2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #50542

Measuring the Reliability and Validity of the CS Goal Affordances and Endorsements Instrument (Fundamental)

Dr. Joseph C Tise, Institute for Advancing Computing Education

Dr. Tise is a Senior Education Researcher at the Institute for Advancing Computing Education. As an educational psychologist, he has training and expertise in quantitative and mixed-methods education research, particularly within computer science and other STEM domains.

Dr. Monica McGill, Institute for Advancing Computing Education

Monica McGill is President & CEO of the Institute for Advancing Computing Education (formerly known as CSEdResearch.org). Have previously worked in industry and academia, McGill is using her experiences as a computer scientist, professor, and researcher to enable others to build a strong foundation of CS education research focused on all learners while also honoring their backgrounds and experiences.

Measuring the Reliability and Validity of the CS Goal Affordances and Endorsements Instrument (Fundamental)

Joseph C. Tise[1] and Monica McGill[2]

[1,2]Institute for Advancing Computing Education
[1]joe@csedresearch.org
[2]monica@csedresearch.org

Abstract

Background. Goal congruity in computer science (CS)—the alignment between the types of goals CS affords and the goals students personally endorse—can influence students' intentions to pursue CS in K–12 classrooms and beyond. However, few instruments exist with evidence of reliability and validity that are designed to assess goal congruity in CS among middle and high school students.

Research Question. Our research question for this study was: *What evidence supports the validity and reliability of the CS Goal Affordances and Endorsements scales for use with middle and high school students?*

Methodology. We adapted four scales (goal affordances-agentic, goal affordances-communal, goal endorsements-agentic, goal endorsements-communal) from a previous study that was not focused on CS. We then conducted a cross-sectional survey using the four adapted scales to examine students' communal and agentic goal affordances for computer science (CS) and their personal endorsements of communal and agentic goals. Participants ($n = 292$) were U.S. middle school ($n = 154$) and high school students ($n = 136$).

Results. Face validity was generated by critique of the scales by a group of 15 researchers, educators, industry representatives, and non-profit representatives from across the U.S. Evidence of validity based on internal structure was generated via confirmatory factor analysis (CFA) and measurement invariance testing. The CFA models exhibited good model-data fit, and scalar measurement invariance held for each model across, race/ethnicity, home language, and grade. By gender, scalar invariance held for the CS goal affordances scale and metric for the personal goal endorsements scale. By timepoint, metric held for the CS goal affordances scale and scalar for the personal goal endorsements scale. We found McDonald's Omega for each of the four subscales to be $\omega = .76$ and above.

Discussion. This goal congruity instrument can be used with confidence by researchers, educators, and evaluators interested in goal congruity within a CS context among students in grades 6-12.

1 Introduction and Background

Coursework that focuses on computer science (CS), a subfield of engineering, is often a pre-requisite or co-requisite to engineering in postsecondary education. In fact, CS can often be found in the engineering departments or schools at postsecondary institutions, mainly due to CS being born in large part from engineering [1, 2]. Due to this, CS and engineering still share fundamental constructs, such as problem solving, trouble shooting, abstraction, and design thinking [3–6].

The fundamental cognitive processes that underlie both fields are taught frequently in K-12 classrooms and lessons that include CS and/or engineering. At the same time, emphasis on CS and engineering education continues to grow. Therefore, it is important that researchers, policy makers, and educators understand the factors that influence student decisions to pursue or abstain from CS and engineering education (at all levels) and professions.

Motivation research and theory illustrates the influence of goals on human behavior and decisions [7–9]. Indeed, many motivation theories assume goals (on some level) direct decisions, such as whether or not to pursue a particular discipline like CS or engineering. Goals can be classified numerous ways, but one increasingly common way is through the lens of *goal congruity theory* [9]. At its heart, goal congruity theory posits that people pursue education and professions that *afford* them opportunities to pursue goals which they personally *endorse* (i.e., value). In this way, goal congruity theory implicates two primary psychosocial constructs: perceived goal *affordances* and personal goal *endorsements*.

Perceived goal affordances hinge on one's perceptions of a domain or profession, and research has shown middle and high school student perceptions of engineering and CS are shaped by culture, experiences, and stereotypes [10, 11]. Building on early foundational research, Diekman et al. demonstrated that people frequently view STEM disciplines as emphasizing agentic qualities such as independence, competition, and personal accomplishment [9]. Subsequent work by Montoya et al. [12] reinforces this pattern, showing that agentic perceptions persist across contexts.

