From Consensus Building to Skill Acquisition: Teaching an Evidence Based Evaluation Method in Advanced Research Classes

Dr. Violet Cheung-Blunden & Dr. Saera Khan

University of San Francisco

Abstract

The ratings and reliability exercise provides instruction for obtaining valid and reliable evaluations. In the peer rating method, students read articles, write comments, share their views in class, and then obtain ratings from other classmates. Students articulate the kinds of comments that help them learn, fine tune the rating criteria, analyze for technical indicators of the ratings, and compare peer and professor ratings. Over the course of the semester, students were presented with peers' ratings of their reflections as well as the overall reliability of the ratings. Inter-rater reliability increased across the semester as well as their ratings for the value of the exercise. In sum, students' ratings of their confidence in this method at the beginning and end of the semester approached significance in two advanced research methods sections suggesting that students became more confident about this method over the course of the semester.

Introduction

In most advanced research methods (ARM) classes, the primary focus is on teaching students experimentation. However, many students will not become experimentalists per se but will likely apply for employment positions in business or intervention programs. Many of these occupations require their ability to evaluate personnel or programs. The purpose of the ratings and reliability exercise is to provide students another research-based tool: obtaining valid and reliable evaluations.

In the peer rating method, students read ten assigned articles relevant to the course topic. Then, students are asked to electronically post their reflections on the readings. The instructions for the reflection assignment emphasize that they must contribute original thought or pose a question designed to engender class discussion.

Procedure

Instructions For Reflection Assignment

These (almost) weekly comments should be a brief comment, critique, idea, or question bearing on some aspect of the assigned readings. During the class meeting, everyone should be prepared to convey the nature of their comment/question to the rest of the class. These comments should not exceed one- two pages in length. Although the specific format of it is up to you, here is a suggestion for how to start it.

- very concisely, what were the core ideas of the readings?
- what are the major conclusions drawn by the authors?
- substantively, what points raised in a reading sparked insight or questions for you OR substantively, what aspects of the reading covered warrant further investigation and briefly how would you study this question

Keep in mind: you should have something of "you" (i.e. original) in it. Summarizing the readings is not acceptable. Credit will be awarded for papers as long as they are posted by the start of class. In order to be fair to all students, no credit can be given for papers turned in after class.

Class Discussion

Prior to class, the instructor reads the reflections so that common themes and interesting points can be brought to class discussion. In class, students discuss the readings and share their views.

Operationalizing "good" and "well-written"

At the beginning of the course a class discussion is devoted to operationalizing what we mean by good and well-written comments. This discussion allows the students to have the same metric by which they are to judge their peer's comments. The class consensus centered around the following characteristics.

Operationalization of "well-written"

- Correct English
- Structure and flow
- Clarity of the main idea
- Concrete connection to the reading

Operationalization of "good"

- Correct understanding of the reading
- Thoughtful analyses
- Novel point

Reflection Ratings

After class, the instructor compiles all the reflections, removes names and assigns coded initials to each one. The instructor posts this document along with an excel sheet with the codes and two rating columns for each writer.

Students are asked to rate, along a scale of 1-5, each anonymous reflection on the basis of two dimensions: "well written" and "good."

This tool was implemented in two different sections of Advanced Research Methods.

Analyses

Sample ratings of comments as "good" (1-5)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AC	16	1.00	5.00	3.8750	1.02470
AH	15	2.00	5.00	3.7333	.79881
AK	15	2.00	5.00	4.0000	1.00000
AM	0				
CA	15	2.00	5.00	3.8000	.86189
DA	15	2.00	5.00	3.9333	.96115
ER	15	2.00	4.00	3.4000	.73679
EY	14	3.00	5.00	3.5000	.65044
GE_1	15	2.00	5.00	3.6000	.82808
GS	15	3.00	5.00	3.6667	.72375
IE	15	2.00	5.00	3.5333	.91548
IO	15	3.00	5.00	4.0000	.84515
NA	0				
NE_1	0				
OD	15	3.00	5.00	4.4000	.73679
UM	0				
YE	15	3.00	5.00	3.4667	.74322
Valid N (listwise)	0				

