

Criterion-Related and Differential Validity of ACT® WorkKeys® in Arkansas

Jeffrey Conway, PhD

Introduction

In the winter of 2020, ACT and the Arkansas Department of Higher Education (ADHE) entered into a data sharing agreement where the ADHE would share educational data on several cohorts of incoming college freshmen and ACT would match that data to its test records. The original purpose of this work was to create high school to college success reports using ACT® test scores as a predictor. The purpose of this paper is to outline an additional use of this data by examining the validity of ACT® WorkKeys® scores in predicting first-year college GPA and to examine if any differential validity, based on gender or ethnicity, existed in predicting first-year college GPA. ACT WorkKeys tests students on foundational workplace skills (e.g., mathematics, reading comprehension, interpretation of graphics). Moreover, students can earn an ACT® WorkKeys® National Career Readiness Certificate® (NCRC®), which certifies skills to potential employers and postsecondary training programs (ACT, 2014). The NCRC is determined based on three assessments: ACT® WorkKeys® Applied Math, ACT® WorkKeys® Workplace Documents, and ACT® WorkKeys® Graphic Literacy.

Data

To start, a unique identifier using first name, last name, and date of birth was created to find matching WorkKeys Assessment scores. Next, a unique dataset was created which contained WorkKeys scores and first-year college GPA for students within the Arkansas educational system at either a 2- or 4-year institution. The students took WorkKeys in high school prior to enrollment. This dataset contained 964 students. The sample was 52% female. The sample was 63% White, 20% Black/African American, 11% Hispanic/Latino, with the remaining proportion of the sample coming from Asian (1%), Native American (0.5%), Pacific Islander (0.5%), and Multiracial groups (4%). Lastly, the distribution of NCRC distinctions were as follows: Non-Qualifier (9%), Bronze (26%), Silver (32%), Gold (22%), and Platinum (11%). Descriptive statistics of the key variables of the study are in Table 1.

Table 1. Descriptive Statistics

Variable	Mean	SD	Scale
GPA	3.25	0.63	0–4.33
WK Applied Math	79.12	4.58	65–90
WK Graphic Literacy	79.26	4.17	65–90
WK Workplace Documents	79.35	4.02	65–90

Analyses

Correlation Coefficient Overview

This section relies heavily on correlation coefficients, and it is important to understand how to interpret correlation coefficients to get the most out of this portion of the paper. Correlations range from -1.00 to $+1.00$ and evaluate the strength of the relationship between two variables. Any correlation above zero means that the two variables being examined move in the same direction (i.e., one value goes up, the other goes up too). The opposite is true for negative correlations (i.e., one value goes up, the other goes down). The farther the correlation coefficient is from zero, the stronger the relationship (e.g., a correlation of -0.50 is stronger than a correlation of $+0.40$). According to Cohen (1988), a correlation of 0.10 represents a small effect, 0.30 a medium effect, and 0.50 a strong effect. To provide some meaningful reference points, several correlations using variables most people are familiar with are as follows:

- Weight and height: 0.42
- Education and income: 0.33
- Interview ratings and job performance: 0.40
- Time spent playing video games and grades: -0.25

