "creative discretion?" Who will rate the reports: faculty, students, or professional staff? How will raters be recruited and trained? Will feedback be provided for every section of the rubric or will general suggestions be made? What resources are available to programs that do not score well? Will the results of the rubric be used for high-stakes decision making, or simply for program improvement?

We acknowledge that assessment is a resource-intensive endeavor requiring money and, particularly, the time of faculty and staff. As such, this process needs to bear fruit in the form of enhanced student learning from improved degree programs. We hope that this rubric can be a resource toward that end. Regardless of whether this particular tool is appropriate for your institution, we recommend that every university incorporates some process of evaluating assessment. Too often this aspect of the assessment cycle is overlooked, an ironic fate for an endeavor rooted in reflection.

References

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice, 17*, 10-22.

Erwin, T. D. (1991). *Assessing student learning and development.* San Francisco, CA: Jossey-Bass.

Suskie, L. (2009). *Assessing student learning: A common sense guide* (2nd Edition). San Francisco, CA: Jossey-Bass.