Inter-Rater Reliability Analyses

Reliability Statistics

Cronbach's Alpha	N of Items
.799	32

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Bac_8G	126.647	164.743	.298	.	.907
Bad_8G	127.176	164.654	.316	.	.907
Bai_8G	127.118	158.860	.494	.	.904
Barn_8G	126.824	155.904	.666	.	.901
Base_8G	126.529	163.390	.501	.	.904
Bour_8G	126.706	162.471	.661	.	.902
Cej_8G	126.529	163.015	.455	.	.904
Dan_8G	127.176	163.404	.375	.	.906
Ju_8G	126.706	162.096	.575	.	.903
Lee_8G	127.118	158.860	.494	.	.904
Lem_8G	127.118	157.235	.661	.	.901
Lut_8G	126.176	164.779	.473	.	.904
Magli_8G	126.118	162.610	.601	.	.903
Pha_8G	125.941	165.934	.369	.	.906
Pag_8G	126.882	161.985	.629	.	.902
Sta_8G	126.000	168.125	.290	.	.906
Wo_8G	127.176	162.779	.259	.	.910
Bac_8WW	126.765	159.566	.471	.	.904
Bad_8WW	127.000	161.875	.396	.	.906
Bai_8WW	126.235	157.191	.690	.	.901
Barn_8WW	126.588	163.132	.427	.	.905
Base_8WW	126.529	163.640	.585	.	.903
Bour_8WW	126.588	168.382	.438	.	.906
Cej_8WW	126.235	162.066	.449	.	.905
Dan_8WW	126.941	162.434	.509	.	.904
Ju_8WW	126.941	165.309	.409	.	.905
Lee_8WW	126.235	158.566	.621	.	.902
Lem_8WW	126.647	165.368	.340	.	.906
Lut_8WW	126.529	161.765	.601	.	.903
Magli_8WW	126.529	167.765	.291	.	.906
Pha_8WW	125.706	168.971	.273	.	.907
Page_8WW	126.882	161.860	.546	.	.903
Sta_8WW	126.118	166.110	.466	.	.905
Wo_8WW	127.118	158.235	.370	.	.908

Students see the class reliability as well as their own own performance as a rater.

Pre and Post Measures of Self-Reported Confidence in the Peer Rating System

Students responded to 11 items pertaining to their confidence in the peer rating method at the beginning and at the end of the semester. The scale ranged from 1 (strongly disagree) to 7 strongly agree).

1. I believe students' comments on the assigned readings can be objectively evaluated.
2. I believe students' rating of their peers' work can reach a reasonable agreement.
3. I believe my peers' rating of my work will be similar to my instructor's rating.
4. I believe students' written comments can be graded with respect to their contribution to the learning experience in classroom
5. I'm confident about my ability to rate my peers' comments
6. I'm concerned about my peers' ratings (I don't trust my peers rating my work)
7. Rating written comments is a skill that can be learned through practice (I expect to learn about evaluating other students' work)
8. I believe that the difference between poorly written and well-written comments is apparent to all of my peers.
9. I believe that my peers can detect the difference between substantive (i.e., "good") and empty (i.e., "superficial") comments.
10. I believe that peer ratings can accurately reflect the quality of students' work
11. I believe that peer ratings should be included as a component of the final grade.

The Cronbach's alphas of the scale for the first section were $.23_{pre}$ and $.85_{post}$; for the second section $.65_{pre}$ and $.80_{post}$.

The difference in the mean scores between pretest and posttest approached significance in both sections (section 1: $t(9)=1.93$, $p=.086$; section 2: $t(13)=1.78$, $p=.098$), suggesting that students became more confident about the peer rating method over the course of the semester.

Conclusion

Through these exercises, students begin to trust the reliability and validity of psychological measurement. A benefit outside a measurement class is the ability dictate a direction to improve student comments as well as class discussions.

RESEARCH POSTER PRESENTATION DESIGN © 2015
www.PosterPresentations.com