Gathering, Analyzing, Interpreting, and Using Faculty Evaluation Data

November 16, 2007
Michael Theall, Ph.D.
Youngstown State University
330-941-1320 mtheall@ysu.edu

ACADEMIC IMPRESSIONS

BASIC ISSUES affecting data decisions

ALL EVALUATION and DEVELOPMENT ARE LOCAL!
Evaluation and development systems will not be complete until they are based on an understand of the work that faculty are expected to do, the skills that are required to do that work, and the criteria to be applied in evaluation!
BASIC ISSUES affecting data decisions

- Reliability
- Validity
- Generalizability
- Feasibility
- Skullduggery

DATA – DATA - DATA

- Purposes – data types
- Data uses
- Data sources
- Research --- [Ratings; Other data]
- Decision-making
- Instrumentation
- Analysis
- Interpretation
ACADEMIC IMPRESSIONS

What information do stakeholders need?

EVALUATION INFORMATION MATRIX

ACADEMIC IMPRESSIONS

Evaluation Purposes and Data

Purposes of Evaluation and Kinds of Data

Formative
(for information, revision, improvement)

Instrumental
(process and activities)

Consequential
(outcomes and effects)

Summative
(for decisions about merit or worth)
## Uses of Data

<table>
<thead>
<tr>
<th>Personnel Decisions</th>
<th>Teaching Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall performance</td>
<td>Assessable modifiable behaviors</td>
</tr>
<tr>
<td>Quantitative outcome oriented</td>
<td>Qualitative process oriented</td>
</tr>
<tr>
<td>Comparative database</td>
<td>Informative database</td>
</tr>
<tr>
<td>Empirical direct unambiguous</td>
<td>Comprehensive detailed suggestive</td>
</tr>
<tr>
<td>Global items</td>
<td>Specific “low inference” items</td>
</tr>
<tr>
<td>Public</td>
<td>Confidential</td>
</tr>
</tbody>
</table>

Supporting data

## Sources of data:

- Student Opinions & Student Learning
- Peer Evaluation (internal & external)
- Administrator and Self-Evaluation
- Advising; Recruiting; Admissions; Retention
- Media Documentation
- Alumni Ratings & Employer Opinions
- Scholarly Work (pubs; presents; citations; interviews; media; invited opinion; testimony; products, performances; exhibits)
ACADEMIC IMPRESSIONS
Sources of data:

- Awards; Honors; Grant$; Donation$
- Professional Activities (consulting & pro-bono)
- Service (department; college; university; national; professional, community)
- Administration (committees; grants; service or Gen Ed course coordination; GA/TA supervision; interim roles)

ACADEMIC IMPRESSIONS
Research Findings and Recommendations:

- Student ratings
- Other sources of evaluation data
Ratings are:

- Multidimensional
- Reliable and Stable
- Primarily a function of the instructor
- Relatively valid as evidence of effective teaching
- Relatively unaffected by a number of variables posed as biases
- Useful as teaching feedback

Marsh, 2007
Additional findings:

- Class size: slight negative (curvilinear)
- Prior interest in subject: positive
- Elective vs. required courses: more positive for electives
- Disciplinary area: consistent differences
- Work/difficulty: slight positive (curvilinear)
- Course level: slight positive for upper division & grad
- Anonymity: ratings more positive if violated

Additional findings:

- Purpose of eval: more positive if manipulated
- Instructor rank: none
- Teacher/student gender: none
- Teacher ethnicity/race: none
- Research productivity: none
- Student locus & performance attributions: none
- Student/teacher personality: none
<table>
<thead>
<tr>
<th>ITEMS</th>
<th>RANKED CORRELATION WITH ACHIEVEMENT</th>
<th>RANKED CORRELATION WITH EVALUATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. preparation and organization</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2. clarity and understandableness</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3. perceived outcome or impact</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4. stimulation of interest in content</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5. encouragement and openness</td>
<td>5-6</td>
<td>11</td>
</tr>
<tr>
<td>6. availability and helpfulness</td>
<td>5-6</td>
<td>16</td>
</tr>
<tr>
<td>7. presentation and speaking skills</td>
<td>7-8</td>
<td>10</td>
</tr>
<tr>
<td>8. clarity of objectives and requirements</td>
<td>7-8</td>
<td>7</td>
</tr>
<tr>
<td>9. subject knowledge</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

