Article

Hammill Institute on Disabilities

Career Development and Transition for Exceptional Individuals
2023, Vol. 46(1) 4–14

DOI: 10.1177/21651434221081229
cdtei.sagepub.com
SAGE

Establishing Construct Validity of a Measure of Adolescent Perceptions of College and Career Readiness

Allison R. Lombardi, PhD[1], Graham G. Rifenbark, PhD[1], H. Jane Rogers, PhD[1], Hariharan Swaminathan, PhD[1], Ashley Taconet, MAT[1], Valerie L. Mazzotti, PhD[2], Mary E. Morningstar, PhD[3], Rongxiu Wu, PhD[1], and Shannon Langdon, M.Ed[1]

Abstract

The purpose of this study was to establish construct validity of a college and career readiness measure using a sample of youth with ($n = 356$) and without ($n = 1{,}599$) disabilities from five high schools across three U.S. states. We established content validity through expert item review, structural validity through initial field-testing, and convergent validity by correlating domain scores with school academic and behavioral data. A four-factor measurement model emerged representing the domains Ownership of Learning, Academic Engagement and Processes, Interpersonal Engagement, and Career Development. Domain scores were significantly correlated with achievement, college admission exam scores, and attendance. Implications for research and practice with an emphasis on transition service delivery via multi-tiered systems of support are discussed.

Keywords

secondary school, college and career readiness, age-appropriate transition assessment, validity, multi-tiered systems of support

Youth with disabilities have increasing career and postsecondary education options that allow for opportunities to be gainfully employed in adult life (Grigal et al., 2019). Despite this progress, disparities persist with regard to preparing these youth for college and careers. National data have consistently indicated the graduation rate for students with disabilities is lower (72%) than youth without disabilities (86%; National Center for Education Statistics [NCES], 2021a, 2021b). Postschool outcomes for youth with disabilities also lag well behind their peers without disabilities. Youth with disabilities remain less likely to attain and maintain competitive, integrated employment or pursue postsecondary education to prepare for long-term careers (Lipscomb et al., 2017; Newman et al., 2011). As such, it is imperative we identify and teach skills needed to be college and career ready while these youth are still in high school to ensure they are prepared for postschool life.

The purpose of this study was to establish construct validity (Messick, 1995) for a college and career readiness (CCR) measure of academic and nonacademic skills critical to adult life and useful for *all* students, with and without disabilities. We describe our efforts to pursue schoolwide data collection via a multi-tiered system of support (MTSS) approach. Specifically, to establish construct validity, we focused on three types of validity evidence: (a) content validity through expert item review, (b) structural validity through pilot field-testing, and (c) convergent validity by correlating resulting domain scores with school academic and behavioral data.

Focusing on CCR for *all* students may ensure students with disabilities have access to schoolwide supports as part of transition services. Students with disabilities benefit from access to and inclusion in general education (McLeskey et al., 2012; Sabornie et al., 2006). Specifically, correlational evidence indicates that more time in general education is positively related to enrollment in postsecondary education and a greater likelihood of employment after high school (Lombardi et al., 2013; Mazzotti et al., 2021; Rojewski et al., 2015). For the purpose of the current study, we focused on youth with disabilities who spend the majority of their school day with their peers without disabilities in general education settings.

[1]University of Connecticut, Storrs, USA
[2]The University of North Carolina at Charlotte, USA
[3]Portland State University, OR, USA

Corresponding Author:
Allison R. Lombardi, University of Connecticut, 249 Glenbrook Road Unit 3064, Storrs, CT 06269, USA.
Email: allison.lombardi@uconn.edu

Given the importance of maximizing time spent in general education settings for youth with disabilities, MTSS is an important framework to consider for transition service delivery. MTSS has been applied to school contexts to support implementation of evidence-based academic and social behavior practices for all students (Simonsen et al., 2010). A central tenet of MTSS is data-based decision-making, which entails the use of systematic screening and progress monitoring data to make critical support decisions for all students (Tier 1), including those with disabilities, and to determine who might be in need of more intensive support (Tiers 2 and 3).

MTSS ensures all students have access to high-quality evidence-based instruction, and data-based decision-making determines which students may need more intensive interventions and supports. This multi-tiered approach is well-established in academic (e.g., Response-to-Intervention) and behavioral (e.g., Positive Behavioral Interventions and Supports [PBIS]) contexts, yet less established with regard to CCR and transition services (Morningstar et al., 2018). In fact, there are no validated measures that help practitioners monitor progress toward CCR among students with and without disabilities. As such, the new CCR measure was designed to provide stakeholders (e.g., teachers, families, students, administrators) with data for progress monitoring and critical decision-making, which may in turn lead to shifts in instructional and curricular focus resulting in meaningful impact on college and/or career outcomes for all students. In this respect, MTSS allows for the inclusion of youth with disabilities in school-wide efforts focused on CCR. The challenge remains, however, in the lack of adequate measures of CCR that are valid and reliable for youth with and without disabilities. Therefore, the focus of the current study was to validate an instrument relevant to students with and without disabilities who might have similar educational experiences pertaining to CCR within general education settings.

CCR Assessment Framework

To underpin the development of the CCR measure, we used a five-domain assessment framework that represents academic and nonacademic skills: (a) academic engagement, (b) process-oriented skills, (c) interpersonal engagement, (d) ownership of learning, and (e) transition competencies (Lombardi et al., 2020). The five domains include nonacademic skills that are important to consider in defining CCR, such as soft skills, character traits, and social and behavioral skills (Gates et al., 2016; Krauss et al., 2016; Lombardi et al., 2015, 2019; Nagaoka et al., 2015; West et al., 2016). Specifically, the five domains evolved from extensive literature reviews within and outside of special education research and soliciting stakeholder feedback on proposed domains (Morningstar, Lombardi, et al., 2017), as well as preliminary measurement modeling using preexisting measures that mapped onto the framework (Lombardi et al., 2018). Drawing from these studies, the CCR assessment framework incorporates five domains of academic and nonacademic skills (Lombardi et al., 2020). Table 1 shows the domain definitions.

