Original Research

Exceptional Children
1–18

DOI: 10.1177/0014402917731557
ec.sagepub.com
SAGE

Modeling College and Career Readiness for Adolescents With and Without Disabilities: A Bifactor Approach

Allison Lombardi, PhD[1], Jennifer Freeman, PhD[1], and Graham Rifenbark, BA[1]

Abstract
Nonacademic skills related to college and career readiness (CCR) have become more prevalent in the literature as proposed conceptual models and frameworks, yet little empirical research exists in their support. We employed latent variable modeling to empirically test a previously proposed six-domain framework of CCR for adolescents with and without disabilities. Results support four specific factors of CCR: Academic Engagement, Critical Learning Processes, Mind-Set, and Transition Knowledge. Using a bifactor model, we confirmed one general factor (CCR) and one specific factor (Transition Knowledge), established measurement invariance on the basis of disability, and found latent mean differences between these groups; students without disabilities had greater overall CCR and transition knowledge. Findings support the use of a CCR measurement model with two potential factor scores in future research and practice and may inform efforts to measure CCR nonacademic skills.

Adolescents with and without disabilities must be college and career ready to be prepared to engage in adult life, and recent policy efforts confirm this prioritization (Mishkind, 2014). Even so, as many as 75% of students lack the necessary academic preparation to enroll and succeed in credit-bearing postsecondary courses (ACT, Inc., 2012; Camara, 2013). This issue is further complicated by the multiple definitions, frameworks, and models of college and career readiness (CCR) that emphasize academic *and* nonacademic skills that are deemed important for employment, postsecondary education, or both (e.g., College and Career Readiness and Success Center, 2014; Conley, 2010; Farrington et al., 2012; Mishkind, 2014). Further, mounting evidence shows students with disabilities have poorer postschool outcomes than their peers without disabilities (Sanford et al., 2011), such as higher course failure and dropout rates (Doren, Murray, & Gau, 2014; National Center for Education Statistics, 2012, see Indicator 33) and fewer opportunities to receive an academically rigorous curriculum in high school (Gregg, 2007), and they self-report using fewer critical thinking skills (Lombardi, Kowitt, & Staples, 2015). Together, these findings demonstrate a persistent problem that *all* students do not receive adequate preparation for college and careers and that, for students with disabilities, this issue is even more pronounced. It is therefore crucial that high schools adequately prioritize CCR and ensure these opportunities are offered schoolwide to all students, with and without disabilities.

[1]University of Connecticut

Corresponding Author:
Allison Lombardi, PhD, University of Connecticut, 249 Glenbrook Rd, Unit 3064, Storrs, CT 06269-3064.
E-mail: allison.lombardi@uconn.edu

The purpose of this study was to empirically test an emerging framework of CCR for adolescents with and without disabilities (Morningstar, Lombardi, Fowler, & Test, 2017) with the objective to clarify a measurement model that could guide the implementation of a multitiered system of support (MTSS) for CCR. In this study, measures that map onto the CCR framework were selected and administered to adolescents with and without disabilities across 13 urban and suburban high schools. Following data collection, latent variable modeling was employed to clarify a measurement model of CCR. The results of this study inform the application of a MTSS framework for CCR by providing an option for schoolwide data collection and decision making.

A Framework of CCR

Typically, CCR is measured with high school achievement indicators, such as grade point average and college admissions exam scores. Yet, the predominant indicators of CCR do not sufficiently align with the knowledge and skills that first-year college students need in order to be successful in entry-level courses (R. S. Brown & Conley, 2007). Researchers agree there are multiple multifaceted skills that are more nonacademic by definition (College and Career Readiness and Success Center, 2014; Conley, 2010; Farrington et al., 2012), and some of these nonacademic skills are positively related to higher achievement (Duckworth, Peterson, Matthews, & Kelly, 2007; Lombardi et al., 2015; West et al., 2016). As such, it is crucial to include nonacademic skills in CCR definitions, models, and frameworks.

Adhering to the urgency around nonacademic skills of CCR and building on the work from researchers outside of special education, Morningstar, Lombardi, Fowler, and Test (2017) developed a six-domain organizing framework of CCR that emphasizes skills both academic and nonacademic skills: academic engagement, mind-sets, learning processes, critical thinking, interpersonal engagement, and transition competencies. This framework is meant to be applied in an inclusive, schoolwide manner, applicable to students with and without disabilities. However, although these researchers conducted a preliminary qualitative study to shape the six domains, the CCR framework has not yet been empirically validated.

CCR Within Multitiered Systems of Support

Lombardi and colleagues (2015) called for the application of MTSS as an implementation framework for CCR in high schools for all students, including those with disabilities. This approach has several advantages: (a) It encourages schoolwide access, which fosters an inclusive approach to ensure that all students, with and without disabilities, have access to CCR; (b) it promotes data-based decision making with regard to schoolwide practices and the level and intensity of small group and individual supports; and (c) fidelity of implementation—or the assurance that interventions are implemented the way that they were intended—is paramount.

All students do not receive adequate preparation for college and careers and…, for students with disabilities, this issue is even more pronounced.

To implement CCR via MTSS, there must be systematic data collection and ongoing analysis and progress monitoring of academic and nonacademic skills. Adequate data collection systems must be put in place and utilized; however, there are no systematic methods currently available for measuring CCR. This is partially due to many competing CCR frameworks and models (e.g., Conley, 2010; Farrington et al., 2012; Morningstar, Lombardi, Fowler, & Test, 2017). Morningstar, Lombardi, and Test (2017) acknowledged this gap and suggested the need for a data dashboard that high schools might use to select validated measures that are relevant to academic and nonacademic CCR skills. As such, a crucial step in the research to date is to empirically test one or more CCR frameworks to establish a validated CCR measurement model that will enable researchers and practitioners to measure the relevant skills.

In this study, we employed factor analytic methods to empirically test the six-domain framework (Morningstar, Lombardi, Fowler, & Test, 2017) as well as examine measurement invariance on the basis of disability. We used a bifactor model approach (Holzinger & Swineford, 1937) to determine whether a general factor (CCR) exists and, if so, the strength of the general factor and the degree to which specific factors explain variance controlling for the general factor. Specifically, we addressed the following research questions:

> *Research Question 1*: Can a measurement model of CCR be established?
> *Research Question 2*: Does the measurement model function similarly for adolescents with and without disabilities?
> *Research Question 3*: What are the differences in the CCR latent parameters for adolescents with and without disabilities?

