Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

Arif Rachmatullah, Jessica Mislevy, & Krystal Thomas
SRI

January 2026

Postsecondary
Teaching with Technology
Collaborative

Suggested Citation

Rachmatullah, A., Mislevy, J., & Thomas, K. (2026). Technical report: Psychometric properties of self-directed learning survey measures in broad access institution contexts. SRI.

The Postsecondary Teaching with Technology Collaborative (the Collaborative) is a U.S. Department of Education, Institute of Education Sciences R&D center, co-led by SRI Education and the Community College Research Center at Teachers College, Columbia University, in partnership with Achieving the Dream. The Collaborative uses research findings to build the capacity of institutions and instructors to establish inclusive learning environments and incorporate technology in ways that improve learning and success in postsecondary online courses. Our research analysis aims to contribute to knowledge and understanding of how instruction can support students to employ a constellation of motivational and metacognitive processes and certain applied learning processes—which we refer to as self-directed learning (SDL)—to manage their learning more effectively in online courses and increase their postsecondary success.

Visit us at https://postseccollab.org/ • Follow us on X: @PostsecCollab

CONTENTS

ABSTRACT

The Postsecondary Teaching with Technology Collaborative (the Collaborative) conducted secondary analyses of survey datasets to advance the conceptualization and measurement of self-directed learning (SDL) in higher education. Drawing on data from two of the Collaborative's prior studies—the Institutional Policy and Practice (IPP) Study and the Rapid Cycle Evaluations (RCEs)—we examined how existing instruments used to measure SDL perform in the context of broad access institutions and across the diverse student populations they serve.

This technical report describes the data sources, harmonization procedures, and psychometric analyses used to gather evidence for the reliability, validity, and fairness of the SDL survey measures. It is intended for researchers and other technical audiences interested in understanding the methodological approach and psychometric evidence that underpin the Collaborative's work. Detailed results for each measure are presented in the appendices.

This report outlines the methods and analyses used to re-examine validity evidence of self-directed learning measures, offering methodological insights and detailed results for researchers.

OVERVIEW

The Postsecondary Teaching with Technology Collaborative (the Collaborative) have used several survey measures to gain insights into students' self-directed learning (SDL) in online settings. However, most of these surveys have been conceptualized and validated with study samples that may not fully reflect students across broad access institutions (e.g., Harrison & Vallin, 2017; Jackson, 2018). Prior research has raised concerns about the limited generalizability and cross-cultural validity of commonly used SDL measures, noting that cultural norms and contextual factors can shape how students interpret and respond to items (Midkiff et al., 2018; Roth et al., 2016). These issues highlight the importance of re-examining the psychometric properties of SDL instruments in more diverse student populations – an effort that directly motivates the Collaborative's analyses described in this report.

Building on these concerns, the Collaborative conducted analyses to further investigate the psychometric properties of the SDL survey measures used in our studies. Our main goal was to gather more insights into the measures' construct validity and generalizability across different demographic factors (i.e., first generation status, gender, and race/ethnicity). We used data collected from two different studies, the Institutional Policy and Practice (IPP) Study and Rapid Cycle Evaluations (RCEs). IPP study was conducted in Fall 2022 and the RCEs was conducted in Fall 2022 to Fall 2023. RCEs in particular involved STEM course instructors participating in

tests of specific, technology-enabled strategies to support students' development of SDL skills and mindsets. Some of the survey measures were administered only in the RCEs and some items had different scale points across the two studies. Therefore, before examining the psychometric properties of the survey measures, data harmonization was done to enable us merging the two datasets and ultimately increasing the sample size. We describe more about the measures and analytical approach in the following sections before we share the results.

MEASURES

The Collaborative have collected data across two studies—RCEs and IPP—on a total of 10 constructs related to SDL, using five measures drawn from existing instruments and prior research. Table 1 presents each survey measure, its associated construct, the number of items, and the response format. As shown in Table 1, the format of certain measures differed between the RCE and IPP studies, requiring us to harmonize the data. Additionally, the RCEs sample includes both pre- and post-intervention data, which we treat as independent groups.

Table 1. Summary of Survey Measures, Constructs, Item Counts, and Response Formats Used in the RCEs and IPP Studies

Survey Measure	Construct	# Items	RCE (n = 2,323)	IPP (n = 246)
Fixed Mindset (Dweck et al., 1995)		3	Scale 1-6	—
Learning Strategies (McGuire et al., 2015)		11	True/False	—
MAI (Schraw & Dennison, 1994)	Comprehension Monitoring	7	True/False	Scale 1-5
	Debugging strategies	5	True/False	Scale 1-5
	Evaluation	6	True/False	Scale 1-5
OSLQ (Barnard et al., 2009)	Goal-setting	5	Scale 1-5	—
	Help seeking	4	Scale 1-5	Scale 1-5
	Time management	RCEs 4; IPP 3	Scale 1-5	Scale 1-5
PERTS (Hanson, 2017)	Self-efficacy	3	Scale 1-6	Scale 1-5
	Sense of Belonging	4	Scale 1-6	Scale 1-5

Note: A dash (—) means that the survey measure or construct was not administered. Detailed participant demographic characteristics are provided in Appendix A.