Concomitantly, goal congruity theory asserts that students' personal goal endorsements (i.e. the extent to which they personally value certain goals) interact with perceived goal affordances to drive behavior and decisions [9, 13]. A person therefore experiences goal congruity when their perceived goal affordances of a certain domain/profession align with their personal goal endorsements. For example, if a student perceives CS/Engineering as a profession which makes people wealthy, allows autonomy in work, and enables the pursuit of power/status, that student will be goal congruent with CS/Engineering to the extent that they personally value those types of professional goals.

Importantly, perceived goal affordances are generally considered domain specific (e.g., one's perceived goal affordances of CS may not be the same as those related to creative writing), while personal goal endorsements are considered domain general (i.e., invariant across contexts/domains). To this end, goal congruity theory has been investigated within STEM disciplines for at least the past decade. Seminal empirical work on this front has shown that communal goal endorsements mediates the well-established relationship between gender and interest in STEM careers–specifically, girls were more likely to endorse communal goals, and higher endorsements of communal goals in turn predicted lower interest in STEM careers ([9]).

Research since then has replicated some of these results. For instance, Lewis et al. found that undergraduate women more highly endorsed communal goals than undergraduate men, and that endorsements of communal goals predicted sense of belonging in computing.

Despite the robust theoretical and empirical work within other STEM disciplines completed heretofore, goal congruity theory is empirically underexplored within K-12 CS contexts. To establish a well-grounded foundation of goal congruity research in K-12 CS, we must be able to effectively measure perceived goal affordances and personal goal endorsements. Such effective measurement demands evidence of reliability and validity. To date, high quality measures tailored to CS and that show strong evidence of validity and reliability have not been produced.

Given the identified gaps in the literature, we sought to gather preliminary evidence of validity and reliability for our adapted measure of CS goal affordances and personal goal endorsements. Specifically, our primary research question was:

> *What evidence supports the validity and reliability of the CS Goal Affordances and Endorsements scales for use with middle and high school students?*

2 Methodology

To answer our research question, we conducted a quantitative study using a cross-sectional survey of middle and high school students across three U.S. states. We obtained external IRB approval prior to data collection. The survey assessed students motivational beliefs, experiences with CS, and numerous demographic characteristics. Surveys were completed online during school hours. In the present study, we analyze the psychometric properties of two motivation measures: students' CS goal affordances, and their personal goal endorsements.

2.1 Participants

Participants ($n = 292$) included students from grades 6-12 across three U.S. states (California, Michigan, and Illinois) (see Table 1). All students enrolled at participating sites were eligible to participate, though one school in California administered the survey only to students taking a CS course. The participating schools were not necessarily equivalent regarding common metrics (e.g., SES, funding, size, locale). Student demographic data included self-reported gender, race/ethnicity, home language, grade level, and individual education plan (IEP) status. Broadly, the sample was mostly Hispanic/Latinx or White, had at least some prior CS education experiences, and spoke English at home.

Participants completed the survey during school hours to ensure equitable access to the internet and electronic devices. Representation was fair across the grades (see Table 2).

2.2 Measures

The demographic variables also served as the grouping variables by which we tested our primary measures for measurement invariance.

Primary measures included two scales which were adapted from prior literature, with the measures we used reported in [15] . The first measured students' perceived *CS goal affordances*, and the second measured their personal endorsements of communal and agentic goals (i.e., how important each type of goal is to them personally). The adapted CS goal affordances measure [16] comprised two factors aligned with the two main focal types of goals: communal and agentic.

Table 1: Sample descriptive statistics for categorical variables

	N	%
Prior CS Experience	216	73.97
Gender: Boy	125	42.81
Gender: Girl	109	37.33
Black or African-American	29	9.93
Asian	11	3.76
Hispanic/Latinx	112	38.36
Native American or Alaska Native	23	7.88
White	58	19.86
Multiracial	40	13.70
Race/Ethnicity self-described or not provided	19	6.51
Support for Learning Differences: Yes	30	10.27
Support for Learning Differences: No	112	38.26
Support for Learning Differences: Unknown	143	48.97
Home language: English	219	75.00
Home language: Not English	69	23.63

Table 2: Grades Represented in the Sample

	N	%
6th	27	9.25
7th	65	22.26
8th	62	21.23
9th	26	8.90
10th	15	5.14
11th	53	18.15
12th	42	14.38

Both subscales were measured by six items each. Students rated how much they agreed or disagreed with each item from 1 (Strongly Disagree) to 7 (Strongly Agree). Importantly, perceived goal affordances are inherently domain specific, so we could not simply use an extant measure of goal affordances which was not contextualized to CS.