### ACADEMIC IMPRESSIONS

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>RANKED CORRELATION WITH ACHIEVEMENT</th>
<th>RANKED CORRELATION WITH EVALUATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. concern for student progress</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>11. teacher enthusiasm for subject</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>12. teacher's fairness</td>
<td>12</td>
<td>14-15</td>
</tr>
<tr>
<td>13. intellectual challenge</td>
<td>13</td>
<td>4 /</td>
</tr>
<tr>
<td>14. concern / respect for students</td>
<td>14-15</td>
<td>12</td>
</tr>
<tr>
<td>15. feedback quality &amp; frequency</td>
<td>14-15</td>
<td>17</td>
</tr>
<tr>
<td>16. nature / value of course material</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>17. nature / usefulness of supplements/aids</td>
<td>17</td>
<td>14-15</td>
</tr>
</tbody>
</table>


### Other sources of data:

- **Peer Evaluation**
  (usually department, but sometimes institutional colleagues)

  Best for teacher knowledge, certain course or curricular issues, assessment issues, currency /accuracy of content, (esp. when used along with student ratings). If on teaching, less reliable and higher on average than student ratings.
### Other sources of data:

- **Administrator Evaluation** *(department chair)*
  - Necessary as part of process, but same problems as peers on teaching (criteria, process, instruments, validation, etc.)

- **Self – Evaluation** *(e.g., in a portfolio)*
  - Provides the most complete picture of teacher thinking & instructional decisions/practices, but difficult to reliably interpret & use

### Other sources of data:

- **External Expert Evaluation** *(almost always by an expert in the same discipline; sometimes by an expert in teaching)*
  - Useful, but require process cautions and careful use/interpretation; having a purpose is important

- **Alumni Ratings** *(at various career stages & times)*
  - Can be useful but generally the same as student ratings given same instrument; can shed light on teaching in terms of content, process, or curricular issues for formative purposes.
Other sources of data:

- **Media Documentation** (usually video recording for teaching, but other forms for overall evaluation)
  
  Excellent for formative purposes; need guidelines for use by others beyond teacher; unambiguous if used carefully to assess low-inference behaviors; other media may vary in quality (e.g., Carl Sagan vs. local news)

- **Awards & Honors** (all types)
  
  Local awards lack standard criteria & decision processes; national awards more prestigious

Other sources of data:

- **Scholarship of Teaching & Learning**
  
  Valid and important **IF** recognized within the dept/college/univ.

- **Employer Opinions of Graduates**
  
  Limited use; better for program evaluation & curricular issues
Other sources of data:

• **Student Learning Outcomes**
  
  Useful for formative (individual) or program purposes (if aggregated for assessment); not recommended for summative decisions. Test scores are not as reliable as ratings from a validated instrument. Criteria vary considerably (e.g., What does “All her students got ‘A’s’” mean?)

---

**FOR FORMATIVE DECISIONS**

• **Student Ratings** (Teaching)
• **Peer-external ratings** (All topics)
• **Self-ratings** (All topics)
• **Videos** (Primarily teaching)
• **Student Interviews** (Primarily teaching)
• **Exit & Alum Ratings** (Primarily teaching)

Adapted from Berk, 2006, p. 45
Gathering, Analyzing, Interpreting, and Using Faculty Evaluation Data
Academic Impressions Web Conference

ACADEMIC IMPRESSIONS
FOR SUMMATIVE DECISIONS
(ANNUAL REVIEW)

• Student Ratings (Teaching)
• Self-peer Ratings (Various topics)
• Administrator Ratings (All issues)
• Scholarship (Disciplinary)
• Scholarship of Teaching
• Service (department/institution/professional)
• Administrative Support (Rating of work performed as administration or service)

Adapted from Berk, 2006, p. 45

ACADEMIC IMPRESSIONS
FOR SUMMATIVE DECISIONS
(promotion & tenure)

• Student Ratings (Teaching)
• Administrator Ratings (All topics)
• Teaching Portfolio
  (All topics with caution & clear criteria)
• Institutional Input (P & T Committees)
• External Input (Some or all topics)

Adapted from Berk, 2006, p. 45
FOR PROGRAM DECISIONS

- Student Ratings
- Assessment Data
- Exit and Alumni Ratings/Data
- Employer Ratings
- Institutional Data

Adapted from Berk, 2006, p. 45

Implementation issues:

- Instrumentation
  - Analysis
- Reporting
- Interpretation
Instrumentation

- Student Rating Instruments
- Peer/Administrator Protocols
- Peer/Administrator Instruments
- Media Documentation
- Checklists/Rubrics
Analysis possibilities for validation