Method

Participants

Participants were adolescents in Grades 9 to 12 at 13 high schools in a Midwestern state. Students with (n = 784) and without (n = 4,253) disabilities were included in the sample. “Disability” was defined as those students who had an individualized education program and received special education services. Of the students with disabilities, the majority fell into the categories of learning disability (44%) and other health impairment (36%). Compared with the national average (National Center for Education Statistics, 2014), this sample included a high percentage of African American students overall (41% in sample, 15.7% nationally) and within special education (49% in sample, 15.3% nationally). Table 1 shows detailed sample characteristics.

Procedures

Within the context of a larger study examining student outcomes in high schools, we recruited participating schools through existing relationships with technical assistance providers connected with the Office of Special Education Programs’ National Technical Assistance Center for Positive Behavioral Interventions and Supports. Interested administrators followed up with researchers to volunteer for the study and signed a data access agreement to release specified school data. We provided participating schools with a parental notification letter, which school administrators sent home at least 1 week prior to the survey administration. The notification letter provided parents with a web link to view the surveys online and the opportunity to view paper-based versions of the survey, which were made available in the front office of participating schools. Parents were given the opportunity to opt their student out of participation by signing and returning the notification letter. In addition, students were given the opportunity to opt out during the assent process on the day of administration. Students who chose not to participate or whose parents opted them out were allowed to work on other classroom activities during the time of administration. All study protocols were approved by the institutional review board for the protection of human subjects.

Students took the compilation of surveys using the online survey program Qualtrics on school-based computers. An administration window for data collection was determined between the researchers and school partners and spanned a 3-week period in the spring. School administrators determined time of day and class periods in which to administer the survey, and they were asked to ensure that all students were given the opportunity to take the survey. The survey was designed to take 30 to 50 min, and students were given one class period to respond. After administration, Qualtrics reports showed an estimated 25-min mean response time across participants. We provided each school with a web link that was unique to its school, which was put on

Table 1. Sample Characteristics.

Group	Sample 1 (n = 2,519)	Sample 2 (n = 2,520)	Responders (n = 5,039)	SPED (n = 784)	Non-SPED (n = 4,253)	Nonresponders (n = 5,696)
Gender						
Male	52	51	51	67	49	44
Female	48	49	49	33	51	56
Ethnicity						
Caucasian	27	28	27	25	28	18
African American	42	40	41	49	40	58
Hispanic/Latino	25	24	24	22	25	18
Asian	5	6	5	2	6	4
Other	1	2	2	1	2	2
Grade						
9	37	34	35	40	35	27
10	26	27	27	26	27	24
11	22	25	23	23	23	23
12	15	14	15	10	15	26
Free/reduced-price lunch						
Yes	61	61	61	69	60	69
No	39	39	39	31	40	31
Disability status						
Yes	15	16	16	100	0	28
No	85	84	84	0	100	72
Disability						
Learning disability	41	47	—	44	—	23
Other health impairment	39	34	—	36	—	33
Emotional disturbance	7	6	—	6	—	10
Autism spectrum disorder	4	5	—	5	—	5
Intellectual disability	6	5	—	5	—	25
Other	3	3	—	4	—	4

Note. Values are presented as percentages. The overall sample (N = 10,735) includes survey responders and nonresponders and contains 2 more observations due to omission of special education (SPED) status, which were evenly split between Samples 1 and 2.

school computers on the administration dates. We provided a script to teachers to introduce the study and inform students of their right to choose not to participate. Students gave assent to participate by answering yes to Question 1, “I would like to take these surveys.” During administration, students were instructed to ask for help if they had any questions or wished to have one or more items read aloud by an adult. Upon this request, we instructed staff members to read aloud the specified item, then step away from the computer and allow the student to respond independently. School personnel were told to provide accommodations (e.g., extra time, translated materials) for any students or families for whom these accommodations were routinely provided. All items were optional, and students were able to discontinue the survey at any point without penalty. During the survey administration window, research team members monitored the number of responses and coordinated follow-up survey administrations with school contacts if needed (e.g., the number of respondents was considerably low after the first 2 weeks). Students were asked to enter their school-issued identification numbers at the beginning of the survey. After survey administration, we matched student responses with extant school data using the identification numbers. Across schools, 7% of participants did not give assent. The mean response rate across the 13 schools was 50% (SD = 18%), with a minimum rate of 9% and a maximum of 75%. To better understand the nature of the response rate and sample representation, Table 1 shows sample characteristics of survey responders and nonresponders.

Measures

A compilation of previously validated measures was selected per the similarities between the operational definitions of the measures and Morningstar and colleagues’ six domains of CCR (Morningstar, Lombardi, Fowler, & Test, 2017). In what follows, each measure is described with the specified CCR domain (or domains) onto which it maps.

Vocational Skills Self-Efficacy. The Vocational Skills Self-Efficacy (VSSE; McWhirter, Raseed, & Crothers, 2000) is a 29-item self-report scale in which students select responses based on their levels of confidence in career preparation tasks (e.g., writing a resume, preparing for a job interview, searching for jobs). The response scale ranges from 1 (*no confidence at all*) to 5 (*complete confidence*). McWhirter et al. (2000) report a Cronbach’s alpha of .97, estimated from a sample of high school sophomores. In the current sample, a Cronbach’s alpha of .977 resulted, whereas for those with and without disabilities, this estimate was .978 and .977, respectively. The VSSE was selected because it maps onto the transition competencies domain.

Key Learning Strategies and Techniques. The Key Learning Strategies and Techniques (KLST; Lombardi, Seburn, & Conley, 2011) subscale is taken from CampusReady, an instrument intended for high schools to evaluate nonacademic skills related to CCR. Specifically, KLST measures study skills and self-monitoring, and it maps onto the academic engagement, learning processes, and interpersonal engagement domains. The response scale ranges from 1 (*not at all like me*) to 5 (*very much like me*). The total number of items on the KLST is 29. Sample items include “I refer to the syllabus or class website to prepare for and complete course assignments,” and “I use unstructured time during the school day to complete assignments.” Lombardi et al (2011) reported a four-factor structure and an overall Cronbach’s alpha of .95, .84 for Goal-Driven Behavior items, .90 for Persistence items, .89 for Self-Monitoring items, and .83 for Study Skills items. In the current sample, Cronbach’s alpha overall was .958; this value was .961 for those with disabilities and .957 for those without disabilities. With respect to subscale internal consistency estimates, all were >.80.