The adapted *personal goal endorsements* measure (which we created from the related items in Belanger et al.) also comprised two factors (communal and agentic). Communal goal endorsements was measured via two items and agentic goal endorsements by four items. Students rated how personally important each of the goal endorsement items were from 1 (Not at all important) to 7 (Extremely Important).

2.3 Analyses

Preliminary data analyses included missing data analyses (patterns and rates of missingness), descriptive statistics, and data cleaning. Very little data were missing (2.0%), and missing data were missing at random (MAR). Descriptive statistics indicated there was sufficient variance in

Table 3: CS Goal Affordances Descriptive Statistics and Items

Item #	Item Text	Mean (SD)	Median
AA1	A job in computer science would let me be in charge.	3.11 (1.53)	3.00
AA2	I think computer scientists make a lot of money.	4.41 (1.71)	5.00
AA3	Computer scientists often get more attention than people in other jobs.	3.30 (1.40)	3.00
AA4	Computer scientists often compete with others as part of their job.	3.72 (1.49)	4.00
AA5	I think a job in computer science would let me be my own boss.	3.39 (1.50)	3.00
AA6	I think a job in computer science would let me get really good at what I do.	3.46 (1.61)	4.00
CA1	I think computer science helps make society better.	3.94 (1.65)	4.00
CA2	I think computer science helps people.	4.22 (1.63)	4.00
CA3*	Computer science is interesting, but it doesn't help society much.	4.92 (1.35)	5.00
CA4	Work that computer scientists do can help meet the needs of the community.	3.77 (1.55)	4.00
CA5*	Computer scientists mostly work alone.	4.50 (1.51)	4.00
CA6	I think a job in computer science would let me connect with others.	3.41 (1.65)	4.00

Note: Items marked * are negatively worded relative to the construct definition. Min-Max for all items was 1 to 7. AA = Agentic Affordance; CA = Communal Affordances.

each item to support further analyses. Normality assumptions were checked via the asymptotic one-sample Kolmogorov-Smirnov and Shapiro-Wilk tests. All items violated these tests of normality, so we used a robust estimator (MLR) for all CFA analyses.

Primary analyses included confirmatory factor analyses and subgroup invariance testing. For whole-sample CFAs, model fit was assessed via common fit indices (CFI/TLI, RMSEA, SRMR, and the χ^2 test of model fit). Good model-data fit is indicated by $CFI \geq .95$, $SRMR \leq .05$, and $RMSEA \leq .06$ [17]. Factor variances were set to 1 so all item loadings could be estimated. To assess measurement invariance (i.e., Configural, Metric, and Scalar; [18]), we leveraged ΔCFI and ΔRMSEA indices, such that measurement invariance was achieved if after applying each successive invariance constraint, CFI decreased by no more than .01 (or increased) and RMSEA increased by no more than .015 (or decreased) [19, 20]. For all analyses except the grade-level invariance testing, middle and high school participants were considered together as a single sample.

3 Results

Descriptive statistics pertaining to the CS goal affordances items are presented in Table 3 and those pertaining to the personal goal endorsement items in Table 4. Notably, we observed sufficient variance in responses across all items, with the average hovering around the mid-point of the scale for most items.

Table 4: Personal Goal Endorsements Descriptive Statistics and Items

Item #	Item Text	Mean (SD)	Median
CE1	I want to help people who need it with my work.	4.41 (1.85)	4.00
CE2	In my job, I want to connect with others.	4.79 (1.79)	5.00
AE1	In my future job, I want to be in charge.	4.14 (1.78)	4.00
AE2	I want to make a lot of money in my future job.	5.50 (1.77)	6.00
AE3	I want people to notice me because of my future job.	4.19 (1.91)	4.00
AE4	I like to compete with others.	3.82 (1.87)	4.00
AE5	I want to be my own boss in my future job.	4.17 (1.85)	4.00
AE6	I want to get really good at what I do in my future job.	5.59 (1.76)	6.00

Note: Due to an administrative error, only two of the four Communal Endorsement items were administered.
CE = Communal Endorsements; AE = Agentic Endorsements.

Min-Max for all items was 1 to 7.

3.1 Confirmatory Factor Analysis & Reliability

Figure 1 shows the final model for the two CS goal affordances scales with standardized coefficients. The model had subpar fit without any correlated errors ($\chi^2 = 208.91, p = .000$; $CFI = .893$, $TLI = .867$, $RMSEA = .100$, $SRMR = .056$), but correlating the errors of items AA1 and CA6, and also CA1 and CA2 improved the model fit substantially. Thus, the final model fit was excellent ($\chi^2 = 15.45, p = .163$; $CFI = .996$, $TLI = .990$, $RMSEA = .38$, $SRMR = .060$). Note that items CA3 and CA5 were negatively worded (relative to the construct definition), so their coefficients are supposed to be negative. Upon reviewing Table 3, we see that items AA1 and CA6 both pertain to having a job in CS and are somewhat similar in syntactic structure. These similarities may be why their errors were correlated. Similarly, we see that CA1 and CA2 are nearly the same in meaning, thus correlating their errors makes sense as well.

