

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Collecting, Analyzing, Interpreting, and Using Faculty Evaluation Data

April 19, 2011

Michael Theall
Youngstown State University
mtheall@ysu.edu

ACADEMIC • IMPRESSIONS 1

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

LEARNING OBJECTIVES

You will be able to identify the most productive ways
to use faculty evaluation data:

- * formatively for improvement and
- * summatively for decision-making

ACADEMIC • IMPRESSIONS 2

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Agenda

- Evaluation purposes and data types
- Collecting data from multiple sources
- Steps to analyzing data
- Using evaluation data
- Guidelines to good practice

ACADEMIC • IMPRESSIONS ³

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

1. Evaluation Purposes and Data Types

ACADEMIC • IMPRESSIONS ⁴

All Evaluation and Development are Local!

- Evaluation without Development is
Punitive
- Development without Evaluation is
Guesswork!

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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!

| SKILL SETS | | FACULTY ROLES | | | |
|---------------------------------------|-----------------------------------|-----------------------|---------------------------------|---------|----------------|
| | | Teaching | Scholarly & Creative Activities | Service | Administration |
| BASE PROFESSION Skill Sets | Content Expertise | | | | |
| | Practice/Clinical Skills | | | | |
| | Research Techniques | | | | |
| Additional IETA-PROFESSION Skill Sets | Instructional Design | | | | |
| | Instructional Delivery | | | | |
| | Instructional Assessment | | | | |
| | Course Management | | | | |
| | Instructional Research Techniques | | | | |
| | Psychometrics/Statistics | | | | |
| | Epistemology | | | | |
| | Learning Theory | | | | |
| | Human Development | | | | |
| | Information Technology | | | | |
| | Technical Writing | | | | |
| | Graphic Design | | | | |
| | Public Speaking | | | | |
| | Almost Always | Communications Styles | | | |
| | Frequently | Conflict Management | | | |
| | Occasionally | Group Process | | | |
| | Almost Never | Resource Management | | | |
| | Personnel Supervision/Management | | | | |
| | Financial/Budget Development | | | | |
| | Policy Analysis & Development | | | | |

Summary matrix <http://www.cednet.com/meta>

Data

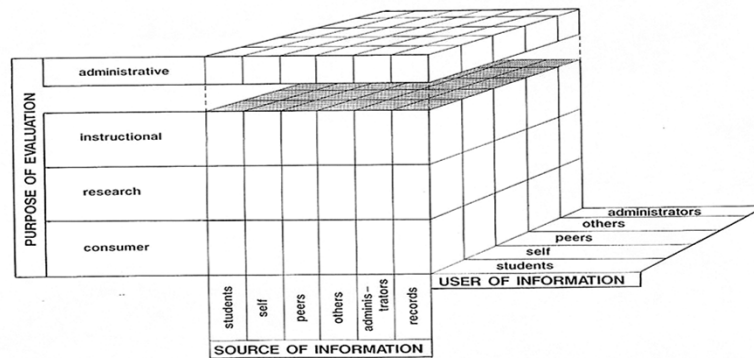
- Purposes – data types
- Data uses
- Research [Ratings; other data]
- Decision-making
- Instrumentation
- Analysis
- Interpretation

Basic Issues affecting data decisions

- Validity
- Reliability
- Generalizability
- Feasibility
- Skulduggery

What information do stakeholders need?

Evaluation Information Matrix: Developing a Synergy for Improved Practice



COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Evaluation Purposes and Data Types

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)

ACADEMIC • IMPRESSIONS 13

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Uses of Data

| Personnel Decisions | Teaching Improvement |
|--|--|
| Overall performance Quantitative outcome oriented Comparative database Empirical direct unambiguous Global items Public | Assessable modifiable behaviors Qualitative process oriented Informative database Comprehensive detailed suggestive Specific “low inference” items Confidential |
| Supporting data | |

ACADEMIC • IMPRESSIONS 14

2. Collecting Data From Multiple Sources

Sources of data:

- Student Opinions & Student Learning
- Peer Evaluation (internal & external)
- Administrator and Self-Evaluation
- Service (department; college; university; national; professional, community)
- Scholarly Work (pubs; presents; citations; interviews; media; invited opinion; testimony; products, performances; exhibits)

Sources of data:

- Awards; Honors; Grant\$; Donation\$
- Professional Activities (consulting & pro-bono)
- Administration (committees; grants; service or Gen Ed course coordination; GA/TA supervision; interim roles)
- Advising; Recruiting; Admissions; Retention
- Alumni Ratings & Employer Opinions
- Media Documentation

17

Research Findings and Recommendations:

- **Student ratings**
- **Other sources of evaluation data**

18

What are ratings?