Key Cognitive Strategies. The Key Cognitive Strategies (KCS; Lombardi, Conley, Seburn, & Downs, 2013) subscale is taken from

CampusReady, an instrument intended for high schools to evaluate nonacademic skills related to CCR. Specifically, KCS maps onto the critical thinking domain. The response scale ranges from 1 (*not at all like me*) to 5 (*very much like me*). The total number of items on the KCS is 57. Sample items include "I can accept critiques and challenges to assertions I've made," and "I check for spelling and grammar errors before turning in work." Lombardi et al. (2013) reported a five-factor structure and a Cronbach's alpha of .96 overall, .88 for Problem Formulation items, .88 for Research items, .88 for Interpretation items, .90 for Communication items, and .93 for Precision Accuracy items. In the current sample, the Cronbach's alpha estimate was .987 overall, and this estimate remained true across disability groups. All subscale internal consistency estimates were >.9.

Grit Scale. The Grit Scale (12-item version; Duckworth et al., 2007) measures two constructs—perseverance and consistency of interests—and maps onto the mind-set domain. The response scale ranges from 1 (*not at all like me*) to 5 (*very much like me*). Duckworth et al. (2007) reported internal consistency estimates for the composite across six studies, and these estimates ranged from .77 to .85. In the current sample, the Cronbach's alpha estimate was .829 for the Perseverance subscale, which was the only subscale retained in the analysis, a decision based on previous evidence of problematic psychometric properties of the Consistency of Interests subscale (Lombardi, Rifenbark, & Freeman, 2017). By disability status, these estimates for those with and without disabilities were .826 and .830, respectively.

Demographic characteristics. We gathered students' grade level, gender, race, free and reduced-price lunch status, and disability status using school extant data records.

Data Analysis

An empirical approach for testing the CCR framework requires the development of a measurement model whereby student item responses are analyzed via factor analysis. Factor-analytic models entertain the notion that survey items collectively measure some unobservable phenomenon or latent construct (i.e., items are manifestations of CCR) with error. Therefore, the observed variance of each manifest variable is partitioned into variance that supports the latent construct (e.g., common or shared variance) and does not support the latent construct (e.g., unique variance, or measurement error). Such partitioning results from a multivariate regression where manifest variables are simultaneously regressed onto an exogenous latent variable, producing structural and latent parameters.

Structural parameters. Measurement models result in structural parameters that pertain to either the variance or the mean structure. With respect to variances, two parameters result: factor loadings (Λ, lambda) and manifest residuals (Θ, theta). Lambda estimates represent variance explained by the latent variable (e.g., variance common to CCR) that determines the importance of a given manifest variable to its construct (e.g., CCR) via regression weights. Variance left unexplained, represented by theta, is unique to each manifest variable. When the mean structure is included, an intercept results for each manifest variable, represented by tau (τ), interpreted as the expected value of the manifest variable when the latent variable equals zero.

Latent parameters. By controlling for the resulting structural parameter estimates, inferences can be made with respect to the distribution of the measured construct (e.g., CCR) in the population sampled. Specifically, anchoring and dispersion parameters result. The latent mean (α, alpha) is interpreted as the expected value of the construct that anchors the distribution. Latent variance (ψ, psi) represents the degree to which the population collectively deviates from the expected value.

Model identification. An important aspect of factor analysis is scaling and identification. Depending on the method of identification employed, scaling of the parameter estimates is done with respect to either the latent variable or a manifest variable. For all models, we used the fixed factor method of identification (T. A. Brown, 2015), whereby the latent variance is fixed to 1.0 and the mean to 0.0; thus, scaling was done with respect to the latent variable, allowing free estimation of all structural parameters. In turn, the intercepts (tau) are interpreted as the estimated value on the manifest variable for someone at mean-level CCR, whereas factor loadings (lambda) can be used to calculate indicator reliability (described later).

Model fit. For all models, we consulted two indices from the incremental perspective for model fit: the comparative fit index (CFI; Bentler, 1990) and the Tucker-Lewis index (TLI; Tucker & Lewis, 1973), where a value ≥0.95 reflects a close-fitting model. In addition, we consulted two indices from the absolute perspective: the root mean squared error of approximation (RMSEA; Steiger & Lind, 1980) and the standardized root mean square residual (SRMR; Bentler, 1995) where values of 0.06 and 0.08, respectively, indicate good fit to the data (Hu & Bentler, 1999).

Model-building process. Using the aforementioned measures, we a priori selected items that mapped onto the six domains of CCR, forming single hypothesized constructs. In the context from which bifactor models are more prevalent, item response theory, a bifactor model can be fit to establish the degree to which the construct is unidimensional (Reise, 2012; Reise, Moore, & Haviland, 2010). In this study, we pooled items from various existing measures to generate the CCR constructs and felt it necessary to examine the psychometric characteristics of each construct prior to estimating the bifactor measurement model. Therefore, using a confirmatory lens, we took an exploratory approach to model building to investigate the six domains of CCR. Accordingly, we generated subsamples to prevent the occurrence of Type II errors (e.g., that our findings were due to chance).

Split sample. We split the full sample into two by taking as many draws from the uniform distribution as there were participants. A distinguishing characteristic of uniform distribution is that its density mass is equivalent across its range. Therefore, each observation was equally likely to fall into either subsample. To ensure an equal split, we used the observed median value of this random variable. Table 1 shows characteristics by Samples 1 and 2. We carried out the full model-building process with Sample 1 and used Sample 2 to confirm the final bifactor measurement model.

Single-factor confirmatory factor analyses. Using Sample 1, we fit single-factor confirmatory factor analysis (CFA) models to determine whether the hypothesized domains were adequate. In the process of fitting single-factor CFAs, we retained items with standardized lambda estimates ≥0.400, which corresponds to a minimum R^2 of 0.16 (e.g., indicator reliability calculated via tracing rules) for each item (Kline, 2011; Shogren & Garnier Villarreal, 2015).