Our sample produced mostly reliable data as measured by McDonald's Omega (agentic affordances $\omega = .89$; communal affordances $\omega = .89$).

Figure 2 shows the final model for the two personal goal endorsements scales with standardized coefficients. The model had subpar fit without any correlated errors ($\chi^2 = 108.75, p = .000$; $CFI = .880$, $TLI = .840$, $RMSEA = .120$, $SRMR = .116$), but empirical modification indices indicated that several items had correlated errors (as shown in the figure), which improved the model fit substantially. The final model fit was good overall, though the SRMR metric was slightly elevated ($\chi^2 = 18.61, p = .136$; $CFI = .992$, $TLI = .983$, $RMSEA = .039$, $SRMR = .098$). It is not immediately apparent why some of these errors were correlated, but we do see similarities in some of the item wordings (e.g., AE1, AE5).

Our sample produced mostly reliable data as measured by McDonald's Omega (agentic endorsements $\omega = .86$; communal endorsements $\omega = .76$). McDonald's Omega is interpreted similarly to Cronbach's alpha, but is generally considered a better measure of reliability because it takes into account specific item loadings and error variances, unlike Cronbach's alpha [21]. Notably, the reliability for communal affordances items was lower than for agentic affordances, though still an acceptable level. This lower reliability may have been partially caused by the scale

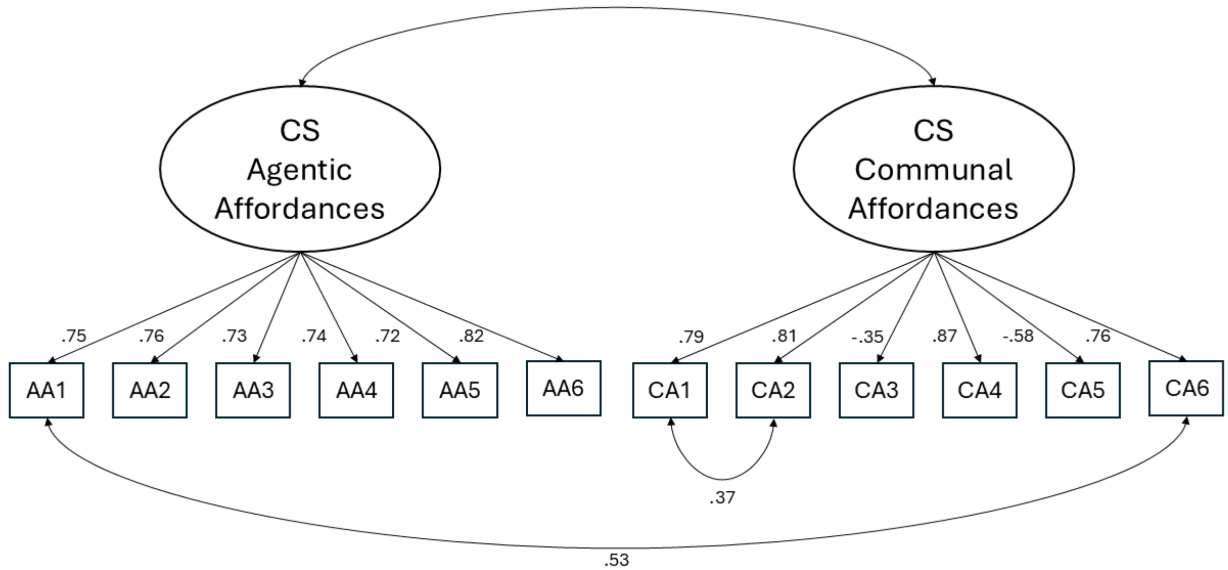

Figure 1: Final Measurement Model for CS Goal Affordances (Standardized Coefficients)