In this study, the primary research objective was to establish construct validity of the College and Career Readiness for Transition (CCR4T), a newly developed CCR measure. In this process, we examined content, structural, and convergent validity, which are three critical aspects to Unified Validity (Messick, 1995). Throughout this study, we had a specific focus on students with disabilities who spend the majority of their school day with their peers without disabilities in general education settings and approached data collection from an MTSS perspective. Specifically, we addressed the research questions (RQs):

RQ1: What is the dimensionality of the CCR4T constructs?
RQ2: Could the number of items be reduced while maintaining adequate coverage of the domains?
RQ3: Do the CCR4T domain scores relate to other relevant academic and behavioral data?

Method

The purpose of this study was to establish construct validity of a new measure. In this section, we describe the item development and refinement process, as well as our efforts to establish three types of construct validity evidence (content, structural, convergent). Throughout, we report how, and when, we made decisions to reduce items.

Measure

The CCR4T measures student perceptions of their own academic and nonacademic skills. The item development team comprised experts in secondary transition, CCR, school-wide PBIS, and measurement. We had the goal to create items that represent the breadth of content of the assessment framework and span the developmental continuum for youth in Grades 9 to 12 with and without disabilities. Initial item development began with 105 items that were the result of a prior study (Lombardi et al., 2018) in which preexisting measures that mapped onto the domains were selected. The team identified gaps in the coverage of these 105 items with the five-domain assessment framework (Lombardi et al., 2020). For example, one domain that was not well covered in the previous study was Interpersonal Engagement. Throughout, we maintained best practices in item-writing (e.g., clarity and simplicity of language, correct grammar, appropriate reading level, appropriate response scale). The team aimed to draft approximately 20 to 30 items per

Table 1. CCR4T Assessment Framework Domain Definitions and Sample Items.

Domain	Definition and sample items
Academic Engagement	Acquisition of academic content through interacting and engaging with the material, including cognitive and behavioral skills that students need to successfully engage with academics. These skills may include attendance, homework completion, active participation in class and less observable skills like making connections between content in different courses. *Sample items:* I connect content learned in different classes; I attend all of my classes.
Process-Oriented Skills	These skills may include test taking, studying, and time management, as well as critical thinking skills such as formulating problems, hypothesizing solutions, collecting evidence, analyzing the evidence, and communicating findings. These skills span across content areas. *Sample items:* I think of more than one way to solve a problem; I use test taking strategies
Interpersonal Engagement	Focuses on social skills with an emphasis on interactions with other individuals as well as understanding within themselves. Students with these skills show responsibility and adaptability across educational and noneducational settings, collaborate with peers, have an awareness in how others may be feeling or perceiving situations, and sense of belonging with the school. *Sample items:* I invite others to join me during activities; I understand how my own culture may be different from my peers.
Ownership of Learning	Entails sense of belonging, growth mindset, ownership of learning, and perseverance. Specifically, that all students have the ability to take academic risks and understand the importance of the growth that comes from making mistakes. *Sample items:* I overcome setbacks when facing challenges; I feel like I belong at school.
Transition Competencies	Focus on projects and activities that facilitate competency in employment, postsecondary education, and independent living, with a focus on understanding shifting cultures and responsibilities within each unique setting. *Sample items:* I can correctly fill out a job application; I know how to prepare a meal.

Note. CCR4T = College and Career Readiness for Transition.

domain to allow for item reduction techniques to be used after an initial field-testing phase.

Concurrently, the team conducted a mixed studies systematic literature review across scholarly and gray literature (Mazzotti et al., 2021). Findings revealed the five domains had a range of coverage across the literature with gaps identified between research, policy, and practice. Notably, only a small percentage of peer-reviewed studies (13%) included students with disabilities in CCR-focused research, confirming the need for a CCR measure.

In addition, the team conducted focus groups among key stakeholders, including state-, district-, and school-level educators (Morningstar et al., 2021). Specifically, stakeholder input enhanced and refined domains and items, ensuring relevancy and usefulness of the CCR4T and resulting in a clear understanding of educational context, experiences, and needs of secondary and transition stakeholders. Furthermore, stakeholders were queried on potential uses of data reports of individual student progress based on valid and reliable data that secondary educational providers can use for program evaluation and improvement. The resulting version of the CCR4T included 134 items. In this version, students rate statements from 1 to 6 (1 = *totally not like me,* 2= *mostly not like me*, 3 = *not like me*, 4 = *like me*, 5 = *mostly like me*, 6 = *totally like me*), which is a Likert-type type rating scale consistent with measuring affective domains in school contexts (McCoach et al., 2013).

Content validity. Prior to the field-testing the CCR4T with youth with and without disabilities, we investigated the content validity of the 134-item version of CCR4T by conducting an expert item review panel and gathering feedback on the items and response scale. Researchers with expertise in secondary special education and transition, PBIS, and school counseling, as well as practitioners (e.g., general and special education teachers) were recruited to join the panel (n = 37). For each item, respondents were given the item stem. Based on the item stem, respondents were asked to select the hypothesized domain the item belonged to (e.g., Process-Oriented Skills, Transition Competencies) or indicate that the item did not fit and should be deleted. They also indicated their certainty with respect to their choice (0 = *not very sure*, 1 = *somewhat sure*, 2 = *very sure*) and whether or not the item was inappropriate for students with disabilities. Raters independently rated items and thus were not aware of each other's ratings at any time throughout the process. We then estimated a content validity ratio (CVR; Lawshe, 1975) for each of the 134 items. Lawshe (1975) provides cutoff values (or critical values) for CVR based on the number of raters. We consulted a table of critical values to determine the number of experts who would need to agree the item is highly indicative of the construct in order for the result to be significantly greater than due to chance (Ayre & Scally, 2014). As the number of raters increase, the critical value for CVR decreases; therefore, based on

37 raters, we used a critical value of .33. In addition, we estimated the expected CVR for each hypothesized domain to evaluate item fit to domain. After estimating the expected CVRs, we conducted a sensitivity analysis to understand the impact of rater certainty on CVRs. Specifically, if a rater indicated they were *not very sure*, then their response for that item was omitted, leaving only responses with a degree of certainty.

