Mar 11th, 9:00 AM - 9:45 AM

A Methodology for Developing and Validating an Assessment Rubric

Sandra Allen
Columbia College Chicago, sallen@colum.edu

John E. Knight
University of Tennessee at Martin, jknight@utm.edu

Follow this and additional works at: http://digitalcommons.georgiasouthern.edu/sotlcommons

Part of the Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, Educational Methods Commons, Higher Education Commons, and the Social and Philosophical Foundations of Education Commons

Recommended Citation

This presentation (open access) is brought to you for free and open access by the Programs and Conferences at Digital Commons@Georgia Southern. It has been accepted for inclusion in SoTL Commons Conference by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.
A Methodology for Developing and Validating an Assessment Rubric

Sandra Allen, M. A., M.B.A.  
Columbia College Chicago  
sallen@colum.edu

John E. Knight, Ph. D.  
Univ. of TN at Martin  
jknight@utm.edu

The SoTL Commons: A Conference for the Scholarship of Teaching & Learning  
Georgia Southern University  
Statesboro, GA

March 11, 2009

©Allen & Knight 2009
Background

Rubrics are:

- Modality for structuring learning outcomes with a formal assessment protocol (Stevens & Levi, 2005)
- More effective when (1) academic concepts and dimensions focus on student learning skills, and (2) dimensions and grading scales consider professional input
- Assessment tools for equitability regardless of instructor or section
Rubrics development:
• Process or series of steps
  • Example: Public relations news release rubric at Columbia College Chicago.
Literature Review

- Basic primer (Stevens & Levi, 2005)
- Advice on feedback (Rucker & Thomson, 2003)
- Expert informants (McNamara, 1995)
- Methodology (Adler & Ziglio, 1996)
- Validity, reliability and fairness (Andrade, 2005; Messick, 1989)
Literature Review (Cont’d)

- (Steel, 2008; McCleneghan, 2006; & College Board, 2004) on PR news release writing
- (Rios, 2008; Tucker, et al, 1997; and McMaster & Espin, 2007) on background information
- (Hardin & Pompper, 2004) lament voids in writing skills for new hire PR practitioners
The Collaborate Professional/Academic
Development and Validation of a Rubric

- Develop Learning Objectives (Course/Technique, etc.)
  - Identify a Representative Sample of Work that Demonstrates Learning Competency
    1. Review professional guidelines.
    2. Academic literature review
    3. Academic discussions
  - Develop an Evaluation Rubric for Uniformity of Teaching in All Sections
    1. Establish professional baseline
    2. Refine with academic input.
    3. Refine with Professional interviews
  - Longitudinally Test Student Learning Against Professional Assessments
- Identify Problems with Sub-Optimal Learning Performance.
  - Improve the Construct Validity of the Rubric
    - Determine the Ability of the Rubric to Differentiate Submissions to the Rubric
    - Improve the Consistency of Rubric Teaching and Student Learning with Data Analysis to Further Validate the Rubric.
- Test for Reliability of Multiple Graders to Develop Consistent Grades
Step 1 Develop Learning Objectives

Reflection stage

- Input sources
  - Faculty
  - Professional in the field
  - Certification and licensing guidelines of the profession
- Application to PR writing
Step 2  Identify a Representative Sample of Work

- Course objectives and related skills
- Select an assignment
- Review professional certification requirements
  - Example: PR news release
  - Why?
  - Example: Business field
Step 3  Develop an Evaluation Rubric for Uniformity of Teaching in All Sections

(Stevens & Levi, 2005) on skill and assignment objective

Then:

- Identify dimensions
- Develop grading scale
- Include collaborative input

Example:
- Fall 2005, PR writing rubric design
- Columbia College Chicago
Step 4  Longitudinally Test Student Learning Against Professional Assessments

- Establish a student learning baseline
- Retest student learning
- Measure areas of progress/regression.
  - Example: PR news release writing tests 2006 & 2007
Step 5  Identify Problems with Sub-Optimal Learning Performance

Identify:
- Areas of emphasis
- Student concerns
- Faculty issues
Step 5 (Cont’d) Application to PR Writing

Example

Faculty:
- Examined pedagogy
- Adopted consistency
- Agreed to use rubric as model
- Agreed to original content of rubric
- Recognized differences
- Recognized void in validation
Step 6  Improve Construct Validity of Rubric

Improved validity implies:

- Rubric dimensions apply equally in emphasis (weight) relative to the best judgments of academicians and professionals.
- Judgments should be examined and reconciled
Step 6 (Cont’d) Application of Delphi Method: Weights for Rubric

- Road to consensus: professional PR practitioners
Step 6 (Cont’d) Rubric Weights: PR Professionals and Faculty

