
Evaluating the Validity of WorkKeys in the Mississippi Community College System

Jeffrey Conway

Introduction

In the fall of 2020, ACT and the Mississippi Community College Board (MCCB) entered into a data sharing agreement in order to examine the validity of ACT's® WorkKeys® Assessments. WorkKeys Assessments measure students' foundational workplace skills. Moreover, students can earn an ACT® WorkKeys® National Career Readiness Certificate® (NCRC®)¹, which certifies skills to potential employers and postsecondary training programs (ACT, 2014). The NCRC is determined on the basis of three assessments: Applied Math, Workplace Documents, and Graphic Literacy. The focus of this technical brief is to summarize the results of several validity studies using a sample from MCCB. This brief examines three different types of validity: (1) convergent, (2) discriminant, and (3) criterion-related. To examine convergent and discriminant validity, the relationship between WorkKeys and the Test of Adult Basic Education (TABE) was examined. To assess criterion-related validity, the relationships between WorkKeys and several educational outcomes were examined.

Data

MCCB shared student data with ACT from its Smart Start course, where students are prepared to obtain new skills for future careers within their region's workforce sector, identifying the career components that are necessary for middle-skill employment. MCCB also provided information on these students regarding post-secondary GPA, TABE scores, credential completion, and high school equivalency attainment. ACT merged this information into a larger dataset containing WorkKeys scores and demographic information. The final dataset was 57% female and 43% male. The majority of the sample was Caucasian (57%), but there were also a significant number of African American students (38%) and a meaningful percentage of Hispanic students (4%). The sample sizes in the study varied by analysis due to missing data (e.g., not all Smart Start students took TABE). The sample size will be mentioned for each specific analysis in the forthcoming Analyses section.

Analyses

Background on Correlation Coefficients

This section relies heavily on correlation coefficients, and it is important to understand how to interpret them to get the most out of this portion of the paper. Correlations range from -1.00 and +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 or -0.40). The following examples provide some meaningful reference points using variables most people are familiar with.

- The correlation between weight and height is 0.42.
- The correlation between education level and income is 0.33.
- The correlation between interview ratings and job performance is 0.40.
- The correlation between time spent playing video games and grades is -0.25.

Convergent and Discriminant Validity

First, in order to study convergent and discriminant validity², the relationship between WorkKeys and TABE was examined. TABE is composed of three assessments: (1) Language, (2) Reading, and (3) Math. In order to meet the requirements for the National Reporting System for Adult Education (NRS), new versions of TABE are consistently in development. In this study, the two most recent version of NRS-approved TABE tests were versions 11 and 12. The expectation regarding convergent validity was that all the assessments would be positively related as they all measure some aspects of intelligence, which is a unidimensional construct (Spearman, 1961). The expectation regarding discriminant validity was that WorkKeys Applied Math would be more strongly related to TABE Math than to TABE Language and Reading. Conversely, WorkKeys Workplace Documents was expected to be more strongly related to TABE Language and Reading than to TABE Math. There were no hypotheses regarding WorkKeys Graphic Literacy, but we do include that assessment in the results. Table 1 shows the correlation results between WorkKeys and TABE. Sample sizes for the analyses ranged from 1,799 to 2,190.

Table 1. Relationship Between WorkKeys and TABE³

TABE Version	Applied Math	Graphic Literacy	Workplace Documents
TABE11 Language	0.47	0.52	0.55
TABE11 Reading	0.52	0.54	0.59
TABE11 Math	0.52	0.47	0.36
TABE12 Language	0.52	0.57	0.59
TABE12 Reading	0.54	0.58	0.63
TABE12 Math	0.57	0.54	0.43

The results for convergent validity were straightforward as all the assessments were significantly positively related to one another. The results for discriminant validity were mixed. Applied Math did not demonstrate discriminant validity as it had similar relationships with each of the different TABE tests across versions. However, Workplace Documents did demonstrate discriminant validity as it showed higher correlations with TABE Language and Reading than with TABE Math.

In more general terms, the results in Table 1 indicate that students who score high on one of the assessments are more likely to score higher on all the other assessments (i.e., a positive correlation). Conversely, students who score low on one of the assessments, for example on WorkKeys Applied Math, would also be expected to have low scores on all the other assessments. Overall, students have similar scores on TABE and WorkKeys. The only area where this relationship breaks down is when comparing Workplace Document scores to TABE Math scores (but not vice versa) across the assessments as the correlations were weaker (i.e., closer to zero).