Based on findings, we (a) combined items from the Academic Engagement and Process-oriented Skills domains to form an Academic domain; (b) moved four items from the Ownership of Learning to other domains (two to Academic, one to Interpersonal Engagement, one to Transition Competencies); and (c) removed items that were redundant and/or had low rater agreement. This process resulted in a 114-item version of the CCR4T to be field-tested. Domain definitions and sample items are listed in Table 1.

Structural validity. After revising the items based on the content validity stage, we field-tested the 114-item version of the CCR4T. We recruited school partners through national professional networks and organizations, including the National Technical Assistance Center for Transition (NTACT, now NTACT: the Collaborative), the Council for Exceptional Children's Division of Career Development and Transition (CEC-DCDT), the National Center on Positive Behavior Interventions and Supports (PBIS Center), and the Transition Coalition. We sent electronic invitations and reminders via the list servs and social media sites of these professional networks (i.e., Twitter, Facebook) and targeted secondary educators, specifically school counselors, administrators, and general and special education teachers. Recruitment materials encouraged viewers to broadly share the invitation with any secondary school practitioners within their own professional and personal networks. Interested school partners completed a data use agreement and parental notification and consent process. All study procedures were approved by the institutional review board of the university of the first author.

Participants. The field-testing sample was drawn from five high schools in three U.S. states: Connecticut, Maryland, and Pennsylvania. Four schools were public and one was private. Four schools were considered suburban and one school was rural. The high schools ranged in student enrollment between 175 and 1,200 students. Students with (n = 356, 18.2%) and without disabilities (n = 1,599, 81.7%) participated in the study (N = 1,955). Of those with disabilities, a range of categories were represented, with the majority considered high-incidence (83%). The sample demographics included students in Grade 9 (27%), Grade 10 (27%), Grade 11 (25%), and Grade 12 (21%). Due to COVID-19, we collected and coded learning modality as remote (i.e., instruction was received entirely in a virtual setting) and in-person (i.e., all or some amount of in-person instruction including various hybrid models such as 2 days/week in-person and 3 days/week virtual). In our sample, 41% of participants were identified as remote and 59% were considered in-person. The racial makeup of study participants was White (35%), Hispanic/Latinx (36%), Black (18%), or other/multiple races and Asian Americans (10%). In addition, 64% of participants were eligible for free or reduced-price meals. Table 2 provides participant characteristics.

Table 2. Characteristics of the Field-Testing Sample.

Characteristics	n^a	%
Gender		
Male	891	45.6
Female	1,064	54.4
Race/ethnicity		
Black	352	18
Asian	57	2.9
Hispanic	710	36.3
White	690	35.3
Other/multi	146	7.5
Grade		
9th	523	26.8
10th	521	26.6
11th	489	25
12th	422	21.6
Free/reduced-price meal status		
No	709	36.3
Yes	1,246	63.7
English language learner status		
No	1,294	89.3
Yes	155	10.7
IEP status		
No	1,599	81.8
Yes	356	18.2
Disability		
Specific learning disability	177	49.7
OHI (ADD/ADHD)	67	18.8
Emotional disturbance	44	12.4
Autism	26	7.3
Intellectual disability	21	5.9
Other/multiple disabilities	21	5.9
Method of instruction		
In-person	1,149	58.8
Remote	806	41.2

Note. IEP = Individualized Education Program; OHI = other health impairment; ADD = attention-deficit disorder; ADHD = attention-deficit/hyperactivity disorder.
[a]506 responses were missing, percentage reflects an n of 1,449.

Procedures. School partners provided the research team with two school-based data files. The first data file contained student email addresses and demographic data. The second data file, collected after the CCR4T survey was administered, contained academic and behavioral variables. Files were shared via a secure file sharing program and merged using the student's email address as a linking variable. Prior to sharing the data files, school partners sent parental notification letters to parents or guardians that explained the purpose of the study, the data collection plan, how the data were to be used, and any potential risks and benefits. Parents or guardians were given the choice to opt out of the study at this time. For those who did not opt out, students received email invitations to their school-issued email account with a personalized link to the CCR4T. Students took the CCR4T using the online program REDCap (Harris et al., 2009). On the first page of the CCR4T, students had to give assent to proceed with the CCR4T items. Researchers and school partners determined an administration window for data collection. School administrators determined time of day and class period(s) in which to administer the CCR4T and they were asked to ensure that all students were given the opportunity to take it, including those with disabilities. This approach aligns with MTSS as a designated Tier 1 universal screening assessment. In three of the five participating schools, there were coordinated efforts to devote a class period to completing the CCR4T (e.g., all math teachers across grades devote a class period to CCR4T administration). The other two schools promoted the CCR4T be taken as a homework assignment. As such, it was primarily administrators who directed classroom teachers to promote the CCR4T. We encouraged administrators to work directly with general and special educators to promote the CCR4T and ensure as many students could take it as possible to truly emulate a schoolwide data collection effort.

Data analyses. Using the overall sample (N = 1,955), we estimated descriptive statistics, inter-item correlations, and item-total correlations together and separately for students with (n = 356) and without (n = 1,599) disabilities. This enabled us to identify problematic items prior to assessing dimensionality. To assess dimensionality and reduce the number of items, we used a two-stage approach that entailed exploratory factor analysis (EFA) and item response theory (IRT).