- Item analysis
- Reliability coefficients
- Correlational analysis
- Factor analysis
- Regression analysis

on entire database and on subsets as soon as enough data is available

REPORTING DATA

- Statistics
- Formats
- Content
- Delivery (paper vs e-copy)
- Distribution
- Interpretation Training
ACADEMIC IMPRESSIONS

Analysis possibilities for reports of results

- Descriptive statistics
  (item distributions in # and %)
- Central tendency (mean, mode, median)

\[
\begin{array}{c}
1 & 2 & 3 & 3 & 4 & 5 \\
1 & 1 & 2 & 3 & 4 & 5 & 5 \\
1 & 2 & 3 & 4 & 5 & 5 \\
1 & 2 & 3 & 4 & 5 & 5 & 5 \\
\end{array}
\]

\[
(3, 3, 3) \\
(3, 1 \& 5, 3) \\
(3.33, 5, 3.5) \\
(3.57, 5, 4)
\]

- Standard deviations (sampling error)
- Enrolled / responded #s and ratio

ACADEMIC IMPRESSIONS

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Minimum acceptable response</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-20</td>
<td>at least 80%</td>
</tr>
<tr>
<td>20-30</td>
<td>at least 75%</td>
</tr>
<tr>
<td>30-50</td>
<td>at least 60%</td>
</tr>
<tr>
<td></td>
<td>75% or more recommended</td>
</tr>
<tr>
<td>50-100</td>
<td>at least 50%</td>
</tr>
<tr>
<td>&gt;100</td>
<td>more than 50%</td>
</tr>
</tbody>
</table>

*providing there is no systematic reason for absence or non-responding which might bias response.
Analysis possibilities for reports of results

- Standard scores
- Comparative data (norms, criterion references, self-ratings)
- Ranges (%ile rank, %ile group, confidence intervals for self and comparison groups)

Application to Decision Making

1. Sufficient TCE data to characterize average performance
2. Decision-makers use valid interpretations of TCE results
3. TCEs are one source among several (portfolio)
4. Criteria are standard, documented, and public.
Quality of Personnel Decision Depends On

- Fair personnel practices
- Interpretive skills of decision-makers
  - Knowledge of evaluation methods
  - Quantitative skills
  - Knowledge of post secondary teaching practice and technique
- The quality of the information decision-makers receive
  - Validity – measures relevant aspects of teaching skill or instructional quality
  - Reliability – precision
  - ‘Comprehensibility’ – message design and contents are appropriate for the skills of users, including any needed decision support

INSTRUCTIONAL REPORT of EDUCATIONAL SATISFACTION: I.R.E.S.)

Universitas pro Omnibus Discipuli et Facultitas in Excelsis

Instructor: U.N. Fortunate
Course #: HIS123
Course name: History of Everything
Term/year: Spring, 1994

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>amount learned</td>
<td>3</td>
<td>16</td>
<td>46</td>
<td>21</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>overall teacher</td>
<td>1</td>
<td>12</td>
<td>40</td>
<td>29</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>overall course</td>
<td>2</td>
<td>8</td>
<td>49</td>
<td>20</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: (A) =5= Best; (F)=6=Worst ... Enrolled: 120; Responded: 53
### INSTRUCTIONAL REPORT of EDUCATIONAL SATISFACTION: I.R.E.S.

Universitas pro Omnibus Discipuli et Facultitas in Excelcis  
Instructor: U.N. Fortunate  
Course #: HIS123  
Course name: History of Everything  
Term/year: Spring, 1994

<table>
<thead>
<tr>
<th>% / # responses</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>O</th>
<th>mean</th>
<th>s d</th>
<th>T</th>
<th>group</th>
</tr>
</thead>
<tbody>
<tr>
<td>amount learned</td>
<td>3/2</td>
<td>16/10</td>
<td>46/29</td>
<td>21/13</td>
<td>14/9</td>
<td>0/0</td>
<td>0/1</td>
<td>1/1</td>
<td>2.64</td>
<td>0.88</td>
<td>27 low</td>
</tr>
<tr>
<td>overall teacher</td>
<td>1/1</td>
<td>12/8</td>
<td>40/25</td>
<td>29/18</td>
<td>24/11</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>2.43</td>
<td>0.96</td>
<td>24 low</td>
</tr>
<tr>
<td>overall course</td>
<td>2/1</td>
<td>18/11</td>
<td>49/31</td>
<td>20/13</td>
<td>13/7</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
<td>2.81</td>
<td>0.93</td>
<td>33 low</td>
</tr>
</tbody>
</table>

