Extra credit assignment

For extra credit, you may write up to 2 different empirical article summaries/reviews on one of the OB topics we talked about this semester – each paper is worth up to 2 points towards your final grade. Each extra credit assignment should be at least 2 pages (double-spaced, 1” margins, 12-point font), following the example given below. Assignments should be handed in (or emailed to me) no later than December 3rd.

For full credit, your review/summary should include the following points:

- An APA or AMJ-style reference at the top of the page
- A thorough summary of the topic and purpose of the study, the procedures, and the results
- A critique including at least 2 highlights or contributions you think the article made, and at least 2 criticisms of the article
- Appropriate spelling and grammar

For reference, you may follow the example paper given on the next page.
Empirical Article Summary/Critique Example


Summary

The authors of this article point out that organizations frequently make decisions about which predictor measures to use in a selection context. They frequently choose between commercially available tests of general aptitudes and tests designed to measure job-specific knowledge, skills, and abilities (KSAs). These tests each have their own advantages and disadvantages. For example, tests of general aptitudes are relevant to success across diverse jobs and are most cost effective. On the other hand, tests of job-specific KSAs tend to be perceived as fair by both the test takers and the courts and therefore, have higher levels of user acceptance. Given the initial investment required to develop job-specific KSA tests, the authors set out to assess the extent to which the tests measure the same thing. Specifically, they examine the overlap in the constructs measured (construct equivalence) by the two types of tests.

The authors gathered responses to a six tests from 3,956 applicants for a journey-level apprenticeship training program at a large manufacturing firm. The applicants sought positions in one of eight skilled trades: tool and die maker, tool maker, die maker, machine repair, millwright, plumber/pipefitter, electrician, and welder. The authors examined the equivalence of three constructs measured by the six tests (3 of each kind). They looked to see if verbal reasoning was equally measured by the Verbal Reasoning subscale of the Differential Aptitude Test (DAT) and a job-specific measure labeled Technical Reading. They examined the extent to which The Numerical Ability subscale of the DAT and a job-specific measure labeled Industrial Math equally measured numerical ability. Finally, they examined the Space and Visualization subtest
of the Employee Aptitude Survey as well as a job-specific measure labeled Following Instructions to see if they measured spatial visualization equally.

The authors conducted a series of confirmatory factor analyses to assess the extent to which the data fit a variety of models. They also performed multiple groups analysis to assess the extent to which the models fist consistently across race and sex subgroups. Overall, the authors supported their hypotheses that the tests measured the same constructs. In addition, the tests assess their underlying constructs equivalently across various race and gender.

Critique

I found this to be an interesting article that addresses a real organizational problem. I think it is very important for organizations to know that there are tests that measure constructs of interest equally well and can be purchased less expensively. I also think it is crucial to conduct a construct equivalence study like this one to demonstrate equivalence before administering the less expensive tests. This article provides an example of how to demonstrate construct equivalence across tests that researchers can use when facing similar questions and problems. For example, the authors cite another study that used construct equivalence for the purpose of maintaining test security.

My only criticism of the article is that the authors do not compare how these tests predict the criterion of interest, job performance. While they appear to measure the same thing, it is important to assess the extent to which they equally predict a designated criterion. Also, how the tests measure the same construct across groups should be further explored. Again, criterion-related validity should be examined for various subgroups. At the same time, I must point out that the authors acknowledge my criticisms as limitations and directions for future research in the discussion section of their article.