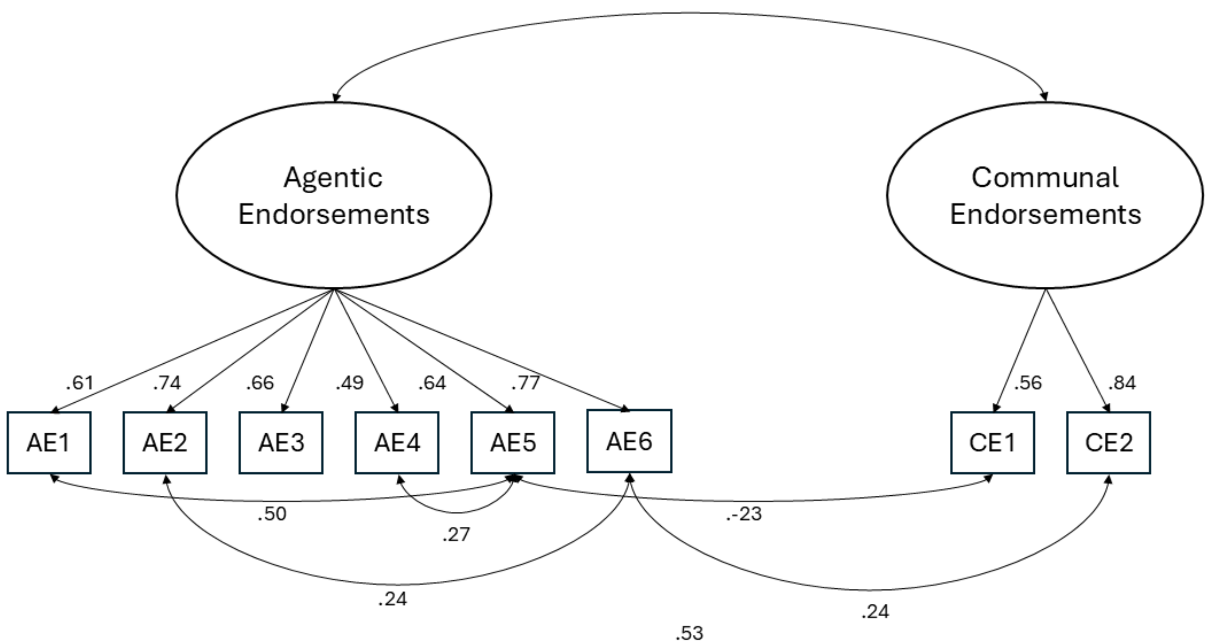

Figure 2: Final Measurement Model for Personal Goal Endorsements (Standardized Coefficients; all were significant)

only having two items compared to the six included for agentic endorsements.

3.2 Measurement Invariance

To evaluate how similarly the measures functioned among subgroups of students, we conducted multiple measurement invariance analyses via a multi-group CFA approach. Specifically, we sequentially assessed three levels of invariance: configural, metric, and scalar. Configural invariance indicates that, among subgroups, the measure has the same number of factors and the same items load onto the same factors. Metric invariance builds upon these constraints to evaluate if the item loadings are equivalent among subgroups, and finally scalar invariance further constrains the model to evaluate if items have equivalent intercepts across subgroups (see Brown for more information on multi-group CFA).

Regarding the CS goal affordances scales, results indicated that we achieved scalar invariance by gender, ethnicity, grade, and home language. Table 5 details the model fit statistics. We observed only metric longitudinal invariance due to the reduced CFI and increased RMSEA values observed after constraining item intercepts.

Regarding the personal goal endorsements scales, results indicated that we achieved scalar invariance by ethnicity, grade, home language, and longitudinally. Unfortunately, only metric gender invariance was achieved, indicating that boy and girl participants did not necessarily interpret the items in the same way. Table 6 details the model fit statistics.

Table 5: CS Goal Affordances Invariance Results

		Metric		**Scalar**		
	N	ΔCFI	ΔRMSEA	ΔCFI	ΔRMSEA	Invariance Achieved
Gender	234	-0.006	0.005	-0.005	0.003	Scalar
Race/Ethnicity	273	-0.002	-0.016	-0.009	-0.008	Scalar
Home Language	285	0.003	-0.008	0.003	-0.005	Scalar
Grade	286	-0.011	-0.006	-0.003	-0.009	Scalar
Longitudinal	125	0.000	0.025	-0.051	0.015	Metric

Note: Some students did not report demographic information used in these analyses. Therefore the sample size was reduced for each analysis accordingly.

Table 6: Personal Goal Endorsements Invariance Results

		Metric		Scalar		
	N	ΔCFI	ΔRMSEA	ΔCFI	ΔRMSEA	Invariance Achieved
Gender	285	0.001	-0.009	-0.019	0.018	Metric
Race/Ethnicity	272	-0.015	0.001	-0.012	-0.003	Scalar*
Home Language	284	-0.007	0.011	-0.006	0.005	Scalar
Grade	285	0.007	-0.03	0.005	-0.018	Scalar
Longitudinal	125	-0.006	-0.001	-0.005	-0.002	Scalar

Note: Some students did not report demographic information used in these analyses. Therefore the sample size was reduced for each analysis accordingly.