Exploratory factor analysis. We estimated an EFA with an oblique rotation because we hypothesized that the constructs would be related to one another; in total, we requested up to an eight-factor solution to ensure dimensionality was fully assessed. Guided by theory, we evaluated a given EFA solution by consulting the data-model fit via the root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker–Lewis index (TLI), and standardized root mean square residual (SRMR); the magnitude of factor loadings; and the estimated latent correlations. Specifically, we considered an RMSEA of 0.05 or below, a CFI or TLI of 0.90 or higher, and an SRMR of 0.06 or lower as evidence of acceptable data-model fit (Hu & Bentler, 1998, 1999). With respect to factor loadings, we regarded estimates equal to 0.4 or higher as evidence of a properly functioning item (McCoach et al., 2013). Finally, we consulted latent correlations as a means for determining whether or not a given EFA solution was overextracted (i.e., nonsignificant correlation among constructs hypothesized to be positively related to one another). All EFAs were conducted in *Mplus* version 8.44 (Muthén & Muthén, 1998–2017) using a robust maximum likelihood (ML) estimator.

Item response theory models. We fitted IRT (Hambleton et al., 1991) models for each of the factors that resulted from the EFA. We estimated both Samejima's graded response model (GRM; Samejima, 1969) and the generalized partial credit model (GPCM; Muraki, 1992) to determine which provided the best data-model fit. These models contain the same number of estimated parameters; however, they arguably stem from different response mechanisms. These unidimensional IRT models were fitted using the mirt (Chalmers, 2012) package in R. We utilized the S-X2 fit statistic proposed by Orlando and Thissen (2000, 2003) and extended to the polytomous case by Kang and Chen (2011). To guide us, we used an alpha level of .01 to determine whether an item had problematic item-fit.

Specification of CCR. Using the final set of items retained from the EFA and the IRT item-fit, we proceeded to fit a multifactor confirmatory factor analysis (CFA) to determine to what extent the CCR constructs could be confirmed. We used the same thresholds for data-model fit as before (e.g., RMSEA less than or equal to 0.05).

Convergent validity. We used the same structural validity data set to next investigate convergent validity. After assessing dimensionality, we examined correlations between the newly confirmed domain scores and the academic and behavioral data that we collected in the second data file. Specifically, we examined academic variables identified as typical indicators of CCR (Allensworth & Clark, 2019, 2020; Hodara & Lewis, 2017; Welch et al., 2017) for youth with and without disabilities (Lombardi et al., 2015), as well as behavioral variables that have been linked to school engagement and academic performance (Feldman et al., 2014; Gottfried & Gee, 2017). These included grade point average, PSAT scores, and attendance data. As such, we selected these observed school variables to bolster the domain scores, which were based on self-report ratings.

Grade point average (GPA). Cumulative GPA was gathered from school records. Generally, partner schools calculated GPA scale ranging from 0.0 = "F" to 4.0 = "A."

There was some variation of the scale between schools. As such, we computed a categorical GPA variable that captured this variation. To do so, we first conducted an initial search of the participating high schools to determine their GPA point value to letter grade conversion. Upon review, we determined each school differed with regard to assigning a point value range to corresponding letter grade (e.g., 3.67–4.33 was the "A" range for one school and 4.0–5.0 was the "A" range for another). As such, we computed a categorical GPA variable that took into account these differences by using if/then statements (e.g., if school ID = 1, then compute "A" = 3.67–4.33). The GPA variable had numeric values of 1 to 5 entered into all analyses, which correspond to a letter-grade range as well as an approximate values of A, B, C, D, and F (A = 5, B = 4, etc.).

PSAT scores. If available, we used PSAT total combined scores gathered from school records. These data were available from four of the five high schools in the sample. The PSAT scores from the most recent semester available were selected for this analysis.

Attendance. Attendance was gathered from school records and was the number of full days a given student was absent during the semester in which the CCR4T was administered. These data were available from four of the five high schools in the sample.

Results

Prior to addressing the RQs, we removed student cases who had more than 80% missing responses (n = 50) before settling on our analytic sample (N = 1,955). We estimated the descriptive statistics and examined the item-total correlations. The item-total correlations for Academic Engagement ranged from .34 to .76, with the lowest corresponding to the reverse-scored item "I am often late for class." With respect to Interpersonal Engagement, item-total correlations ranged from .19 to .75; each of the four reverse-scored items had item-total correlations of .4 or smaller. Based on these results, we eliminated all five reverse scored items. The item-total correlations for Ownership of Learning ranged from .59 to .76; for Process-Oriented Skills, estimates ranged from .51 to .76; and for Transition Competencies, estimates ranged from .49 to .79. Results are organized by RQ.

RQ1: Assessing Dimensionality

To address our first RQ, we assessed dimensionality by conducting a linear EFA using ML estimation and treating the 1 to 6 response categories as continuous (Rhemtulla et al., 2012). In this process, we tested three- through eight-factor solutions and noticed as the number of factors extracted increased, the data-model fit improved. Due to these results, we determined the four-factor solution was optimal because it conformed to a simple structure and yielded factor loadings that were equal to 0.40 or greater for each of the four factors (McCoach et al., 2013). With respect to model fit, the RMSEA and SRMR were estimated to be 0.043 and 0.028, respectively, whereas the CFI and TLI were estimated to be 0.867 and 0.855. This solution also contained sensible latent correlations, with estimates ranging from .46 to .67. Finally, the four-factor solution was most consistent theoretically with our assessment framework and content validity results. This process resulted in retaining 85 items.