<table>
<thead>
<tr>
<th>Rubric Category</th>
<th>Professional/Practitioner Rating</th>
<th>Faculty Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are the key items organized effectively?</td>
<td>35</td>
<td>15</td>
</tr>
<tr>
<td>2. In the headline/lead combination, is the message clear and compelling?</td>
<td>15</td>
<td>25</td>
</tr>
<tr>
<td>3. In the headline/lead combination, is the news angle compelling?</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>4. Does the news release use a convincing journalistic style?</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>5. Are the basics in place – grammar, sentence and paragraph mechanics?</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>
Step 7  Determine Rubric’s Ability to Differentiate Between Submissions

<table>
<thead>
<tr>
<th></th>
<th>PaperX</th>
<th>PaperY</th>
<th>Averages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professionals (n=5)</td>
<td>45.4</td>
<td>77.3</td>
<td>61.35</td>
</tr>
<tr>
<td>Faculty (n=6)</td>
<td>42.6</td>
<td>78.74</td>
<td>60.67</td>
</tr>
<tr>
<td>Averages</td>
<td>43.9</td>
<td>78.05</td>
<td>60.97</td>
</tr>
</tbody>
</table>

Two-way ANOVA: Grade versus Fac. vs. Prof., X or Y

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fac. vs. Prof.</td>
<td>1</td>
<td>2.8</td>
<td>2.80</td>
<td>0.01</td>
<td>0.910</td>
</tr>
<tr>
<td>X or Y</td>
<td>1</td>
<td>6915.6</td>
<td>6915.61</td>
<td>32.69</td>
<td>0.000</td>
</tr>
<tr>
<td>Interaction</td>
<td>1</td>
<td>25.2</td>
<td>25.22</td>
<td>0.12</td>
<td>0.734</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>4230.7</td>
<td>211.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>11174.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

S = 14.54  R-Sq = 62.14%  R-Sq(adj) = 56.46%
Step 7 (Cont’d) Numerical Analysis of Data

- Significant differences existed in papers.
- No significant differences in mean grades
- Conclusion:
  - Academic graders on the average provide a realistic grading perspective relative to what students need in workplace.
- The standard deviation of the model of 14.54 indicates a precision of grading using this model on any one paper would be too large.
Step 7 (Cont’d) Further Statistical Analysis

- The variance of grades administered by faculty was significantly lower than by professionals ($p = .01$)
  - Conclusion: The precision of assessment is better when only faculty grade papers.

- The variance of grades for better papers was significantly lower than poorer papers ($p = .04$)
  - Conclusion: Poor papers need proper assessment without discouraging student learning.
Step 7 (Cont’d) Further Statistical Analysis

- When good papers are graded, the best estimate of the precision of assessment was $\pm 5.2$ points with 68% confidence (within a letter grade in 2/3rds of the cases).
- When poor papers are graded, the best estimate of the precision of assessment was $\pm 14.36$ points with 68% confidence ... far from acceptable.
- Improved precision in assessment is indicated as necessary.
Step 8  Improve Consistency of Assessment

- Discussions of rubric terms and applications need further discussion and clarification across all instructors.
  - Common format of numerical feedback.
  - Common presentation for rubric in all sections
- Use Delphi Method for examining variability in assessment of a common paper.
Step 8 (Cont’d)  Improve Rubric Consistency and Precision

<table>
<thead>
<tr>
<th>Instr.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Ave.</th>
<th>S.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>79</td>
<td>100</td>
<td>79</td>
<td>85</td>
<td>88</td>
<td>84</td>
<td>82</td>
<td>85</td>
<td>91</td>
<td>85.89</td>
<td>6.57</td>
</tr>
<tr>
<td>2nd</td>
<td>83</td>
<td>94</td>
<td>84</td>
<td>85</td>
<td>88</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>90</td>
<td>86.89</td>
<td>3.41</td>
</tr>
<tr>
<td>Consensus</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88.00</td>
<td></td>
</tr>
</tbody>
</table>

- Results of assessment of a common paper by all 9 instructors along with variance reduction when the Delphi Method increases common assessment understanding.
Step 8 (Cont’d) Improve Rubric Consistency and Precision

- Precision of assessment increases from 1\textsuperscript{st} to 2\textsuperscript{nd} assessment ($p = .082$ for $n = 9$).
- Precision on 2\textsuperscript{nd} grading results in an assessment precision of $\pm 6.8$ with 95\% confidence.
Conclusions

- Processes to develop and improve rubric validation.
- Consider differences between practicing professionals and academicians
- Teaching focus:
  - Learning outcomes
  - Student & faculty confidence
  - Reliability & precision