- 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

19

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

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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*

21

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

| ITEMS | RANKED CORRELATION WITH ACHIEVEMENT | RANKED CORRELATION WITH EVALUATIONS |
|---|-------------------------------------|-------------------------------------|
| 1. preparation and organization | 1 | 6 |
| 2. clarity and understandableness | 2 | 2 |
| 3. perceived outcome or impact | 3 | 3 |
| 4. stimulation of interest in content | 4 | 1 |
| 5. encouragement and openness | 5-6 | 11 |
| 6. availability and helpfulness | 5-6 | 16 |
| 7. presentation and speaking skills | 7-8 | 10 |
| 8. clarity of objectives and requirements | 7-8 | 7 |
| 9. subject knowledge | 9 | 9 |

Dimensions of teaching table adapted by Michael Theall with permission from Kenneth Feldman. Source: Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisession validity studies. *Research in Higher Education*, 30, 583-645.

22

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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

Dimensions of teaching table adapted by Michael Theall with permission from Kenneth Feldman. Source: Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30, 583-645.

23

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

FACTORIAL INVARIANCE

learning / value

workload / difficulty

enthusiasm

breadth of coverage

organization

individual rapport

assignments

group interaction

exams / grading

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:

- **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?)

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

- **Awards & Honors** (all types)

Local awards lack standard criteria & decision processes; national awards more prestigious but consider the source

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

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Questions?



ACADEMIC • IMPRESSIONS 31

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

3. Uses of Evaluation Data

ACADEMIC • IMPRESSIONS 32

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

FOR FORMATIVE DECISIONS

- **Student Ratings** (Teaching)
- **Peer-external ratings** (All topics)
- **Self-ratings** (All topics)
- **Document reviews** (e.g., syllabi)
- **Videos** (Primarily teaching)
- **Student Interviews** (Primarily teaching)
- **Exit & Alum Ratings**(Primarily teaching)

Adapted from Berk, 2006, p. 45

ACADEMIC • IMPRESSIONS ³³

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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 ³⁴

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

35

FOR PROGRAM DECISIONS

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

Adapted from Berk, 2006, p. 45

4. Steps to Analyzing Data

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

Reporting Data

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

Analysis Possibilities For Reports of Results

- Descriptive statistics (item distributions in # and %)
- Central tendency (mean, mode, median)
 - 1 2 3 3 4 5 (3, 3, 3)
 - 1 1 2 3 4 5 5 (3, 1 & 5, 3)
 - 1 2 3 4 5 5 (3.33, 5, 3.5)
 - 1 2 3 4 5 5 5 (3.57, 5, 4)
- Standard deviations (sampling error)
- Enrolled / responded #s and ratio

Class Size - Minimum Acceptable Response

| | |
|--------|-------------------------|
| 5-20 | at least 80% |
| 20-30 | at least 75% |
| 30-50 | at least 60% |
| | 75% or more recommended |
| 50-100 | at least 50% |
| | 66% or more recommended |
| >100 | more than 50% |

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.

Instructional Report of Educational Satisfaction (IRES)

Case Study - Universitas pro Omnibus Discipuli et Facultitas in Excelcis

Instructor: U.N. Fortunate
Course #: HIS123
Course name: History of Everything
Term/year: Spring, 2010

| | A | B | C | D | E | F | O |
|------------------------|---|----|----|----|----|---|---|
| <u>amount learned</u> | 3 | 16 | 46 | 21 | 14 | 0 | 1 |
| <u>overall teacher</u> | 1 | 12 | 40 | 29 | 18 | 0 | 0 |
| <u>overall course</u> | 2 | 8 | 49 | 20 | 11 | 0 | 0 |

46Note: (A) =5= Best; (F)=6=Worst ... Enrolled: 120; Responded: 53

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Instructional Report of Educational Satisfaction (IRES)

Universitas pro Omnibus Discipuli et Facultas in Excelcis

Instructor: U.N. Fortunate
Course #: HIS123
Course name: History of Everything
Term/year: Spring, 2010

| % / # responses > | A | B | C | D | E | F | O | mean | s d | T | grp |
|-------------------|-----|-------|-------|-------|-------|-----|-----|------|------|----|-----|
| amount learned | 3/2 | 16/10 | 46/29 | 21/13 | 14/9 | 0/0 | 1/1 | 3.27 | 0.88 | 24 | low |
| overall teacher | 1/1 | 12/8 | 40/25 | 29/18 | 18/11 | 0/0 | 0/0 | 3.47 | 0.96 | 27 | low |
| overall course | 2/1 | 18/11 | 49/31 | 20/13 | 11/7 | 0/0 | 0/0 | 3.22 | 0.93 | 23 | low |

Raw score: (A) =5= Best; (E) =1= Worst; F= Not applicable; O = Omitted;
Enrolled = 120; Responded = 63: (sample adequate)
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