Bifactor versus second-order model. When factors are highly correlated with one another, a second-order model can be fit, as evidenced by Lombardi and colleagues (Lombardi et al., 2011; Lombardi et al., 2013) when validating the KLST and the KCS. Such a representation has the following consequences: First, the effect that an item has on the second-order factor (e.g., CCR) is transmitted through its first-order factor (e.g., Transition Knowledge); therefore no inferences can be made with respect to its direct effect on the second-order factor. Second, first-order factors are endogenous, and their variances are represented by disturbance parameters (i.e., variance that is not explained by the second-order factor); therefore, determining the effect of a first-order disturbance on outcomes (e.g., controlling for CCR, the effect of Transition Knowledge) in a structural equation model is intractable (Reise, 2012). It was for these reasons that we chose to model CCR using the traditional bifactor model (Holzinger & Swinefold, 1937).

Traditional bifactor model. This model requires each item to load onto a specific factor (e.g., no cross-loadings) as well as the general factor (e.g., CCR). Specific factors are orthogonal to one another (e.g., correlations fixed to 0.0), allowing construct reliability (e.g., omega hierarchical) to be appropriately estimated while easing the interpretation of the specific factors. Finally, the indicator–(specific) factor ratio should be similar. Therefore, after estimating the single-factor CFA models, we estimated the bifactor model at the item level, rather than using parcels to avoid masking multidimensionality (Reise, Scheines, Widaman, & Haviland, 2013). Items that did not have a significant loading on its specific factor were removed from the model to preserve the 1:1 ratio of loadings on the specific and general factor.

Construct reliability. Due to our use of latent variable modeling with a bifactor model, coefficient alpha (Cronbach, 1951) was no longer appropriate, as it does not take into account structural or latent parameters and it consistently overestimates reliability (Zinbarg, Yovel, Revelle, & McDonald, 2006). Instead, we report coefficient omega hierarchical (ω_h), as its calculation takes into account structural parameters estimates to determine reliability.

After arriving at the final bifactor model, we specified and estimated the factor structure using data from Sample 2. To determine whether the factor structure held, we compared model fit indices, chi-square estimates, item effects on each of their constructs, and construct reliability estimates across samples. We determined the utility of the final factor-analytic model based on model fit and construct reliability estimates of the resulting factors.

Invariance. To determine the functionality of the final bifactor measurement model across disability groups, we used the multiple-group CFA approach (Jöreskog, 1971; Sörbom, 1974). In this process, we tested measurement invariance and latent parameter invariance.

Measurement invariance. First, confirmation that the pattern of free and fixed parameters are the same across groups is required, referred to as *form invariance*. Second, it is necessary to determine whether the lambda matrix (Λ) can be constrained to be the same across groups, known as *metric invariance* (i.e., items equally discriminate between those who are high and low on CCR, regardless of disability group membership). Third, it is necessary to ascertain whether or not the tau (τ) vector can be constrained to be the same, referred to as *scalar invariance*. Constraints on the structural parameters were deemed tenable if the change in CFI (ΔCFI) was $\leq$0.01 (Cheung & Rensvold, 2002). Therefore, if Λ is constrained to be the same across disability groups without significant degradation of model fit (ΔCFI $\leq$ 0.01), then metric invariance has been met.

Latent parameter invariance. To test for group differences in latent parameters, its respective structural constraint must be at least partially invariant across disability groups. Similar to measurement invariance, latent parameters are constrained to be the same across disability groups. To assess whether latent parameters differ across groups, we used chi-square difference tests, as these models are nested and such constraints will not have a large impact on global fit indices (Millsap, 2012).

Due to Likert scales being inherently ordinal, we set out to determine if severe departures from normality occurred in our data and to inform us on the appropriateness of a normal theory approach to estimation. To this end, we conducted a descriptive analysis on all survey items and assessed departures from normality via kurtosis and skew. These analyses were carried out with the *psych* package (Revelle, 2016) in R version 3.3.0 (R Core Team, 2016). All factor-analytic models were estimated with Mplus version 7.31 (Muthén & Muthén, 2015). We utilized a user-defined

R function to carry out all chi-square difference tests.

Results

After subjecting the manifest variables to a descriptive analysis, we observed slight departures from normality, including one item with skew and four items with excess kurtosis. As such, all models were estimated with the robust maximum likelihood (Rhemtulla, Brosseau-Liard, & Savalei, 2012), and we tested for latent parameter invariance via chi-square difference tests (Satorra & Bentler, 2001).

Research Question 1: Explication of CCR Constructs

We subjected all items to construct-specific CFA based on the aforementioned six-domain framework. We used Sample 1 (n = 2,519) for both the single-factor CFAs and the bifactor model. Upon investigating learning processes and critical learning in a simultaneous fashion, we decided to collapse these original domains to form the specific factor Critical Learning Processes, as the resulting latent correlation between these domains was exceedingly large (r = .846) and modification indices resulted in nonnegligible cross-loadings with expected parameter change (T. A. Brown, 2015) estimates approaching 0.400. In a similar vein, the domains interpersonal engagement and mind-set were collapsed to form the specific factor Mind-Set, as the latent correlation between the two was large (r = .818), whereas modification indices suggested four cross-loadings with expected parameter change estimates ≥0.14. Therefore, the specific factors were identified as Academic Engagement, Critical Learning Processes, Mind-Set, and Transition Knowledge.

Academic Engagement CFA. A total of 28 items spanning certain KCS and KLST sub-subscales were utilized (Precision Accuracy, Self-Monitoring, and Persistence). With respect to model fit, the RMSEA was 0.067, 90% confidence interval (90% CI) [0.065, 0.068]; the SRMR was 0.053; the CFI was 0.873; and the TLI was 0.863. All standardized loading estimates were at least 0.560 (KLST 9: "I refer to the syllabus or class website to prepare for and complete course assignments") and as large as 0.826 (KCS 46: "It is important to me to be precise in my schoolwork"); all of which were significantly different from zero.

Critical Learning Processes CFA. We used a total of 27 items across certain KCS and KLST sub-subscales in this model (Study Skills, Interpretation, and Research). Model fit indices showed that the RMSEA was 0.058, 90% CI [0.056, 0.060], the SRMR was 0.042, the CFI was 0.906, and the TLI was 0.899. All parameter estimates were statistically different from zero. The least discriminating item was KLST18, "I journal or blog about what I learn"; its estimate was 0.398. The most discriminating item was KCS18, "When faced with a problem to solve, I can identify the information sources I need to help me find a solution"; its estimate was 0.820.

