
**Researcher Notes**

**Categories addressed:** Personnel selection.

**Participants:** Review of Literature.

**Method:** A meta-analysis of 111 interrater reliability coefficients and 49 coefficients alphas from selection interviews. The sample was drawn from 82 studies, with a total sample size of 20,639. The analysis included study design, interviewer training, and 3 dimensions of interview structure (standardization of questions, of response evaluation, and of combining multiple ratings).

**Independent variable:** Interview structure (structured and unstructured, standardization of questions, response evaluation and multiple ratings), job analysis and interviewer training.

**Dependent variable:** Internal consistency of panel vs. separate interviews on applicant performance.

**Findings:**

- Reliabilities for panel interviews (M=.77, variance=.019) were greater than for separate interviews (M=.53, variance=.038). However, results showed that standardization of questions had effects for both separate-and panel interview coefficients but a considerably stronger effect for separate interviews.
- Estimates of upper limits of validity were .67 for highly structured interviews, .56 for moderate structured interviews, and only .34 for interviews with low structure. In this sense, the best way to increase validity is likely to be increasing structure.
- One explanation is that standardized questions provide a more consistent sample of applicant performance. Standardized questions should therefore result in more consistent measurement of applicant qualifications across multiple interviews than should unstandardized questions. In other words, interviewers can evaluate the information more successfully because standardized questions help interviewers to rate better.

**Practice Implications:**

- The researchers suggest the use of standardized questions because they have a dramatic effect on psychometric properties of interview ratings.
- Results clearly support mechanical (e.g. averaging or summing) combination of ratings (consistent with psychometric theory). They found no evidence that multiple ratings increased reliability when they were combined subjectively. In
sum, they recommend multiple ratings (ideally for each question) that are combined mechanically.

- Researchers also recommend the use of some form of interviewer training, because reliabilities were higher when interviewers were trained.