Criterion-Related Validity

Second, criterion-related validity was examined by calculating the correlations between WorkKeys, NCRC, high school equivalency attainment, GPA, and credential attainment. Criterion-related validity focuses on the ability of an assessment to predict outcomes of interest. Means and standard deviations for the key variables in this section are listed in Table 2.

Table 2. Means and Standard Deviations

Key Study Variables	Mean	SD
Applied Math	76.14	3.86
Graphic Literacy	76.92	3.49
Workplace Documents	78.00	3.57
NCRC	2.45	0.91
High School Equivalency	0.73	0.45
GPA	2.63	1.01
Credentials	1.13	0.40

High School Equivalency Attainment

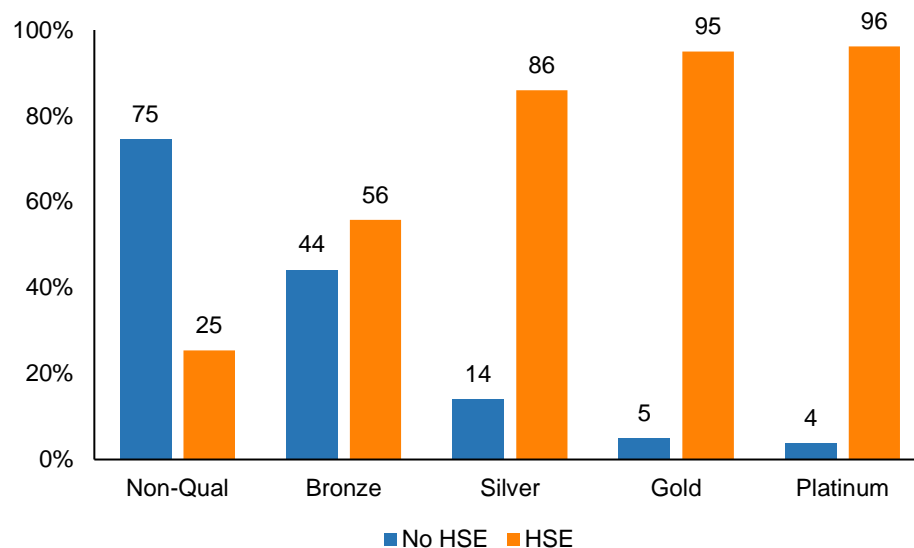
The correlations between the three WorkKeys assessments, NCRC, and high school equivalency attainment (coded as 0 for not attained; 1 for attained) are listed in Table 3. The sample sizes for all correlation analyses in Table 3 were 1,377. The correlations in Table 3 are indicative of a moderate to strong relationship between WorkKeys, NCRC, and a student earning their high school equivalency. Hence, students with high scores on WorkKeys are more likely to earn their high school equivalency.

Table 3. Relationships Between WorkKeys, NCRC, and High School Equivalency⁴

Predictors	High School Equivalency
Applied Math	0.39
Graphic Literacy	0.36
Workplace Documents	0.35
NCRC	0.41

An easier way to interpret the correlations in Table 3 is to calculate the percentage of students at each NCRC level who earned their high school equivalency. Those results are captured in Figure 1. The figure shows a consistent stair-step pattern where the percentages in the orange bars, representing the percentage of Smart Start students who earned their high school equivalency, consistently increased as a function of higher NCRC levels. A related analysis using a dichotomous coding of NCRC (0 = No NCRC; 1 = NCRC Credential) showed that students who earned their NCRC were three times more likely (25% vs. 75%) to earn their high school equivalency ($\chi^2(1) = 84.93, p < .01$).

Figure 1. Percentage of Students Who Earned High School Equivalency (HSE) by NCRC Distinction



To combine elements of the first two sets of analyses, two hierarchical logistic regressions were run to see if the three WorkKeys Assessments could provide incremental prediction for high school equivalency attainment over and above the three TABE assessments⁵ and vice versa. The first hierarchical logistic regression showed that WorkKeys provided significant incremental prediction of high school equivalency attainment beyond TABE. In this regression, TABE accounted for 29% of the variance in high school equivalency completion, and the amount of variance accounted for increased to 34% after adding the WorkKeys assessments to the model ($\Delta R^2 = 5\%$, $p < 0.01$). In the reverse analysis, the WorkKeys Assessments accounted for 28% of the variance, and the amount of variance accounted for increased to 34% after adding the TABE assessments to the model ($\Delta R^2 = 6\%$, $p < 0.01$). The sample sizes for each of these analyses was 688. In sum, both TABE and WorkKeys scores strongly predict high school equivalency completion, and each test provides a unique ability to predict this outcome beyond the other.