Mind-Set CFA. We utilized a total of 35 items spanning certain subscales from the KCS and KLST subscales and the Grit Scale (Problem Formulation, Perseverance, Goal Driven Behaviors, and Communication) in this model. The RMSEA was estimated to be 0.070, 90% CI [0.069, 0.072]; the SRMR was 0.056; and the CFI and TLI were 0.831 and 0.820, respectively. The most discriminating item was KCS38, "I can construct well-reasoned arguments to explain issues or answer questions"; its estimate was 0.811. The least discriminating item was KLST4, "I participate in study groups outside of class"; its estimate was 0.427.

Transition Knowledge CFA. We evaluated all 29 items from the VSSE scale in this unidimensional construct. With respect to fit, the RMSEA was estimated to be 0.073, 90% CI [0.071, 0.075]; the SRMR was 0.043; and the CFI and TLI were 0.857 and 0.846, respectively. The most discriminating item was VSSE 17, "Find available options in a given decision-making situation"; its estimate was

0.838. The least discriminating item was VSSE24, “Make a monthly budget for myself that would include all bills, payment of debts, spending money for food, clothing, and entertainment, etc.”; its estimate was 0.701. All other structural parameter estimates were found to be statistically different from zero.

Bifactor model. Given the results of the single-factor CFAs, we modeled all 119 items using a bifactor approach in which we specified four specific factors and a general factor (CCR). In terms of fit, the RMSEA was estimated to be 0.036, 90% CI [0.036, 0.037], whereas the SRMR was 0.042 and the CFI and TLI resulted in estimates of 0.861 and 0.856, respectively. All factor loading estimates on the general factor were statistically different from zero. Several specific factor loadings were found to be nonsignificant and were removed. For the Academic Engagement factor, we removed items KLST 13 (p = .235), 23 (p = .304), and 24 (p = .490). For the Critical Learning Processes factor, we removed KLST 17 (p = .224) and 19 (p = .632); KCS 34 (p = .243); and, finally, Items 4 (p = .777) and 12 (p = .804) from the Grit Scale. All Transition Knowledge–specific loadings were significantly different from zero and ranged from 0.466 (VSSE 24) to 0.604 (VSSE 9).

After removal of these eight offending items, model fit improved. Specifically, the CFI and TLI were estimated to be 0.866 and 0.861, respectively, whereas the RMSEA and SRMR were estimated to be 0.037, 90% CI [0.036, 0.037], and 0.041, respectively. We found all the general factor loading estimates to be statistically different from zero, and the specific factor loading estimates were all statistically different from zero for the Transition Knowledge and Critical Learning Processes factors. In the Academic Engagement factor, Item 10 (p = .246) from the KLST was found to be nonsignificant. In the Mind-Set factor, a total of 17 items were found to be nonsignificant; however, five items were removed due to large p values, which were KCS 1 (p = .903), 2 (p = .959), 3 (p = .940), 4 (p = .727), and 13 (p = .639).

In sum, we removed 14 items and proceeded to fit the bifactor model with the remaining 105 items. With respect to fit, an RMSEA estimate of 0.037, 90% CI [0.036, 0.037], an SRMR of 0.041, a CFI of 0.872, and a TLI of 0.866 resulted. A pattern matrix of the final model with item text as well as annotated Mplus code for fitting the bifactor model are available as online supplemental files.

College and Career Readiness. In the general factor (CCR), all factor loadings were significantly different from zero, and estimates ranged from 0.448 (KLST18, “I journal or blog about what I learn”) to 0.784 (KCS18, “When faced with a problem to solve, I can identify the information sources I need to help me find a solution”). The mean loading estimate was 0.658. From this factor solution, ω_h was estimated to be 0.949.

Academic Engagement. The items that regressed onto this specific factor were all statistically different from zero. Factor loadings for items from the Precision Accuracy subscale were positive and ranged from 0.150 (KCS 52) to 0.391 (KCS 45), whereas factor loading estimates stemming from the Self-Monitoring and Persistence subscales were negative, ranging from −0.177 (KLST 26) to −0.065 (KLST 22). For this specific factor, ω_h was 0.03; thus, no evidence of multidimensionality was observed. The average factor loading on the general factor was 0.703.

Critical Learning Processes. Controlling for the general factor, mean factor loading estimates were 0.173 for Study Skills, −0.239 for Interpretation, and 0.266 for Research. The mean general factor loading estimate was 0.716. For this specific factor, ω_h was 0.035; therefore, no evidence of multidimensionality was observed.

Mind-Set. Controlling for the general factor, mean factor loading estimates were −0.100 for Problem Formulation, 0.046 for Perseverance, 0.138 for Goal Driven Behaviors, and −0.338 for Communication. The average factor loading estimate on the general factor was 0.679.

For this specific factor, ω_h was 0.018; therefore, no evidence of multidimensionality was observed.

Transition Knowledge. Substantial evidence of multidimensionality was observed for this specific factor, as ω_h was 0.492. The mean general factor loading estimate was 0.553, whereas the mean specific factor loading estimate was 0.551. Therefore, these independent factor loadings exhibit near tau equivalence between these orthogonal constructs.

Confirmation of bifactor model. We then specified the same bifactor model using response data from Sample 2 (n = 2,520). All four specific factors remained consistent with the model estimated with Sample 1. This model was found to replicate the RMSEA estimate, 0.037; however, the SRMR was slightly larger at 0.044 (Sample 1 model was 0.041). Both the CFI and TLI estimates were slightly smaller (0.872 vs. 0.867 and 0.866 vs. 0.861 in Samples 1 and 2, respectively). Both models contain the same number of degrees of freedom, 5,250; the χ^2 value for Sample 1 was 22,748.734, whereas it was 23,401.400 for Sample 2. Therefore, the specified bifactor model was found to fit better with response data from Sample 1; however, these differences are negligible. The ω_h estimates were similar across samples, as Sample 1 and 2 estimates were as follows: 0.949 versus 0.937 for CCR, 0.03 for Academic Engagement (in both samples), 0.035 versus 0.054 for Critical Learning Processes, 0.018 versus 0.03 for Mind-Set, and 0.492 versus 0.463 for Transition Knowledge.

Research Question 2: Invariance of Bifactor Measurement Model

Due to the similarities of the bifactor model across the samples, we used all response data (n = 5,039) to estimate the confirmed bifactor structure. We found this model to have similar fit to the data: RMSEA was estimated to be 0.036, 90% CI [0.036, 0.037]; SRMR was 0.044; and 0.871 and 0.866 were observed for CFI and TLI, respectively. On 5,250 degrees of freedom, χ^2 was estimated to be 39,639.917. To determine whether there were large departures with respect to model fit by group, we estimated independent models using response data from those without disabilities (n = 4,179) and those with disabilities (n = 766). No significant departures were observed; therefore, we proceeded with the investigation of measurement invariance. The fit of these models can be found in Table 2.