With the resulting 85 items, we fitted an EFA and requested two- through five-factor solutions. Again, the four-factor solution achieved the strongest theoretical and empirical support. Specifically, simple structure was observed with no dual loadings and each factor contained loadings that were .40 or greater. Among the constructs, the estimated latent correlations ranged from .48 to .70, indicating moderate to strong relationships. In addition, data-model fit was found to be acceptable with the RMSEA and SRMR estimated to be 0.044 and 0.027, respectively, whereas the CFI and TLI were estimated to be 0.882 and 0.870, respectively. We used the final results to rename the domains and the final four-factor model to be Ownership of Learning, Academic Engagement and Processes, Interpersonal Engagement, and Career Development. Notably, the hypothesized Transition Competencies domain was renamed Career Development because the remaining items pertained only to career development activities (job interviewing, career search and exploration, career interests, etc.) and items pertaining to broader adult roles and responsibilities (e.g., daily living) were dropped.

RQ2: Evaluation of Item-Fit to Inform Item Reduction

Our second research question pertained to item reduction. It was particularly important to reduce the number of items to establish a brief, feasible measure that could be flexible to meet the needs of school contexts while simultaneously maintaining adequate coverage of the domains. We fitted unidimensional IRT models to further investigate item fit. We fitted two competing models for each construct: the GRM (Samejima, 1969) and the GPCM (Muraki, 1992). The GRM model was superior in terms of global model fit. We, then, examined Orlando and Thissen's (2002, 2003) S-X2 fit statistics, where a significant result indicates poor item fit.

After consulting S-X2 values, items with poor fit in the Ownership of Learning factor were flagged. Specifically, p-values for these items were less than 0.001, and the items were removed from the scale. In total, 22 items were retained on the Ownership of Learning factor. For Academic

Engagement and Processes, S-X2 values were significant for two items. One item contained a low item-total correlation. Therefore, these three items were removed, leaving a total of 23 items for this factor. For Interpersonal Engagement, S-X2 values were significant for two of the six items on this factor. However, due to the small number of items on this construct, we elected to retain all six items. Finally, for Career Development, two items from this construct had significant S-X2 values, as well as high local dependence indices. Therefore, we removed these items and retained a total of 14 items for this factor. Overall, we retained 64 items after evaluating item fit based on the IRT mode fit statistics.

As a final step to assessing dimensionality and evaluating item-fit (RQ1 and RQ2), we estimated a correlated factors model using the final 64 items spanning the four constructs. This resulted in a significant chi-square test of exact fit; however, acceptable approximate fit resulted. Namely, the RMSEA and SRMR were estimated to be 0.045 and 0.057, respectively; whereas the CFI and TLI were estimated to be 0.852 and 0.846, respectively. Although the CFI and TLI estimated were below 0.95, this finding replicates patterns observed in the methodological literature. Specifically, as the number of degrees of freedom increases, the RMSEA tends to reflect better fit, whereas indices such as CFI and TLI tend to show poorer fit (Ding et al., 1995; Kenny et al., 2015). The standardized factor loading estimates from this fitted model ranged from 0.51 to 0.83 and therefore, indicator reliability ranged from .26 to .68. When consulting the latent correlations among the four constructs, estimates ranged from .65 and .87 with the latter corresponding to the association between Academic Engagement and Processes and Ownership of Learning. Due to the high correlations, we investigated a higher-order specification of CCR. This model was specified such that the four identified constructs, namely, Ownership of Learning, Academic Engagement and Processes, and Career Development, served as lower-order factors and their shared variance was explained by a higher-order factor of CCR. The lower factors were identified using the marker variable approach where the factor loading and the manifest intercept for the first indicator for each construct were fixed to 1.0 and 0.0, respectively.

After fitting this model with the robust maximum likelihood estimator, $\chi^2(1{,}948, N = 1{,}955) = 9{,}690.77$, acceptable global model fit was observed: The RMSEA and SRMR were estimated to be 0.045 and 0.058, respectively. The CFI and TLI were estimated to be 0.851 and 0.846, respectively. The standardized factor loadings ranged from 0.62 to 0.79 for Ownership of Learning, 0.59 to 0.77 for Academic Engagement and Processes, 0.55 to 0.83 for Interpersonal Engagement, and 0.509 to 0.805 for Career Development. In terms of the paths from the higher-order CCR factor to the lower-order factors, the gamma paths were as follows: .72 for Interpersonal Engagement, .84 for Career Development, .92 for Ownership of Learning, and .94 for Academic Engagement and Processes. Overall, the process of conducting the EFA and IRT analyses to confirm dimensionality (RQ1) and inform item reduction (RQ2) as well as fit the final correlated factors model provides structural validity evidence of the CCR4T domain scores.

RQ3: Correlations With Academic and Behavioral Data

To address RQ3, we examined interrelationships between the domain scores and selected academic and behavioral data routinely collected by schools. Table 3 shows results of correlations between the four CCR4T domain scores and categorical GPA, PSAT scores, and attendance. Pearson correlations were computed for all variables except for categorical GPA, which was polychoric. Overall, GPA had statistically positive correlations with PSAT/SAT and four CCR4T domain scores ($p < .01$), and had the strongest correlation with Academic Engagement ($r = .33, p < .01$) compared with other three domains. The PSAT/SAT had statistically positive correlations with the four domain scores as well, but correlations were weaker compared with GPA. Attendance was statistically negatively correlated with all the variables ($p < .1$) except for the Career Development domain. These results provide evidence of convergent validity for the CCR4T domain scores.

Discussion

The purpose of this study was to establish construct validity of the CCR4T. Ultimately, our process leading up to the resulting four domain scores addresses critical aspects of Unified Validity (Messick, 1995) by establishing three types of evidence (content, structural, convergent). Moreover, we reduced the number of items from 134 to 64, resulting in a measure with sufficient item parsimony. The four-factor solution provides structural validity evidence and builds on previous CCR measurement work (Lombardi et al., 2018). Specifically, this prior study did not account for the social and emotional learning skills comprised within the Interpersonal Engagement domain. The Ownership of Learning domain is a new and important outcome of the study because it was not included in past measurement research. Also, we renamed one domain from Transition Competencies to Career Development to better reflect the remaining set of items. In particular, we found the items pertaining to daily living did not fit with the overall measurement model. With these changes, the four-factor model represents substantial progress in measuring adolescent perceptions of CCR in a valid and reliable way.