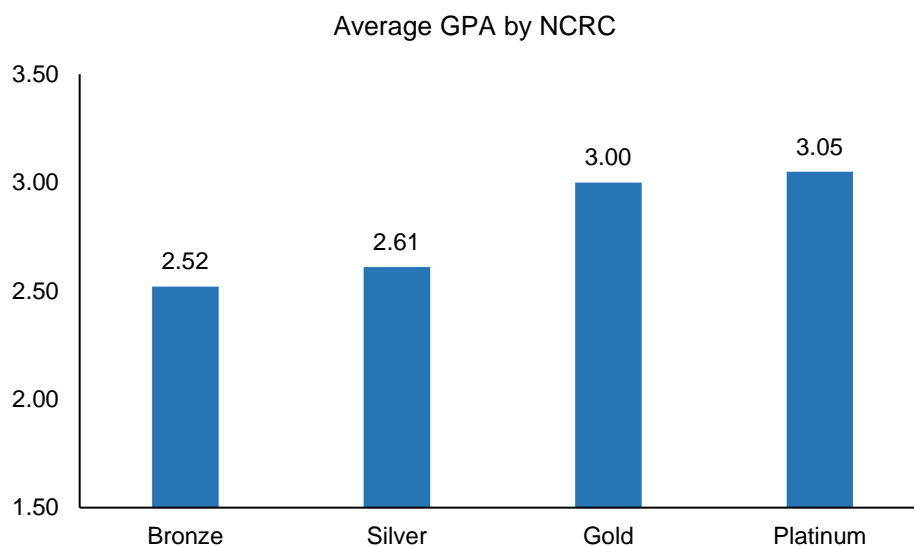
GPA

For the GPA analysis, MCCB provided ACT with Smart Start students' most recent year's GPA, and 455 of those students also had WorkKeys scores. Table 4 shows the correlations between WorkKeys, NCRC, and GPA. All of the correlations in Table 4 were statistically significant ($p < 0.01$) and in the positive direction. Hence, students with higher scores on the NCRC and its component assessments had, on average, higher GPAs.

Table 4. Relationships Between WorkKeys, NCRC, and GPA

Score	GPA
Applied Math	0.16
Graphic Literacy	0.14
Workplace Documents	0.12
NCRC	0.17

The correlations in Table 4 reflect a small relationship between WorkKeys and GPA. However, a larger study on GPA with the state of Arkansas recently found a correlation of 0.41 between NCRC and first-year college GPA using a sample of 964 students (ACT, 2023). As was the case with high school equivalency, the correlations are easier to interpret by calculating the average GPA within each NCRC level (see Figure 2). Figure 2 shows that as the NCRC level advances, the average GPA increases as well. In Figure 2, the non-qualifier level is omitted because there were only two students with GPA information who were non-qualifiers. The lack of non-qualifiers in this analysis likely created range restriction and suppressed the correlation. This could explain the differences in correlation between this study and the one in Arkansas. The sample sizes of the other NCRC levels ranged from 27 (Platinum) to 188 (Silver).

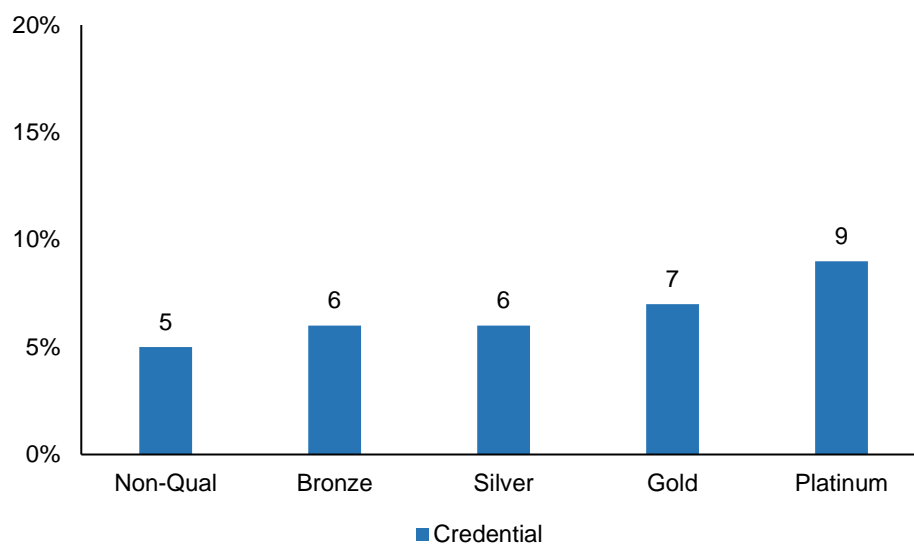
Figure 2. Average GPA by NCRC Distinction