We specified the bifactor structure such that the pattern of fixed and free parameters was the same across groups. The fit of this form-invariant model resulted in an RMSEA of 0.038, 90% CI [0.037, 0.038], an SRMR of 0.042, a TLI of 0.861, and a CFI of 0.867, which was used to compare subsequent models. Constraints on the factor loadings and manifest intercepts were tenable, as change in CFI estimates were 0.001 and 0.002 from their respective comparison model. Attaining metric and scalar invariance affords the opportunity for differences in latent parameters across groups to be tested. Table 2 shows detailed results of invariance testing across students with and without disabilities.

Research Question 3: Test of Bifactor Latent Parameters

Using the metric invariant bifactor model as the comparison, we constrained factor variances to be the same across groups. By doing so, a total of 5 degrees of freedom were gained and resulted in a change in χ^2 of 0.304 and was tenable as 0.304 is smaller than its critical value, 11.07. Next, we constrained the latent means to be the same across groups; therefore, 5 degrees of freedom were gained from the scalar invariant model. The resulting change in χ^2 was 219.912, which is significant on 5 degrees of freedom; therefore, this constraint was not tenable. We further decomposed the latent means and found that the smallest difference in latent means (CCR) could not be constrained to be the same (χ^2 = 9.642, df = 1, p = .002); therefore, all latent means differed across groups.

Table 2. Invariance Testing Across Groups.

Step	Model	χ^2	*df*	Scaling	CFI	TLI	RMSEA	SRMR	ΔCFI / χ^2 (*df*), *p*	Tenable?
Single-group models										
	Full sample	39,639.917	5,250	1.442	0.871	0.866	0.036	0.041	—	—
	Disability group	10,526.463	5,250	1.3285	0.873	0.868	0.036	0.044	—	—
	Nondisability group	36,199.812	5,250	1.4271	0.866	0.860	0.038	0.042	—	—
Measurement invariance										
0	Independent null	289,298.917	10,920							
1	Form	47,644.316	10,500	1.3778	0.867	0.861	0.038	0.042		
2	Metric	48,008.486	10,705	1.374	0.866	0.863	0.038	0.043	−0.001	Yes
3	Scalar	48,589.604	10,805	1.3706	0.864	0.863	0.038	0.044	−0.002	Yes
Tests of latent parameters										
4.0.0	Variances	48,005.343	10,710	1.3741	0.866	0.863	0.038	0.044	0.304 (5), p = .998	Yes
4.1.0	Means	48,747.251	10,810	1.3704	0.864	0.862	0.038	0.044	219.912 (5), p < .001	No
4.1.1	General	48,599.246	10,806	1.3706	0.864	0.863	0.038	0.044	9.642 (1), p = .002	No
4.1.2	General + TK	48,621.004	10,807	1.3705	0.864	0.863	0.038	0.044	45.980 (2), p < .001	No
Final model										
5.0.0	Scalar + latent variances	48,585.964	10,810	1.3708	0.864	0.863	0.038	0.045		

Note. $\Delta\chi^2$ test per Satorra and Bentler (2001). CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean squared error of approximation; SRMR = standardized root mean square residual; TK = transition knowledge.

As such, the most parsimonious model is the scalar invariant model with added constraints on the latent variances. We interpret the difference in latent means as standardized differences due to using the fixed factor method of identification. Small effects were observed for CCR (d = 0.142), Critical Learning Processes (d = 0.374), and Transition Knowledge (d = 0.250); all of which favor those without disabilities. However, medium effects were found for Academic Engagement (d = −0.357) and Mind-Set (d = −0.451) favoring those with disabilities. Based on this final model, ω_h was estimated and was similar for those with and without disabilities: 0.03 for Academic Engagement, 0.02 for Mind-Set, and 0.04 for Critical Learning Processes. However, differences were found for CCR, 0.862 versus 0.956, and Transition Knowledge, 0.414 versus 0.492.

Discussion

In this study, we empirically tested an emerging framework to establish a measurement model of CCR for adolescents with and without disabilities. Findings indicate that the original six-domain framework was most accurately modeled as four specific factors: Academic Engagement, Critical Learning Processes, Mind-Set, and Transition Knowledge. Some of the characteristics of two domains from the initial Morningstar framework (Morningstar, Lombardi, Fowler, & Test, 2017) blended in the process of constructing specific factors; namely, interpersonal engagement and mind-set came together as Mind-Set, and critical thinking and learning processes came together as Critical Learning Processes. Academic Engagement and Transition Knowledge were consistent with the definitions in the Morningstar framework. This was in part empirically driven, due to strong latent correlations between the respective domains; however, this decision was also rendered as an effort to conform to the traditional bifactor model. Specifically, we circumvented the possibility of nonnegligible cross-loadings while attempting to ensure an equivalent indicator–specific factor ratio across all specific factors.

We then utilized a bifactor model where one general CCR factor was specified with four specific factors based on the results of the single-factor CFAs. After removal of problematic items, the model showed good fit to the data and was subsequently confirmed with an independent set of responses. Our results show the explication of CCR as a general factor, with only Transition Knowledge exhibiting evidence of multidimensionality due to a substantial amount of variance that was explained by this specific factor while controlling for CCR ($\omega_{h.S1}$ = 0.492 and $\omega_{h.S2}$ = 0.463). Therefore, this bifactor measurement model proves useful for informing intervention and support needs in the four specific factors, yet only the factor scores for general CCR and Transition Knowledge should be used to drive data-based decisions.

Subsequently, we found that the CCR measurement model is equivalent for students with and without disabilities and that the scaling of the distribution of these scores is identical across groups. However, these populations differ with respect to the point at which their distribution of nonacademic CCR skills is anchored (e.g., latent means were noninvariant across groups). CCR scores for adolescents without disabilities were 0.142 *SD* above those with disabilities. The effect size for this difference was small, indicating that after controlling for CCR construct-irrelevant variance, those without disabilities are more prepared for college and careers than those with disabilities. We found a larger discrepancy between groups in Transition Knowledge, where students without disabilities are on average 0.250 *SD* above those with disabilities. This effect was small, illustrating that after controlling for individual CCR scores, students without disabilities are on average more confident in their abilities to prepare for their future careers. Given that students with disabilities are legally required under the Individuals with Disabilities Education Act (2006; § 1414 d.1.A.i.VIII.aa) to receive transition services in employment and postsecondary education, these results suggest that these services might not be well aligned with nonacademic skills of CCR.