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Two evaluations of HIS 345

| | mean | s d | T | group | |
|-----------------|------|------|----|---------|------------------------|
| amount learned | 3.35 | 0.87 | 45 | low-mid | term/yr = spring, 1995 |
| overall teacher | 2.76 | 0.76 | 35 | low | instr = UNFortunate |
| overall course | 2.85 | 0.90 | 37 | low | course = His 345 |
| | | | | | resp/enr = 29/61 |
| amount learned | 3.97 | 1.40 | 56 | hi-mid | term/yr= fall, 1995 |
| overall teacher | 3.57 | 1.30 | 47 | mid | Instr = UNFortunate |
| overall course | 3.63 | 1.24 | 50 | mid | course = his 345 |
| | | | | | resp/enr = 20/42 |
| | | | | | % resp=48 |

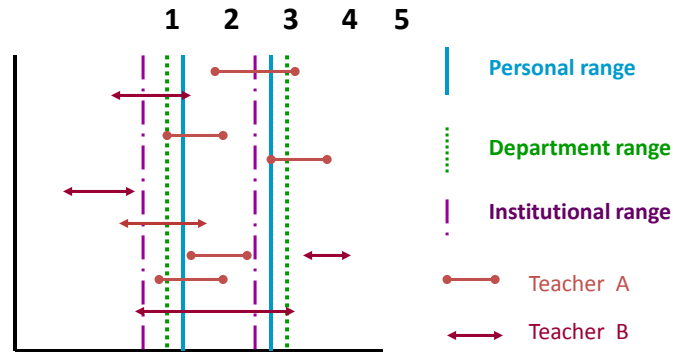
COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Enrollment profiles for HIS 345 in two semesters

| | Fr | So | Jn | Sn | Tot | |
|------------------|----|----|----|----|-----|---|
| original enr. | 6 | 17 | 15 | 23 | 61 | term/yr = spring, 1995 Instr = UN Fortunate |
| final enr. | 5 | 14 | 12 | 20 | 51 | course = his 345 resp/enr = 29/51 |
| eval respondents | 5 | 13 | 11 | 0 | 29 | % resp=57 |
| | | | | | | |
| original enr. | 3 | 11 | 12 | 16 | 42 | term/yr = fall, 1995 Instr = UN Fortunate |
| final enr. | 2 | 7 | 8 | 12 | 29 | course = his 345 resp/enr = 20/29 |
| eval respondents | 2 | 4 | 5 | 9 | 20 | % resp=69 |

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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



Questions?



5. Guidelines to Good Practice

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;**

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.**

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.**

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**

Basic References/Resources

Basic References

- Arreola, R. A. (2007) *Developing a Comprehensive Faculty Evaluation System*. (3rd ed.) Bolton, MA: Anker Publishing Company.
- Berk, R. A. (2006) *Thirteen strategies to measure college teaching*. Sterling, VA: Stylus Publishing.
- Cashin, W. E. (1990) Students do rate different academic fields differently. In M. Theall & J. Franklin (Eds.) "Student ratings of instruction: issues for improving practice." *New Directions for Teaching and Learning # 43*. San Francisco: Jossey Bass.
- Centra, J. A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*. San Francisco: Jossey-Bass.

- Marsh, H. W. (2007) Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases, and usefulness. In R. P. Perry & J. C. Smart (Eds.) *The scholarship of teaching and learning in higher education: an evidence-based perspective*. New York: Springer.
- Scriven, M. (1967) "Methodology of evaluation." In R. Tyler, R. Gagne, & M. Scriven (Eds.) *Perspectives of curriculum evaluation*. Chicago: Rand McNally & Company.
- Theall, M, Abrami, P. A. & Mets, L. (2001) (Eds.), "The student ratings debate. Are they valid? How can we best use them?" *New Directions for Institutional Research No. 109*. San Francisco: Jossey Bass.
- Theall, M. & Arreola, R. A. (2006) "The Meta- Profession of teaching." *NEA Higher Education Advocate*, 22 (5), 5-8, June.
- Theall, M., & Franklin, J. L. (Eds.) (1990). Student ratings of instruction: Issues for improving practice. *New Directions for Teaching and Learning #43*. San Francisco: Jossey Bass.
- Other New Directions volumes on evaluation/ratings: N. D. Teaching & Learning #: 80, 83, 87, 96; N. D. Institutional Research #: 114

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/>

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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

ACADEMIC • IMPRESSIONS 63

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

Questions and Answers



ACADEMIC • IMPRESSIONS 64

COLLECTING, ANALYZING, AND USING FACULTY EVALUATION DATA

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

ACADEMIC • IMPRESSIONS ⁶⁵