*The two criteria were contradictory (ΔCFI indicated Configural, while ΔRMSEA indicated Scalar)

4 Discussion

In this study, we presented initial evidence of reliability and validity for a newly adapted measure of CS goal affordances and personal goal endorsements. Extant measures in the CS and engineering education literature (e.g., [9, 23]) have generally not undergone rigorous psychometric analyses, particularly related to measurement invariance across subgroups of students at the middle and high school levels. It is critically important to do such analyses because truly understanding the factors that influence student decisions to pursue (or abstain from) CS education hinges on accurate measurement of their CS-related motivations. From the CFAs, we saw that item residuals were not fully independent in either the CS goal affordances or the personal goal endorsements scales. This could indicate that some other unidentified underlying factor accounts for significant variance in the item responses. Or, it could indicate that students were unable to sufficiently distinguish the differences among these correlated items. Still, correlating these item residuals resulted in good model fit to the observed data.

We also observed scalar invariance for the CS goal affordances scale regarding gender, race/ethnicity, home language, and grade. For the personal goal endorsements scale, we observed scalar invariance by race/ethnicity, home language, grade, and time point. Importantly, the fact that we observed scalar invariance for both measures by grade level indicates that considering the sample as one (i.e., grouping middle and high schoolers together) for other analyses is statistically warranted. As noted in Table 6, the race/ethnicity invariance results for the personal goal endorsements scale were mixed: based on the ΔCFI metric, the scale exhibited only configural invariance, but based on the ΔRMSEA metric, it exhibited scalar invariance. These contradictory findings may have been caused by relatively small subsample sizes within some of the race/ethnicity categories (e.g., *Asian*) and/or the relatively numerous number of categories upon which we assessed invariance (six different race/ethnicity designations).

Thus, the results pertaining to the race/ethnicity invariance analyses should be interpreted cautiously. Notably, we only observed metric invariance by gender for the goal endorsements scales. This is important because to meaningfully compare means across subgroups (e.g., girls vs. boys), scalar invariance must be achieved [24]. Without scalar invariance, it is impossible to say whether gender differences in goal endorsements are caused by true differences or by disparate interpretations of the scale. Though our exact measure was not used in prior research which found gender differences in goal endorsements (e.g., [9, 14, 23]), it still raises a possible confound in the results of these and other studies because the measures used in those studies did not undergo measurement invariance testing (or at least such analyses were not reported). Future researchers therefore must assess measurement invariance to more confidently interpret results.

We observed metric invariance for the CS goal affordances scale by time point and metric invariance for the personal goal endorsements scale by gender. These results indicate that students were not interpreting the items on these scales similarly across time (for the CS goal affordances scale) or gender (for the personal goal endorsements scale). Technically speaking, one should not compare latent means across subgroups for which scalar invariance was not achieved. Therefore, researchers ought to be cautious with their use of these scales to compare means across time and gender, respectively.

4.1 Limitations

Several limitations arose throughout the study. First, our race/ethnicity variable was somewhat limited in representation. For example, none of our sample identified as *Middle Eastern/North African* or *Hawaiian/Pacific Islander*. Further, we disaggregated the *Asian* category into *Asian (East), Asian (South), Asian (Southeast),* and *Asian (Central)*, but due to low response rates in each of these disaggregated categories, we had to aggregate them into one *Asian* subgroup for invariance testing. This admittedly excludes potentially important cultural nuances among these race/ethnicity categories. Future research with this measure should be conducted to ensure it maintains invariance among these groups which were underrepresented (or not represented at all) in the present study.

Second, our sample drew from only six schools in the U.S. While we did achieve an acceptable sample size for our analyses (generally), our sample was not fully representative of middle and high school students in the U.S. Thus, future research ought to expand this representation across more states and geographic locales.

Third, the majority of our sample (74%) reported having some CS learning experiences in school at some point in their education. It could be that students without any CS learning experiences have very different perceptions of CS (i.e., perceived goal affordances of CS), and since these students were relatively sparse in our sample, their data may have been overshadowed by the majority who had some CS learning experiences. Again, a more representative sample in future research would help ameliorate this concern.

Further related to representativeness, the sample had relatively few 10th grade students participate ($n = 15$). Although invariance analyses indicated scalar invariance by grade level for both scales, a larger sample of 10th grade students would strengthen our confidence in the robustness of the scales.