Table 3. Correlations Between CCR4T Domain Scores and Selected Academic and Behavioral Variables.

Variable	*n*	1	2	3	4	5	6	7
1.GPA	1,714	1						
2. PSAT	1,397	.590**	1					
3. Absence	1,718	−.331**	−.127**	1				
4. Academic engagement	1,641	.339**	.102**	−.095**	1			
5. Ownership of learning	1,636	.304**	.212**	−.073**	.843**	1		
6. Interpersonal engagement	1,715	.157**	.081**	−.048*	.643**	.631**	1	
7. Career development	1,664	.079**	.057*	.003	.741**	.733**	.622**	1
M			847.15	3.91	4.24	4.79	4.23	4.26
SE			4.47	0.22	0.02	0.02	0.03	0.02

Note. Correlations between GPA and other variables are polychoric correlation; all other correlations are Pearson. CCR4T = College and Career Readiness for Transition; GPA = grade point average.
$^*p < .05$. $^{**}p < .01$.

We established convergent validity by examining the relationship between the four domain scores and school academic and behavioral variables. Our findings confirm the domain scores are positively related to GPA and PSAT scores and negatively related to attendance (number of days absent in the semester in which the CCR4T was taken). Notably, these findings are novel to the transition assessment literature base; school-based variables such as GPA, college admissions exam scores, and attendance are rarely if ever included in validation studies of age-appropriate transition assessments. As such, many self-report age-appropriate transition assessments lack this type of validity evidence. Moreover, within the context of MTSS, best practices encourage data teams to examine multiple sources of self-report and school data that include similar variables (Flannery et al., 2019; Freeman et al., 2021). Findings from the current study support the notion that age-appropriate transition assessment could be better aligned with schoolwide efforts to promote and support CCR for all students (Morningstar et al., 2018).

We anticipate the CCR4T will be most useful for youth with disabilities who spend the majority of their school day in general education settings. Using MTSS as an underpinning to data collection, we used schoolwide or "universal" approaches to recruitment and data collection. In doing so, we emphasized all youth with disabilities should be included in data collection efforts. This approach resulted in a final sample comprised of youth with disabilities with a range of disability types, the majority of which are considered high-incidence (83%). However, we did not want to exclude students on the basis of category alone. As shown in Table 2, the final analytic sample represented a range of disability categories.

Limitations

While our study advances the measurement efforts toward comprehensive CCR skills for high school students with and without disabilities, there are limitations to consider in the interpretation of the findings. First, the sample was limited to five high schools in three states. School personnel coordinated the dates and class periods in which the CCR4T was administered. While we encouraged school partners to include youth with and without disabilities, the focus on general education settings to administer the CCR4T influenced which students had access to the survey. Our goal in developing the CCR4T is to not only provide a reliable and valid measure of CCR for secondary students with and without disabilities, but to also provide a reliable and valid age-appropriate transition assessment for students with disabilities who are likely to spend the majority of their school day with their peers without disabilities in general education settings. Given the wide variability in instruction and the highly segregated nature of school and transition experiences among students with low-incidence disabilities (Morningstar, Kurth, & Johnson, 2017), it was beyond the scope of this research to study these subpopulations in-depth. Others have continued to pursue research aligning skills and supports needed for youth with low-incidence disabilities to be college and career ready (Morningstar, Zagona, et al., 2017). Therefore, the focus on students with high-incidence disabilities was intentional and our study procedures were organized such that those youth with disabilities who spend the majority of their time in general education classrooms were included. Field-testing the CCR4T on students with a range of disabilities, including targeted efforts to increase participation from students with low-incidence disabilities, is warranted.

With regard to methodological limitations, the CFI and TLI were estimated to be 0.851 and 0.846 which are below the cutoff recommended by Hu and Bentler (1998, 1999) for incremental model fit values; however, the RMSEA and SRMR were estimated to be 0.045 and 0.057, respectively, demonstrating the absolute fit indices were well within the recommended cut off range. This finding is not uncommon as the simulated model used by Hu and Bentler, which in turn informed their recommendations, departs for the measurement model for the CCR4T and therefore does not generalize to this context. Specifically, the number of indicators per factor, the magnitude of standardized loadings, and the number of degrees of freedom all have an impact on global fit indices (Ding et al., 1995; Kenny et al., 2015). Finally, we were not able to evaluate the functioning (e.g., invariance testing) of the CCR4T on the basis of disability status. This was due to the small sample of students with disabilities ($n = 356$); specifically, the explication of CCR requires 196 parameters and would make this effort considerably underpowered.

Implications for Future Research

There are immediate next steps for future research to consider in light of our findings. First, a large-scale validation effort of the CCR4T should take place, which will require a larger sample size across more high schools. It will be important to conduct a differential item functioning (DIF) analysis within the IRT framework to more closely examine item fit or conduct invariance testing via multiple-group confirmatory factor analysis to ensure the CCR4T functions similarly for students with and without disabilities. Also, after collecting a large enough sample, expert reviewers may consider setting cutoff scores within each domain to attach qualitative descriptors to score ranges (e.g., does not meet, meets, exceeds). These types of descriptors will be helpful to educators in using scores to make support decisions about students.