Raw score: (A) =5= Best; (E) =1= Worst; F= Not applicable; O = Omitted;  
T-score: Standardized score where 40 – 60 = mean, and each 10 points in each direction is one standard deviation  
Group score: 0-10% = low; 10-30% = low middle; 30-70% = middle; 70-90% = high middle; 90-100% = high

---

### Two evaluations of HIS 345

<table>
<thead>
<tr>
<th>mean</th>
<th>s d</th>
<th>T</th>
<th>group</th>
<th>term/yr</th>
<th>instr</th>
<th>course</th>
<th>resp/enr</th>
</tr>
</thead>
<tbody>
<tr>
<td>amount learned</td>
<td>3.35</td>
<td>0.87</td>
<td><strong>45</strong></td>
<td>low-mid</td>
<td>term/yr = spring, 1995</td>
<td>instr = UNFortunate</td>
<td>course = his 345</td>
</tr>
<tr>
<td>overall teacher</td>
<td>2.76</td>
<td>0.76</td>
<td><strong>35</strong></td>
<td>low</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overall course</td>
<td>2.85</td>
<td>0.90</td>
<td><strong>37</strong></td>
<td>low</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>mean</th>
<th>s d</th>
<th>T</th>
<th>group</th>
<th>term/yr</th>
<th>instr</th>
<th>course</th>
<th>resp/enr</th>
</tr>
</thead>
<tbody>
<tr>
<td>amount learned</td>
<td>3.97</td>
<td>1.40</td>
<td><strong>56</strong></td>
<td>hi-mid</td>
<td>term/yr = fall, 1995</td>
<td>instr = UNFortunate</td>
<td>course = his 345</td>
</tr>
<tr>
<td>overall teacher</td>
<td>3.57</td>
<td>1.30</td>
<td><strong>47</strong></td>
<td>mid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overall course</td>
<td>3.63</td>
<td>1.24</td>
<td><strong>50</strong></td>
<td>mid</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ACADEMIC IMPRESSIONS

Enrollment profiles for HIS 345 in two semesters

<table>
<thead>
<tr>
<th>Fr</th>
<th>So</th>
<th>Jn</th>
<th>Sn</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>original enr.</td>
<td>6</td>
<td>17</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>term/yr = spring, 1995</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instr = UN Fortunate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>course = his 345</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>final enr.</td>
<td>5</td>
<td>14</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>resp/enr = 29/51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eval respondents</td>
<td>5</td>
<td>13</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>% resp=57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fr</th>
<th>So</th>
<th>Jn</th>
<th>Sn</th>
<th>Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>original enr.</td>
<td>3</td>
<td>11</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>term/yr = fall, 1995</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instr = UN Fortunate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>course = his 345</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>final enr.</td>
<td>2</td>
<td>7</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>resp/enr = 20/29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eval respondents</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>% resp=69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ACADEMIC IMPRESSIONS

Graphic display of 95% confidence intervals for individuals vs. comparison groups

1 2 3 4 5

- Personal range
- Department range
- Institutional range

Teacher A
Teacher B
ACADEMIC IMPRESSIONS

Guideline #1
(do your homework)

• Establish the purpose of the evaluation and the uses and users of ratings beforehand;
• Include all stakeholders in decisions about evaluation process and policy;
• Keep a balance between individual and institutional needs in mind;
• Build a real "system" for evaluation, not a haphazard and unsystematic process;

ACADEMIC IMPRESSIONS

Guideline #2
(establish protection for all)

• Publicly present clear information about the evaluation criteria, process, and procedures.
• Establish legally defensible process and a system for grievances;
• Establish clear lines of responsibility/reporting for those who administer the system;
• Produce reports that can be easily and accurately understood.
ACADEMIC IMPRESSIONS

Guideline # 3
(make it positive, not punitive)

- Absolutely include resources for improvement and support of teaching and teachers;
- Educate the users of ratings results to avoid misuse and misinterpretation;
- Keep formative evaluation confidential and separate from summative decision making;
- In summative decisions, compare teachers on the basis of data from similar situations;
- Consider the appropriate use of evaluation data for assessment and other purposes.

Guidelines # 4
(verify & maintain the system)

- Use, adapt, or develop instrumentation suited to institutional/individual needs;
- Use multiple sources of information from several situations;
- Keep ratings data and validate the instruments used;
- Invest in the evaluation system and evaluate it regularly;
- Seek expert, outside assistance when necessary or appropriate.
Questions?

