

Undergraduate Psychology Program

Inter-Rater Reliability for Undergraduate Integrative Paper and Marianist Educational Values Paper

In order to determine inter-rater reliability for the psychology undergraduate *Integrative Paper* rubric and the *Marianist Educational Values Paper* rubric, papers from eleven students in the Spring 09 Senior Seminar in Psychology course (Psy 490) were assessed, using the total score from the rubrics, by two members, one who currently teaches the class, of the faculty. Intra-class consistency correlation coefficients were derived using SPSS (see *SPSS Base 10.0 Applications Guide*, 366-368; also see *Reliability Analysis*, <http://faculty.chass.mcsu.edu/garson/PA765/reliab.htm>, 9-12).

Integrative Paper

The inter-rater reliability coefficient is .7468. Both the single measure ($r=.5959$) and average measure ($r=.7468$) intraclass correlation were significant ($p=.0204$).

Marianist Educational Values Paper

The inter-rater reliability coefficient is .8581. Both the single measure ($r=.7514$) and average measure ($r=.8581$) intraclass correlation were significant ($p=.0024$).

Analysis

The results are indicative of **acceptable** inter-rater reliability for both rubrics (see Stemler, Steven E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9(4). Retrieved June 4, 2009 from <http://PAREonline.net/getvn.asp?v=9&n=4>. This paper has been viewed 35,985 times since 3/1/2004.).

Scores that were deemed to be questionably far apart by a third faculty member resulted in the two judges sitting down and going over the individual papers, questions and scores to determine the basis of the significant discrepancy. This resulted in or will result in a more detailed explanation for certain aspects of the rubric relative to specific questions.

Any faculty member who teaches the Senior Seminar in Psychology course (Psy 490) will have to be pre-qualified (inter-rater reliability) on the rubrics, using the papers from the same eleven students, for both papers, before teaching the course. Their assessment will be compared with the two faculty members' assessment that was used for this study. If their inter-rater reliability does not meet the acceptable inter-rater reliability, they will not be allowed to teach the course. Appropriate instruction will be provided both before and after, if necessary, on how to utilize the rubrics.