After these immediate next steps in research, a long-term plan to evaluate the uses of the CCR4T as a student-level measure to support school personnel in developing and evaluating CCR programming for secondary students across the four domains will be important to prioritize. In a similar vein, future research is needed related to professional development emphasizing the integration of the CCR4T data within team-based decision-making that involves special educators and their colleagues, such as school counselors. School-based teams could use CCR4T domain scores for progress monitoring purposes and to make data-based decisions about all students, with and without disabilities. In fact, stakeholders from the focus group research explicitly discussed this as a potent outcome of the CCR4T (Morningstar et al., 2021).

Implications for Policy and Practice

CCR first emerged from educational policy directives and as yet has divergent skill emphases within education, particularly with regard to operationalizing and measuring nonacademic skills (Green et al., 2021). The recent findings from focus groups with stakeholders (i.e., state education agencies, district administrators, school educators) reflected alignments and connections toward using the CCR4T for evaluating, developing, and implementing programs and practices that facilitate CCR for all secondary students (Morningstar et al., 2021). Stakeholders described examples of how it would ultimately support educators and students to align instruction to students' self-reported CCR strengths while addressing gaps in skills and experiences. Operationalizing a CCR-focused approach such as the CCR4T can help bridge long-standing contextual factors unique to secondary schools and adolescent learners.

Specifically for students with disabilities, CCR4T data could be used as part of the age-appropriate transition assessment process to inform measurable postsecondary goals as well as student goals in the Individualized Education Programs (IEPs). For those with and without disabilities, domain scores could be used for progress monitoring purposes in the spirit of data-based decision-making. Integrating the CCR4T within broad, schoolwide data collection efforts will ensure all students have access to CCR supports. An important long-term developmental step will be to build an online portal that houses the CCR4T and data reports for various users, such as students, parents, teachers, and administrators. Ideally, these users will have different types of accounts that enable them to log in, take the assessment, view data reports of progress over time, and use the data toward other systems and processes that facilitate planning, such as the IEP process and school counseling supports. The CCR4T will be a validated measure that has the potential to underlie such a system and could be a useful planning tool that fosters collaboration between general and special educators, school counselors, and students and their families.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A190170 to the University of Connecticut. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

ORCID iDs

Allison R. Lombardi https://orcid.org/0000-0002-7254-8820
Ashley Taconet https://orcid.org/0000-0003-2407-5114
Valerie L. Mazzotti https://orcid.org/0000-0003-2616-8859

References

Allensworth, E. M., & Clark, K. (2019). *Are GPAs an inconsistent measure of college readiness across high schools? Examining* *assumptions about grades versus standardized test scores*. University of Chicago Consortium on School Research.

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, *49*(3), 198–211. https://doi.org/10.3102/0013189X20902110

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79–86. https://doi.org/10.1177/0748175613513808

Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

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*(2), 119–143. https://doi.org/10.1080/10705519509540000

Feldman, M. A., Ojanen, T., Gesten, E. L., Smith-Schrandt, H., Brannick, M., Totura, C. M. W., Alexander, L., Scanga, D., & Brown, K. (2014). The effects of middle school bullying and victimization on adjustment through high school: Growth modeling of achievement, school attendance, and disciplinary trajectories. *Psychology in the Schools*, *51*(10), 1046–1062. https://doi.org/10.1002/pits.21799

Flannery, K. B., Kato, M. M., & Horner, R. H. (2019). *Using outcome data to implement multi-tiered behavior support (PBIS) in high schools*. Technical Assistance Center on Positive Behavioral Interventions and Supports.

Freeman, J., Kern, L., Lombardi, A. R., Morningstar, M. E., & Mazzotti, V. L. (2021). *PBIS high school implementation: Leveraging features of the PBIS framework to support college and career readiness*. Technical Assistance Center on Positive Behavioral Interventions and Supports.

Gates, S., Lippman, L., Shadowen, N., Burke, H., Diener, O., & Malkin, M. (2016). *Key soft skills for cross-sectoral youth outcomes*. YouthPower. https://www.youthpower.org/resources/key-soft-skills-cross-sectoral-youth-outcomes

Gottfried, M. A., & Gee, K. A. (2017). Identifying the determinants of chronic absenteeism: A bioecological systems approach. *Teachers College Record*, *119*(7), 1–34.

Green, S., Sanczyk, A., Chambers, C., Mraz, M., & Polly, D. (2021). College and career readiness: A literature synthesis. *Journal of Education*. Advance online publication. https://doi.org/10.1177/00220574211002209

Grigal, M., Papay, C., Smith, F., Hart, D., & Verbeck, R. (2019). Experiences that predict employment for students with intellectual and developmental disabilities in federally funded higher education programs. *Career Development and Transition for Exceptional Individuals*, *42*(1), 17–28. https://doi.org/10.1177/2165143418813358

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (*Vol. 2*). SAGE.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hodara, M., & Lewis, K. (2017). *How well does high school grade point average predict college performance by student urbanicity and timing of college entry?* (REL 2017-250). U.S. Department of Education.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424

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

Kang, T., & Chen, T. T. (2011). Performance of the generalized S-X2 item fit index for the graded response model. *Asia Pacific Education Review*, *12*, 89–96. https://doi.org/10.1007/s12564-010-9082-4

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Krauss, S. M., Pittman, K. J., & Johnson, C. (2016). *Ready by design: The science and art of youth readiness*. Washington, DC: The Forum on Youth Investment.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575.

Lipscomb, S., Hamison, J., Liu Albert, Y., Burghardt, J., Johnson, D. R., & Thurlow, M. (2017). *Preparing for life after high school: The characteristics and experiences of youth in special education. Findings from the National Longitudinal Transition Study 2012. Vol. 2: Comparisons across disability groups. Full report* (NCEE 2017-4018). National Center for Education Evaluation and Regional Assistance.