Students without disabilities are on average more confident in their abilities to prepare for their future careers.

These findings are crucial to future schoolwide efforts to improve nonacademic CCR skills and further provide empirical support that students with disabilities should be included in such efforts. Although previous researchers have made connections between MTSS and high school implementation (e.g., Bohanon, Gilman, Parker, Arnell, & Sortino, 2016; Freeman et al., 2016; Shogren, Wehmeyer, & Lane, 2016), this study specifically addresses CCR within this context. The implementation of MTSS relies on systematic data collection and analysis that is coordinated and utilized by school-based teams. The multidimensional nature of nonacademic CCR has led researchers to propose models with academic and nonacademic skills, yet little empirical evidence has supported these conceptual models. Our findings help to empirically disentangle CCR constructs, particularly with regard to general nonacademic CCR and Transition Knowledge. These findings should be considered in future schoolwide efforts to measure nonacademic skills, as well as provide direction and clarification with regard to a data dashboard, a recently proposed concept for high school implementation of CCR via MTSS (Morningstar, Lombardi, & Test, 2017). Further, given the discrepancies found between students with and without disabilities on general CCR and Transition Knowledge, an essential next step is to examine the extent to which secondary special education and transition services align with broader CCR definitions, models, and frameworks.

Limitations

In interpreting the findings of this study, several important limitations should be considered. First, with regard to model fit, the RMSEA and SRMR both pointed toward a close-fitting model, whereas the CFI and TLI did not (<0.95) based on the cutoffs provided by Hu and Bentler (1998, 1999). However, Hu and Bentler's recommended cutoffs were based on simulations that investigated Type I and II error rates of common fit indices in the context of a three-factor model informed by 15 observed variables, representing an underparameterized model. Hu and Bentler (1998) question the generalizability of their suggested cutoffs to other modeling contexts; namely, overparmeterized models (e.g., bifactor models). Therefore, without simulation, it is difficult to strictly apply their cutoffs to our context, as it is reasonable to expect fit indices to fluctuate due to sampling error (Pornprasertmanit, 2014). Furthermore, the model fit pattern that emerged from this study complements the work of Ding, Velicer, and Harlow (1995), who concluded that as the ratio of indicators-to-factors increases, estimates for both CFI and TLI will decrease. This example illustrates why parceling (Little, Cunningham, Shahar, & Widaman, 2002) results in better CFI and TLI estimates (e.g., number of parameters to be estimated is reduced); however, parceling was not appropriate in our context. In light of our indicators-factors ratio and the use of a bifactor model, we did not discount the fit of our hypothesized models, and our findings show that the fit indices were stable across different samples (e.g., Samples 1 and 2; disability and nondisability groups).

Second, we intentionally selected measures that mapped onto the six CCR domains, yet some aspects may not have been fully covered. In a similar vein, all measures were self-report and were not supplemented with teacher/parent perceptions or observation measures. Third, the sample characteristics and sampling procedures posed some limitations. School personnel conducted schoolwide data collection, and although we requested and schools reported that all students were given an opportunity to take the survey, the mean response rate across schools was 50%. We are not able to determine precisely why there were nonresponders (although it is confirmed that 7% did not assent). We might surmise the nature of schoolwide data collection presented some challenges that resulted in the inability to ensure that all students had access to the survey (e.g., tracking absences and/or who has and has not taken the survey over multiple weeks among multiple staff). Furthermore, the sample characteristics of the nonresponders show notable differences from the responders in certain demographic categories (see Table 1); namely, students with disabilities who responded were overrepresented in the LD category and underrepresented in the ID category. This discrepancy suggests a potential flaw in the data collection approach; in other words, although they were told to administer the survey schoolwide to all students, some school personnel may have assumed that certain students with disabilities should not be included in a schoolwide CCR survey. In addition, our sample was not entirely reflective of national trends in race and disability category; particularly, it included a higher proportional representation of African Americans, within and outside of special education, who were represented at 41% and 49%, respectively, as compared with national averages of 15.7% and 15.3%. There were disability categories that matched national averages (e.g., learning disability); yet, for other categories, representation in the sample was far below national trends (e.g., emotional disturbance, autism spectrum disorder). As such, despite the large sample size overall of students with (n = 784) and without (n = 4,253) disabilities, the findings may not be generalizable on a national scale.

Implications for Research

The results of this study provide empirical support and an initial approach to measuring overall CCR and Transition Knowledge factors. Future research using samples that match the national trends is needed to replicate and expand these findings and to continue to refine and test the measurement model. Further factor-analytic testing will be a crucial next step to confirm the CCR measurement model. A vital next step is to return to the theoretical six domains and determine whether the combination of measures selected in this current study adequately cover all six domains. Potentially, some domains and subdomains may have been better represented over others. In addition, confirming invariance across other student groups (e.g., race, socioeconomic status, specific disability categories) and the measures' sensitivity to change is an important next step. As well, participating schools should be provided with an efficient way to administer schoolwide as well as analyze and interpret survey results, as running advanced statistical models could be beyond the capacity of many school districts. Finally, future research efforts should focus on the link between survey responses and distal outcomes, such as college completion and employment.

Implications for Practice

Implementation of CCR via MTSS requires systematic collection and evaluation of data. Previously, schools have not had empirically supported options for measuring an overall, nonacademic CCR construct. The combination of measures used in this study could provide schools with a means for evaluating student perceptions of their CCR in an efficient and feasible way, and this information could be helpful to determine both the efficacy of current CCR practices and the need for other schoolwide practices, as well as to identify subgroups or individual students who may need more intensive support. Overall, this multitiered approach may help to better align secondary special education and transition services with schoolwide CCR efforts, and it has the potential to ensure that students with disabilities have access to the same CCR opportunities as their peers without disabilities.

Supplemental Material

The supplemental material is available in the online version of the article.

References

ACT, Inc. (2012). *The condition of college and career readiness*. Retrieved from www.act.org/readiness/2012

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238. doi:10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.