Credentials

MCCB provided ACT with information on 240 individuals who had received at least one credential. Credentials provide an alternative to a 2- or 4-year degree by providing information on job-relevant skills in a shorter period of time and at a lower cost. The participants in the current study reported earning over 50 unique credentials in areas such as commercial driving,

digital literacy, forklift operation, welding, and medicine. The main difficulty with the credential analysis was that the remaining participants in the dataset had no information on credentials. Thus, it could not be confirmed if the remaining participants truly had not earned any credentials or had just not reported them. As a result, the credential analysis has a low base-rate of individuals who earned a credential (about 6%). With those caveats in mind, a chi-square test of independence was used to compare the relationship between earning a credential (0 = No; 1 = Yes) and NCRC level (0 = Non-Qualifier; 4 = Platinum) for 3,053 participants. The result of this analysis was non-significant ($\chi^2(4) = 2.78, p = n.s.$). To provide further insight into this analysis, Figure 3 shows the percentage of participants who earned a credential at each NCRC level. Figure 3 does show a stair-step pattern where percentages increase as NCRC level increases. However, the differences between each level are small.

Figure 3. Credential Completion Percentage by NCRC Level



Summary

The results of the convergent and discriminant validity analyses using TABE were mixed. Strong evidence of convergent validity was found as all WorkKeys Assessments were moderate to strongly related to all TABE assessments (across two versions of TABE). Evidence of discriminant validity was found for WorkKeys Workplace Documents but not WorkKeys Applied Math. The evidence of criterion-related validity of WorkKeys in predicting high school equivalency completion and GPA was strong. WorkKeys was positively and significantly related to both outcomes. In addition, WorkKeys and TABE each had essentially the same ability to predict high school equivalency attainment. The analyses using credential completion were in the hypothesized direction (higher NCRC scores associated with earning a credential) but were not statistically significant and did not represent a meaningful effect size.

One of the key strengths of this study was the sample sizes we were able to obtain for all the analyses; however, this study suffers from several limitations which can be improved upon in

future work. First, a better test of discriminant validity for WorkKeys would be using non-cognitive scales (e.g., openness to experience) to show they do not correlate at all, or at least less, with any of the WorkKeys Assessments (i.e., a correlation closer to 0). Second, there was range restriction in the GPA values we obtained as many low scorers on WorkKeys likely did not continue their education. As mentioned previously, a larger study in Arkansas found much higher correlations between WorkKeys and GPA.

Third, and lastly, credential completion was a difficult variable to analyze in this study because we could not confirm whether a student with a status of no credential truly had not earned any credentials or had just not reported them. As a result, the base-rate of credential completion was quite low. In addition, different credentials might have different relationships with WorkKeys scores. For instance, WorkKeys Applied Math might be a better predictor of earning a forklift certification than WorkKeys Workplace Documents. There were also several certificates, such as the CPR certification, that might be unrelated to any of the WorkKeys assessments. Future studies should attempt to address these limitations.

Overall, the validity evidence for WorkKeys was quite good, though research design in future studies can help improve discriminant validity and criterion-related validity in predicting credential completion. Future studies on the relationship between WorkKeys and GPA with larger sample sizes would also help determine the true population relationship.

References

- ACT. (2014). *ACT National Career Readiness Certificate™*.
<http://forms.act.org/certificate/pdf/NCRC-InformationFlyer.pdf>
- ACT (2023). *Criterion-Related and Differential Validity of WorkKeys in Arkansas* [Manuscript in preparation].
- Spearman, C. (1961). "General Intelligence" Objectively Determined and Measured. In J. J. Jenkins & D. G. Paterson (Eds.), *Studies in individual differences: The search for intelligence* (pp. 59–73). Appleton-Century-Crofts.

Notes

¹ The scoring for NCRC is as follows: Non-Qualifier, Bronze, Silver, Gold, and Platinum.

² Convergent validity evidence exists when similar constructs are positively related whereas discriminant validity evidence exists when dissimilar constructs are unrelated or display smaller relationships.

³ All correlations significant at $p < .01$

⁴ All correlations significant at $p < .01$

⁵ For these analyses, TABE version 12 was used because it was the most recent NRS approved version of TABE in our dataset.



ABOUT ACT

ACT is a mission-driven, nonprofit organization dedicated to helping people achieve education and workplace success. Grounded in 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