Michael Theall, Ph.D.
Youngstown State University
330-941-1320
mtheall@ysu.edu

A PLACE TO BEGIN

EXAMINE THE PRIORITIES AND NEEDS OF THE STAKEHOLDERS IN AN OPEN AND PUBLIC PROCESS OF DIALOGUE AND CONSENSUS BUILDING

CONSIDER A WIDE VIEW OF THE ROLES AND RESPONSIBILITIES OF THE FACULTY AT YOUR INSTITUTION

IDENTIFY THE SKILLS FACULTY NEED AS THE BASIS FOR EVALUATION AND PROFESSIONAL ENRICHMENT
Gathering, Analyzing, Interpreting, and Using Faculty Evaluation Data
Academic Impressions Web Conference

Summary matrix http://www.cedanet.com/meta

<table>
<thead>
<tr>
<th>SKILL SETS</th>
<th>FACULTY ROLES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teaching</td>
</tr>
<tr>
<td>BASE PROFESSION</td>
<td>Content Expertise</td>
</tr>
<tr>
<td>Skill Sets</td>
<td>Practice/Clinical Skills</td>
</tr>
<tr>
<td></td>
<td>Research Techniques</td>
</tr>
<tr>
<td></td>
<td>Instructional Design</td>
</tr>
<tr>
<td></td>
<td>Instructional Delivery</td>
</tr>
<tr>
<td></td>
<td>Instructional Assessment</td>
</tr>
<tr>
<td></td>
<td>Course Management</td>
</tr>
<tr>
<td>Additional META-PROFESSION</td>
<td>Instructional Research Techniques</td>
</tr>
<tr>
<td>Skill Sets</td>
<td>Psychometrics/Statistics</td>
</tr>
<tr>
<td></td>
<td>Epistemology</td>
</tr>
<tr>
<td></td>
<td>Learning Theory</td>
</tr>
<tr>
<td></td>
<td>Human Development</td>
</tr>
<tr>
<td></td>
<td>Information Technology</td>
</tr>
<tr>
<td></td>
<td>Technical Writing</td>
</tr>
<tr>
<td></td>
<td>Graphic Design</td>
</tr>
<tr>
<td></td>
<td>Public Speaking</td>
</tr>
<tr>
<td></td>
<td>Almost Always</td>
</tr>
<tr>
<td></td>
<td>Occasionally</td>
</tr>
<tr>
<td></td>
<td>Almost Never</td>
</tr>
<tr>
<td>Almost Always</td>
<td>Communications Styles</td>
</tr>
<tr>
<td>Occasionally</td>
<td>Conflict Management</td>
</tr>
<tr>
<td>Occasionally</td>
<td>Group Process</td>
</tr>
<tr>
<td>Almost Never</td>
<td>Personnel Management</td>
</tr>
<tr>
<td></td>
<td>Resource Management</td>
</tr>
<tr>
<td></td>
<td>Supervision/Management</td>
</tr>
<tr>
<td></td>
<td>Financial/Budget Development</td>
</tr>
<tr>
<td></td>
<td>Policy Analysis &amp; Development</td>
</tr>
</tbody>
</table>

Basic References


ACADEMIC IMPRESSIONS

Basic References


53

ACADEMIC IMPRESSIONS

Basic References


Other New Directions volumes on evaluation/ratings:
N. D. Teaching & Learning #s: 83, 87, 96; N. D. Institutional Research #: 114

54
For more information, go to:

http://www.cedanet.com/meta
for documents and materials about the
‘meta-profession’ of the professoriate
http://ntlf.com/pod/index.html
for a review of the research and an
extended/annotated bibliography

---

For information about evaluation instruments:

Student Instructional Report (SIR / SIR II)
Educational Testing Service
609-921-9000
www.ets.org

Instructional Development and Educational
Assessment Survey (IDEA)
The IDEA Center
Kansas State University
800-255-2757; idea@ksu.edu
http://www.idea.ksu.edu/
For information about evaluation instruments:

Course-Instructor Evaluation Questionnaire (CIEQ)
Lawrence Aleamoni
University of Arizona
www.cieq.com

Student Evaluation of Educational Quality (SEEQ)
Herbert W. Marsh
Oxford University
herb.marsh@edstud.ox.ac.uk
Evaluation Information Matrix