Correlation Coefficient Results

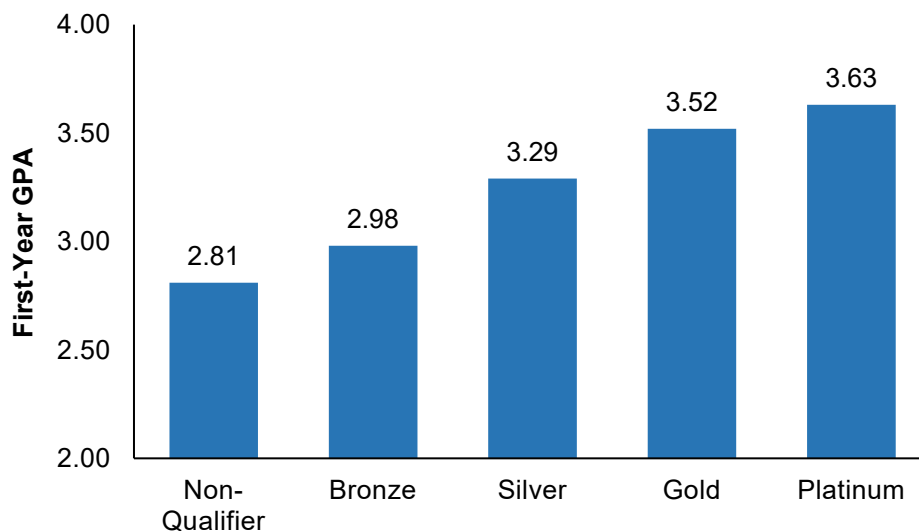
The correlations between WorkKeys scores and first-year college GPA are listed in Table 2. All the correlations in Table 2 were statistically significant ($p < .01$). Interestingly, the correlation between each WorkKeys Assessment and GPA was the same (0.40). These correlations indicate a strong positive correlation between WorkKeys scores and first-year college GPA.

Table 2. WorkKeys Validity Coefficients

WorkKeys Assessment	GPA
Applied Math	0.40
Graphic Literacy	0.40
Workplace Documents	0.40
NCRC*	0.41

* Treated as an ordinal variable with five equal interval levels

In order to provide a different vantage point for the results in Table 2, Figure 1 shows the average GPA at each NCRC level. Figure 1 displays a stair-step pattern where average GPA increases as a function of NCRC level. Also of note, the largest jump in GPA occurs between the Bronze and Silver NCRC levels.

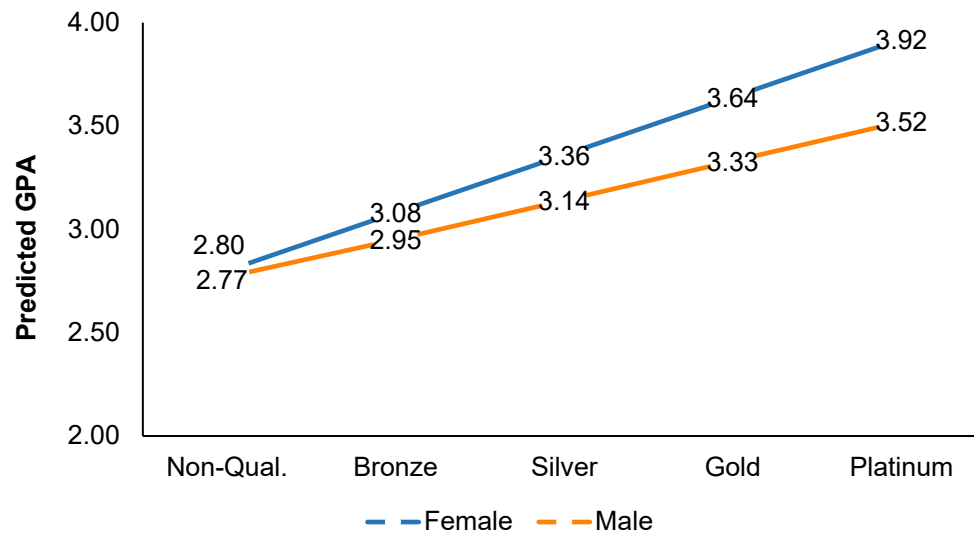
Figure 1. Average First-Year GPA by NCRC Level

Differential Validity Overview

Differential validity occurs when the relationship between test scores and an outcome varies by subgroup, which could suggest that the interpretation of scores should vary by subgroup. Differential validity is examined using regression where interactions between test scores and subgroups are tested to determine whether best-fit regression lines differ by subgroup. According to the *Standards for Educational and Psychological Testing* (2014, p. 70), “In cases where there is credible evidence of differential validity, developers should provide clear guidance to the test user about when and whether valid interpretations of scores for their intended uses can or cannot be drawn for individuals from these subgroups.” Based on this standard, it is important to test whether WorkKeys Assessments create differential validity when predicting first-year GPA based on gender and race/ethnicity. The differential validity for WorkKeys has not been empirically tested before.