Finally, due to an administrative error, we inadvertently administered only two of the four communal endorsements items from the personal goal endorsements scale. Though that subscale still exhibited strong reliability and adequate model-data fit, we had less coverage of this theoretical component compared to agentic endorsements. Given the two omitted items were worded very similarly to the two included items, we think we captured communal endorsements sufficiently–with admittedly less coverage than intended.

5 Conclusions

Taken together, this study shows that the CS goal affordances scale and the personal goal endorsements scale both demonstrate strong psychometric properties. Researchers and educators can use these scales with confidence and empirical evidence of reliability and validity. Still, we recommend that invariance be assessed each time new data are collected and latent factors are included in analytic models. This is because invariance may not necessarily hold for every sample. Nevertheless, our results provide a solid starting foundation and preliminary evidence of validity and reliability for these measures.

Although these scales are not the first to be developed/adapted for a CS context, they are among the first to be rigorously examined for reliability, validity, and subgroup measurement invariance. These measures can contribute to further research used to refine goal congruity theory within middle and high school CS contexts, and/or be used to provide baseline data for educators interested in their students perceptions of CS affordances and personal goals.

6 Acknowledgments

This project was supported by funding from Amazon Future Engineer, Google, Microsoft Philanthropies, and Siegel Family Endowment. We thank the PROACT-CS Consortium members who helped shape the study and the instrumentation, including representatives from Michigan, California, and Illinois Departments of Education; Chicago Public Schools; Computer Science Teachers Association; Code.org; and researchers from University at Buffalo (Dr. Adrienne Decker) and New York City Research Alliance (Dr. Cheri Fancsali).

References

[1] Claude E. Shannon. A symbolic analysis of relay and switching circuits. *Electrical Engineering*, 57(12): 713–723, December 1938. ISSN 0095-9197, 2376-7804. doi: 10.1109/EE.1938.6431064. URL https://ieeexplore.ieee.org/document/6431064/.

[2] A. Newell, A. Perlis, and H.A. Simon. What is computer science? *Science*, (157):1373–4, 1967. URL https://www.cs.cmu.edu/~choset/whatiscs.html.

[3] Noemi V. Mendoza Diaz, Russ Meier, Deborah A. Trytten, and So Yoon Yoon. Computational Thinking Growth During a First-Year Engineering Course. In *2020 IEEE Frontiers in Education Conference (FIE)*, pages 1–7, October 2020. doi: 10.1109/FIE44824.2020.9274250. URL https://ieeexplore.ieee.org/document/9274250. ISSN: 2377-634X.

[4] David Jonassen, Johannes Strobel, and Chwee Beng Lee. Everyday Problem Solving in Engineering: Lessons for Engineering Educators. *Journal of Engineering Education*, 95(2):139–151, April 2006. ISSN 10694730. doi: 10.1002/j.2168-9830.2006.tb00885.x. URL https://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.2006.tb00885.x.

[5] Sean Brophy, Stacy Klein, Merredith Portsmore, and Chris Rogers. Advancing Engineering Education in P-12 Classrooms. *Journal of Engineering Education*, 97(3):369–387, July 2008. ISSN 1069-4730, 2168-9830. doi: 10.1002/j.2168-9830.2008.tb00985.x. URL https://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.2008.tb00985.x.

[6] Alannah Oleson, Brett Wortzman, and Amy J. Ko. On the Role of Design in K-12 Computing Education. *ACM Transactions on Computing Education*, 21(1):1–34, March 2021. ISSN 1946-6226. doi: 10.1145/3427594. URL https://dl.acm.org/doi/10.1145/3427594.

[7] Andrew J. Elliot and Marcy A. Church. A Hierarchical Model of Approach and Avoidance Achievement Motivation. *Journal of Personality and Social Psychology*, 72(1):218–232, 1997.

[8] Albert Bandura. Human agency in social cognitive theory. *American Psychologist*, 44(9):1175–1184, 1989. ISSN 0003-066X. doi: 10.1037/0003-066X.44.9.1175. Place: US ISBN: 1935-990X(Electronic),0003-066X(Print).

[9] Amanda B. Diekman, Elizabeth R. Brown, Amanda M. Johnston, and Emily K. Clark. Seeking Congruity Between Goals and Roles: A New Look at Why Women Opt Out of Science, Technology, Engineering, and Mathematics Careers. *Psychological Science*, 21(8):1051–1057, August 2010. ISSN 0956-7976, 1467-9280. doi: 10.1177/0956797610377342. URL http://journals.sagepub.com/doi/10.1177/0956797610377342.

[10] Sapna Cheryan, Allison Master, and Andrew N. Meltzoff. Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology*, 6(FEB):1–8, 2015. ISSN 16641078. doi: 10.3389/fpsyg.2015.00049.