Developing a Synergy for Improved Practice
Student Ratings of Instruction

provide valid, reliable, and useful information about teaching performance when CORRECTLY

**Constructed...**
- Well-written items
- Appropriate number of items
- Items valid for intended use
  - Administrative decision-making (merit, P&T, assignment, etc.)
  - Teaching improvement
  - Student-consumer feedback
  - Program and staff development
  - Institutional research

**Administered...**
- Accurate course information
  - Teacher identity
  - Course format & enrollment
- Standard procedure for
  - Who collects data
  - When
  - How
- Effective data management
- Verification and follow-up

**Analyzed & Reported...**
- Report format is valid for intended use; and provides estimates of, or controls for:
  - Sample adequacy
  - Sample quality
  - Error (precision)
  - Systematic variation
- Report includes meaningful, valid comparisons where needed for administrative use.

**Interpreted & Applied...**
- Interpretations grounded in explicit evaluation processes that account for:
  - What is "valued" and/or preferred outcomes. Explicit definition of teaching (e.g., content vs. instructional methods)
  - Standard method used for interpretation
  - Other valid data sources
  - "Valid & reliable" users

---

**Questionnaires:**
- Global or "overall" items or empirically valid factor scores for administrative use
- "Low inference", behaviorally descriptive items for diagnostic feedback
- Special purpose forms for lab, studio, discussion formats
- Intensive, mid-semester form available with consultation
- Optional instructor-made or selected items

**Operations:**
- Instructor and course identification matches college or department course files
- Special attention to public relations and communications with clients
- Secure packaging materials with emphasis on clear instructions
- Documentation of student monitor identity
- Careful date-stamping
- No data released for use without confirmation of proper administration

**Reports of Results:**
- Multiple report formats each designed to support specific decision-making purpose of user
  - Descriptive for teaching improvement
  - Comparative for admin.
  - Ongoing evaluation of the "usability" of report design and use documentation
  - Ongoing analysis of evaluation data to detect sources of systematic variation or bias
  - Timely reporting

**Services:**
- Training provided to department heads and committees on methods for assessing teaching performance
- Supporting services for evaluating instructional quality (e.g., video, peer, portfolio, etc.)
- Workshops and training activities for specific needs of ratings users
- Individual consultation on interpretation and use of ratings provided to faculty and administrators
<table>
<thead>
<tr>
<th>ITEMS</th>
<th>RANKED CORRELATION WITH ACHIEVEMENT</th>
<th>RANKED CORRELATION WITH EVALUATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. preparation and organization</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2. clarity and understandableness</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3. perceived outcome or impact</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4. stimulation of interest in content</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5. encouragement and openness</td>
<td>5-6</td>
<td>11</td>
</tr>
<tr>
<td>6. availability and helpfulness</td>
<td>5-6</td>
<td>16</td>
</tr>
<tr>
<td>7. presentation and speaking skills</td>
<td>7-8</td>
<td>10</td>
</tr>
<tr>
<td>8. clarity of objectives and requirements</td>
<td>7-8</td>
<td>7</td>
</tr>
<tr>
<td>9. subject knowledge</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>10. concern for student progress</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>11. teacher enthusiasm for subject</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>12. teacher's fairness</td>
<td>12</td>
<td>14-15</td>
</tr>
<tr>
<td>13. intellectual challenge</td>
<td>13</td>
<td>4 /</td>
</tr>
<tr>
<td>14. concern / respect for students</td>
<td>14-15</td>
<td>12</td>
</tr>
<tr>
<td>15. feedback quality &amp; frequency</td>
<td>14-15</td>
<td>17</td>
</tr>
<tr>
<td>16. nature / value of course material</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>17. nature / usefulness of supplements/ aids</td>
<td>17</td>
<td>14-15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SKILL SETS</th>
<th>Teaching</th>
<th>Scholarly &amp; Creative Activities</th>
<th>Service</th>
<th>Administration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BASE PROFESSION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skill Sets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Content Expertise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Practice/Clinical Skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Research Techniques</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Additional</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>META-PROFESSION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skill Sets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional Delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructional Research</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychometrics/Statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epistemology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning Theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information Technology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical Writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphic Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public Speaking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communications Styles</td>
<td>Almost Always</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflict Management</td>
<td>Frequently</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Process</td>
<td>Occasionally</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource Management</td>
<td>Almost Never</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personnel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervision/Management</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial/Budget Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy Analysis &amp; Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary matrix [http://www.cedanet.com/meta](http://www.cedanet.com/meta)