Differential Validity Results

The first differential validity analysis was focused on gender. The regression test using gender produced a significant interaction between NCRC level and gender ($\beta = -0.25, p < .01$). The direction of the interaction favored men over women. Hence, NCRC level overpredicts the GPA level of men (versus women). The results of this analysis are plotted in Figure 2, which shows a stronger slope in the blue line for women as opposed to the orange line for men, creating a difference in intercepts for the separate regression equations. Stated simply, women outperform predictions, relative to men, regarding GPA. As women continue to outnumber men in higher education, this small effect could serve to close this gap. This is a common finding across ACT research and one reason for the difference could be that women tend to have higher psychosocial factors (Ndum et al., 2015).

Figure 2. Differential Validity Analysis for Gender

Two differential validity analyses were run based on race/ethnicity. First, an analysis comparing Black/African American test takers to White test takers did not produce a statistically significant interaction ($\beta = -0.08$, $p = n.s.$). Second, an analysis comparing Hispanic/Latino test takers to White test takers also did not produce a statistically significant interaction ($\beta = -0.07$, $p = n.s.$). These analyses show that the relationship between NCRC levels and first-year GPA do not vary based on race/ethnicity for Black/African American and Hispanic/Latino test takers.

Summary

On the whole, WorkKeys shows a strong positive relationship with first-year GPA in a sample of 964 students in Arkansas. Students who score high on WorkKeys are also much more likely to have higher GPAs. In addition, the differential validity analyses showed that WorkKeys underpredicts GPA for females and shows no significant differences for Black/African American and Hispanic/Latino test takers (as compared to White test takers). Thus, WorkKeys predicts first-year GPA and does so in an equitable fashion based on these data.

Like all studies, this research has limitations which open avenues for future research. First, future studies should consider differential validity for students who identify as Asian, Native American, and Pacific Islander (assuming large enough sample sizes can be collected for each subgroup). Similarly, the results for the Black/African American and Hispanic/Latino groups should be replicated in another sample with an even larger sample if possible. High statistical power is required to find significant interaction effects in regression. Additional groups such as English learners, students with disabilities, and first-generation college students should also be tested. Second, results of this study should be replicated and expanded using job performance ratings as the primary criterion. While WorkKeys is designed around workplace contexts, it measures some core academic skills that should also be related to success in college.

First-year GPA was used in the current study because of the ease with which a large sample could be collected. Collecting a large enough sample of standardized job performance ratings presents many challenges, including the ever-present criterion problem often encountered in applied psychology (Austin & Villanova, 1992). The gist of the criterion-problem is that job performance is difficult to measure as it is a multidimensional construct and biases in ratings are common. Replication of the present study using job performance ratings would provide additional evidence that WorkKeys scores predict the same chance at achieving success in the workplace across different groups of examinees.

References

- ACT. (2014). *ACT National Career Readiness Certificate™*. Retrieved from <http://forms.act.org/certificate/pdf/NCRC-InformationFlyer.pdf>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917–1992. *Journal of Applied Psychology*, 77(6), 836–874.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Routledge.
- Ndum, E. A., Allen, J., Way, J. D., & Casillas, A. (2015). *Effects of psychosocial factors on gender differences in first-year college English and mathematics grades*. ACT.



ABOUT ACT

ACT is a mission-driven, nonprofit organization dedicated to helping people achieve education and workplace success. Grounded in more than 60 years of research, ACT is a trusted leader in college and career readiness solutions. Each year, ACT serves millions of students, job seekers, schools, government agencies, and employers in the U.S. and around the world with learning resources, assessments, research, and credentials designed to help them succeed from elementary school through career.

For more information, visit act.org