[11] Balamuralithara Balakrishnan and Foon Siang Low. Learning Experience and Socio-Cultural Influences on Female Engineering Students' Perspectives on Engineering Courses and Careers. *Minerva*, 54(2):219–239, June 2016. ISSN 0026-4695, 1573-1871. doi: 10.1007/s11024-016-9295-8. URL `http://link.springer.com/10.1007/s11024-016-9295-8`.

[12] Amanda Kay Montoya, Allison Master, and Sapna Cheryan. Increasing Interest in Computer Science through Group Work: A Goal Congruity Approach, May 2020. URL `https://osf.io/ahgfy`.

[13] Kathryn L. Boucher, Melissa A. Fuesting, Amanda B. Diekman, and Mary C. Murphy. Can I Work with and Help Others in This Field? How Communal Goals Influence Interest and Participation in STEM Fields. *Frontiers in Psychology*, 8, May 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.00901. URL `https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.00901/full`.

[14] Colleen Lewis, Paul Bruno, Jonathan Raygoza, and Julia Wang. Alignment of Goals and Perceptions of Computing Predicts Students' Sense of Belonging in Computing. In *Proceedings of the 2019 ACM Conference on International Computing Education Research*, pages 11–19, Toronto ON Canada, July 2019. ACM. ISBN 978-1-4503-6185-9. doi: 10.1145/3291279.3339426. URL `https://dl.acm.org/doi/10.1145/3291279.3339426`.

[15] Joseph C. Tise and Monica M. McGill. Predicting Middle and High School Students' Intentions to Pursue CS from their Communal and Agentic Goal Affordances and Endorsements. In *Proceedings of the 2025 ACM Conference on International Computing Education Research V.1*, pages 82–92, Charlottesville USA, August 2025. ACM. ISBN 9798400713408. doi: 10.1145/3702652.3744211. URL `https://dl.acm.org/doi/10.1145/3702652.3744211`.

[16] Aimee L. Belanger, Mansi P. Joshi, Melissa A. Fuesting, Erica S. Weisgram, Heather M. Claypool, and Amanda B. Diekman. Putting Belonging in Context: Communal Affordances Signal Belonging in STEM. *Personality and Social Psychology Bulletin*, 46(8):1186–1204, August 2020. ISSN 0146-1672. doi: 10.1177/0146167219897181. URL `https://doi.org/10.1177/0146167219897181`.

[17] Li-tze Hu and Peter M. Bentler. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1):1–55, January 1999. ISSN 1070-5511. doi: 10.1080/10705519909540118. URL `https://doi.org/10.1080/10705519909540118`. Publisher: Routledge.

[18] Keith F. Widaman and Steven P. Reise. Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, and S. G. West, editors, *The science of prevention: Methodological advances from alcohol and substance abuse research*, pages 281–324. American Psychological Association, Washington, DC, US, 1997. ISBN 978-1-55798-439-5. doi: 10.1037/10222-009.

[19] Gordon W. Cheung and Roger B. Rensvold. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2):233–255, 2002. ISSN 1070-5511. doi: 10.1207/S15328007SEM0902_5.

[20] Fang Fang Chen. Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3):464–504, July 2007. ISSN 1070-5511, 1532-8007. doi: 10.1080/10705510701301834. URL `https://www.tandfonline.com/doi/full/10.1080/10705510701301834`.

[21] Andrew F. Hayes and Jacob J. Coutts. Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But.... *Communication Methods and Measures*, 14(1):1–24, January 2020. ISSN 1931-2458, 1931-2466. doi: 10.1080/19312458.2020.1718629. URL `https://www.tandfonline.com/doi/full/10.1080/19312458.2020.1718629`.

[22] Timothy Brown, A. *Confirmatory Factor Analysis for Applied Research*. Methodology in the Social Sciences Series. Guilford Press, New York, 2 edition, 2015. ISBN 978-1-4625-1779-4.

[23] Kathleen Isenegger, Kari L. George, Paul Bruno, and Colleen M. Lewis. Goal-Congruity Theory Predicts Students' Sense of Belonging in Computing Across Racial/Ethnic Groups. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, pages 1069–1075, Toronto ON Canada, March 2023. ACM. ISBN 978-1-4503-9431-4. doi: 10.1145/3545945.3569834. URL `https://dl.acm.org/doi/10.1145/3545945.3569834`.

[24] Robert J. Vandenberg and Charles E. Lance. A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, 3(1):4–70, January 2000. ISSN 1094-4281, 1552-7425. doi: 10.1177/109442810031002. URL `https://journals.sagepub.com/doi/10.1177/109442810031002`.