Bohanon, H., Gilman, C., Parker, B., Arnell, C., & Sortino, G. (2016). Using school improvement and implementation science to integrate multi-tiered systems of support in secondary schools. *Australasian Journal of Special Education*, 1–18. doi:10.1017/jse.2016.8

Brown, R. S., & Conley, D. T. (2007). Comparing state high school assessments to standards for success in entry-level university courses. *Educational Assessment*, *12*, 137–160. doi:10.1080/10627190701232811

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Camara, W. J. (2013). *College persistence, graduation, and remediation* (Research Notes No. RN-19). New York, NY: College Entrance Examination Board.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

College and Career Readiness and Success Center. (2014). *The college and career readiness and success organizer*. Washington, DC: American Institutes for Research. Retrieved from www.ccrscenter.org/ccrs-landscape/ccrs-organizer

Conley, D. T. (2010). *College and career ready: Helping all students succeed beyond high school*. San Francisco, CA: Jossey-Bass.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Ding, L., Velicer, W. F., & Harlow, L. L. (1995). Effects of estimation methods, number of indicators per factor, and improper solutions on structural equation modeling fit indices. *Structural Equation Modeling: A Multidisciplinary Journal*, *2*, 119–143. doi:10.1080/10705519509540000

Doren, B., Murray, C., & Gau, J (2014). Salient predictors of school dropout among secondary students with learning disabilities. *Learning Disabilities Research & Practice*, *29*, 150–159. doi:10.1111/ldrp.12044

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*, 1087–1101. doi:10.1037/0022–3514.92.6.1087

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance. A critical literature review*. Chicago, IL: University of Chicago Consortium on Chicago School Research.

Freeman, J., Simonsen, B., McCoach, D. B., Sugai, G., Lombardi, A., & Horner, R. (2016). Relationship between school-wide positive behavior interventions and supports and academic, attendance, and behavior outcomes in high schools. *Journal of Positive Behavior Interventions*, *18*, 41–51. doi:10.1177/1098300715580992

Gregg, N. (2007). Underserved and unprepared: Postsecondary learning disabilities. *Learning Disabilities Research & Practice*, *22*, 219–228. doi:10.1111/j.1540-5826.2007.00250.x

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*, 41–54. doi:10.1007/BF02287965

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. doi:10.1037/1082-989X.3.4.424

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. doi:10.1080/10705519909540118

Individuals with Disabilities Education Act, 20 U.S.C. §§ 1400 *et seq* (2006 & Supp. V. 2011).

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*, 409–426. doi:10.1007/bf02291366

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*, 151–173. doi:10.1207/S15328007SEM0902_1

Lombardi, A. R., Conley, D. T., Seburn, M., & Downs, A. M. (2013). College and career readiness assessment: Validation of the key cognitive strategies framework. *Assessment for Effective Intervention*, *38*, 163–171. doi:10.1177/1534508412448668

Lombardi, A. R., Kowitt, J., & Staples, F. (2015). Correlates of college readiness and critical thinking for secondary students with and without disabilities. *Career Development and Transition for Exceptional Individuals*, *38*, 142–151. doi:10.1177/2165143414534888

Lombardi, A. R., Rifenbark, G. G., & Freeman, J. (2017). *Exploring grit in adolescents with and without disabilities*. Manuscript submitted for publication.

Lombardi, A. R., Seburn, M., & Conley, D. T. (2011). Development and initial validation of a measure of academic behaviors associated with college and career readiness. *Journal of Career Assessment*, *19*, 375–391. doi:10.1177/1069072711409345

McWhirter, E. H., Rasheed, S., & Crothers, M. (2000). The effects of high school career education on social-cognitive variables. *Journal of Counseling Psychology*, *47*, 330–341. doi:10.1037/0022-0167.47.3.330

Millsap, R. E. (2012). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

Mishkind, A. (2014). *Definitions of college and career readiness: An analysis by state*. Washington, DC: American Institutes for Research

Morningstar, M. E., Lombardi, A., Fowler, C. H., & Test, D. W. (2017). A college and career readiness framework for secondary students with disabilities. *Career Development and Transition for Exceptional Individuals*, *40*, 79–91. doi:10.1177/2165143415589926

Morningstar, M. E., Lombardi, A. R., & Test, D. (2017). *Including college and career readiness within a multi-tiered systems of support framework*. Manuscript submitted for publication.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Center for Education Statistics. (2012). *The condition of education 2012* (Publication No. NCES 2012-045). Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/fastfacts/display.asp?id=16

National Center for Education Statistics. (2014). *Digest of education statistics: 2014*. Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/programs/digest/2015menu_tables.asp

Pornprasertmanit, S. (2014). *The unified approach for model evaluation in structural equation modeling* (Doctoral dissertation). University of Kansas, Lawrence.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org/

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696. doi:10.1080/00273171.2012.715555

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, *92*, 544–559. doi:10.1080/00223891.2010.496477

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, *73*, 5–26. doi:10.1177/0013164412449831

Revelle, W. (2016). *psych: Procedures for personality and psychological research*. Evanston, IL: Northwestern University. Retrieved from http://CRAN.R-project.org/package=psych Version=1.6.4

Rhemtulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*, 354–373. doi:10.1037/a0029315

Sanford, C., Newman, L., Wagner, M., Cameto, R., Knokey, A. M., & Shaver, D. (2011). *The post–high school outcomes of young adults with disabilities up to 6 years after high school: Key findings from the National Longitudinal Transition Study–2 (NLTS2)* (Publication No. NCSER 2011-3004). Menlo Park, CA: SRI International.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514. doi:10.1007/BF02296192

Shogren, K. A., & Garnier Villarreal, M. (2015). Developing student, family, and school constructs from NLTS2 data. *The Journal of Special Education*, *49*, 89–103. doi:10.1177/0022466913513336

Shogren, K. A., Wehmeyer, M. L., & Lane, K. L. (2016). Embedding interventions to promote self-determination within multi-tiered systems of supports. *Exceptionality*, 1–12. doi:10.1080/09362835.2015.1064421

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology*, *27*, 229–239. doi:10.1111/j.2044-8317.1974.tb00543.x

Steiger, J. H., & Lind, J. C. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the Annual Spring Meeting of the Psychometric Society, Iowa City, IA.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10. doi:10.1007/BF02291170

West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F., & Gabrieli, J. D. (2016). Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling. *Educational Evaluation and Policy Analysis*, *38*, 148–170. doi:10.3102/0162373715597298

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ωh. *Applied Psychological Measurement*, *30*, 121–144. doi:10.1177/0146621605278814

Manuscript received March 2017; accepted August 2017.