Lombardi, A. R., Doren, B., Gau, J. M., & Lindstrom, L. E. (2013). The influence of instructional settings in reading and math on postsecondary participation. *Journal of Disability Policy Studies*, *24*(3), 170–180. https://doi.org/10.1177/1044207312468766

Lombardi, A. R., Freeman, J., & Rifenbark, G. G. (2018). Modeling college and career readiness for adolescents with and without disabilities: A bifactor approach. *Exceptional Children*, *84*(2), 159–176. https://doi.org/10.1177/0014402917731557

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*(3), 142–151. https://doi.org/10.1177/2165143414534888

Lombardi, A. R., Monahan, J., & Morningstar, M. (2020). Integrating college and career readiness into transition education. In K. Shogren & M. L. Wehmeyers (Eds.), *Handbook of adolescent transition education for youth with disabilities* (2nd ed., pp. 268–281). Routledge.

Lombardi, A. R., Rifenbark, G. G., Freeman, J., & Harvey, M. (2019). Measuring grit in adolescents with and without disabilities. *Journal of Disability Policy Studies*, *30*(2), 67–77. https://doi.org/10.1177/1044207319863635

Mazzotti, V. L., Morningstar, M. E., Lombardi, A., Kwiatek, S., Taconet, A., Buddeke, K., Monahan, J., & Harris, R. (2022). Policy and Practice Considerations to Support College and Career Readiness for Youth with Disabilities: A Systematic Mixed Studies Review [Manuscript submitted for publication]. Department of Special Education and Child Development, University of North Carolina at Charlotte.

McCoach, D. B., Gable, R. K., & Madura, J. P. (2013). *Instrument development in the affective domain*. Springer.

McLeskey, J., Landers, E., Williamson, P., & Hoppey, D. (2012). Are we moving toward educating students with disabilities in less restrictive settings? *The Journal of Special Education*, *46*(3), 131–140. https://doi.org/10.1177/0022466910376670

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Morningstar, M. E., Kurth, J. A., & Johnson, P. E. (2017). Examining national trends in educational placements for students with significant disabilities. *Remedial and Special Education*, *38*(1), 3–12. https://doi.org/10.1177/0741932516678327

Morningstar, M. E., Lombardi, A., Fowler, C., & Test, D. W. (2017). A college and career readiness framework for secondary students with disabilities. *Career Development and Transition for Exceptional Individuals*, *40*(2), 79–91. https://doi.org/10.1177/2165143415589926

Morningstar, M. E., Lombardi, A., Mazzotti, V. M., Buddeke, K., Langdon, S., Taconet, A., Kwiatek, S., & Rifenbark, G. (2022). *Perspectives of college and career readiness among educational stakeholders*. Manuscript in preparation.

Morningstar, M. E., Lombardi, A., & Test, D. (2018). Including college and career readiness within a multitiered systems of support framework. *AERA Open*, *4*(1). 1–11. https://doi.org/10.1177/2332858418761880

Morningstar, M. E., Zagona, A. L., Uyanik, H., Xie, J., & Mahal, S. (2017). Implementing college and career readiness: Critical dimensions for youth with severe disabilities. *Research and Practice for Persons with Severe Disabilities*, *42*(3), 187–204. https://doi.org/10.1177/1540796917711439

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *ETS Research Report Series*, *1992*(1), i–30. https://doi.org/10.1002/j.2333-8504.1992.tb01436.x

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.).

Nagaoka, J., Farrington, C. A., Ehrlich, S. B., & Heath, R. D. (2015). *Foundations for young adult success: A developmental framework. Concept paper for research and practice*. University of Chicago Consortium on Chicago School Research.

National Center for Education Statistics. (2021a). *Public high school graduation rates*. https://nces.ed.gov/programs/coe/indicator/coi

National Center for Education Statistics. (2021b). *Students with disabilities*. https://nces.ed.gov/programs/coe/indicator/cgg

Newman, L., Wagner, M., Knokey, A. M., Marder, C., Nagle, K., Shaver, D., & Wei, X. (2011). *The post high school outcomes of youth adults with disabilities up to 8 years after high school: A report from the NLTS-2* (NCSER 2011-3004). National Center for Special Education Research. https://ies.ed.gov/ncser/pubs/20113005/pdf/20113005.pdf

Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement*, *24*(1), 50–64. https://doi.org/10.1177/01466216000241003

Orlando, M., & Thissen, D. (2003). Further investigation of the performance of S-X2: An item fit index for use with dichotomous item response theory models. *Applied Psychological Measurement*, *27*(4), 289–298. https://doi.org/10.1177/0146621603027004004

Rhemtulla, M., Brosseau-Liard, P. É., & 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*(3), 354–373. https://doi.org/10.1037/a0029315

Rojewski, J. W., Lee, I. H., & Gregg, N. (2015). Causal effects of inclusion on postsecondary education outcomes of individuals with high-incidence disabilities. *Journal of Disability Policy Studies*, *25*(4), 210–219. https://doi.org/10.1177/1044207313505648

Sabornie, E. J., Evans, C., & Cullinan, D. (2006). Comparing characteristics of high incidence disability groups: A descriptive analysis. *Remedial and Special Education*, *27*(2), 95–104. https://doi-org/10.1177/07419325060270020701

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, *34*, 1–97. https://doi.org/10.1002/j.2333-8504.1968.tb00153.x

Simonsen, B., Shaw, S. F., Faggella-Luby, M., Sugai, G., Coyne, M. D., Rhein, B., Madaus, J. W., & Alfano, M. (2010). A schoolwide model for service delivery: Redefining special educators as interventionists. *Remedial and Special Education*, *31*(1), 17–23. https://doi.org/10.1177/0741932508324396

Welch, M. J., Abulhab, A., & Therriault, S. B. (2017). *College and career readiness in Boston: Understanding and tracking competencies and indicators*. American Institutes for Research. https://www.air.org/sites/default/files/Boston-CCR-AIR-Lit-Rev-FINAL-2017APRIL.pdf

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*(1), 148–170. https://doi.org/10.3102/0162373715597298