ANALYTICAL APPROACH

Data Harmonization

Before conducting psychometric analyses, we performed a data harmonization process (Tsai et al., 2024; Quan & Wang, 2025), particularly for the MAI and PERTS measures, due to differences in response formats between the RCE and IPP studies. For the MAI items, we adopted the binary True/False format used in the RCE study. To align the IPP data, which used a 5-point Likert scale, we recoded responses: 1 and 2 were mapped to 0 (False), and 3, 4, and 5 were mapped to 1 (True). This decision was based on the observation that the frequency distribution of True and False responses in the RCE data closely aligned with the distribution of 1–2 and 3–5 responses, respectively, in the IPP data.

For the PERTS survey, we standardized on the 5-point Likert scale used in the IPP study. To convert the 6-point scale from the RCE data, we combined responses as follows: 3 and 4 were recoded as 3, 5 as 4, and 6 as 5. This approach was also informed by distributional similarities between the combined categories in RCE and the 5-point scale in IPP.
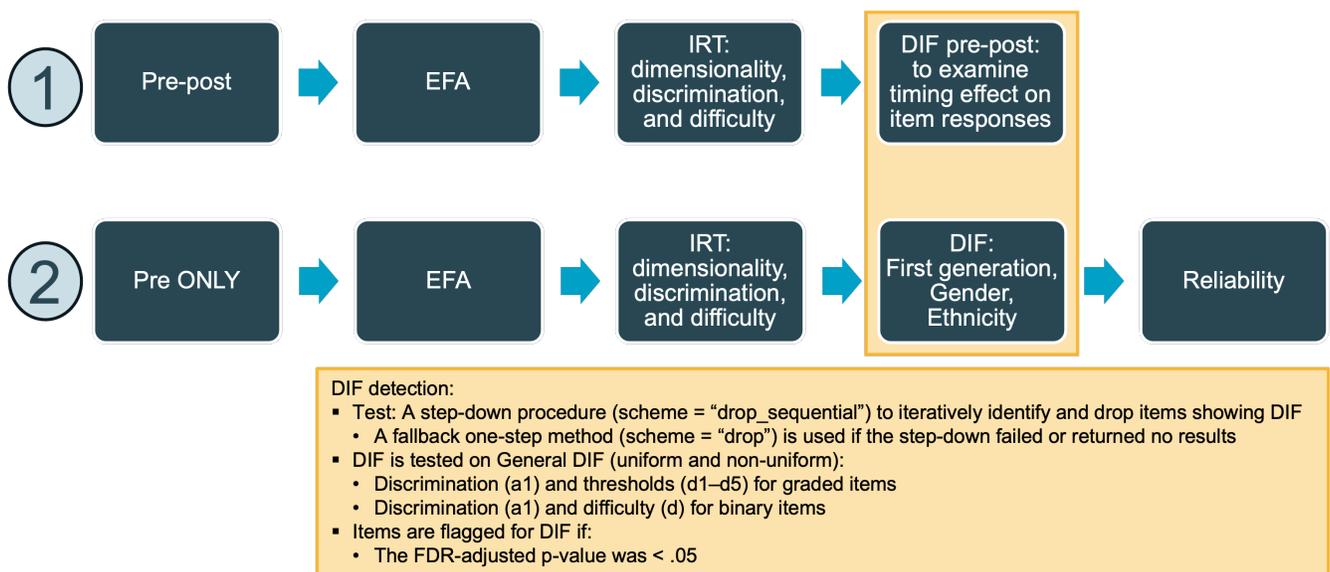
Psychometric Analysis

We conducted a multi-step psychometric analysis to evaluate the quality and fairness of five measures using pre- and post-intervention data from the IPP and RCE studies. To enable cross-study comparisons, we treated the IPP dataset as representing the "pre" time point. Our analysis began with Exploratory Factor Analysis (EFA) to identify item groupings, followed by Item Response Theory (IRT) modeling to assess dimensionality and item-level properties.

Model fit and item fit were evaluated using multiple statistical indices. For IRT model fit, we examined the root mean square error of approximation (RMSEA), with RMSEA values ≤ .08 indicating good fit. Item-level fit was assessed using the $S–X^2$ statistic (Orlando & Thissen, 2000), which compares observed and model-predicted response patterns; items with significant p-values ($p < .05$) after false discovery rate (FDR) correction were flagged for potential misfit. Item discrimination and difficulty parameters (denoted as a, d) were also evaluated to characterize how well each item differentiated among respondents and captured the latent trait continuum.

We then conducted Differential Item Functioning (DIF) analysis on the combined pre-post data to detect potential response shifts due to timing or context. For the RCE study, we conducted additional pre-post DIF analyses separately for the treatment and comparison groups to determine whether observed DIF was driven by the intervention itself or by test administration timing. Noteworthy, the modeling was done assuming time is not nested within person, because such a model is more complicated and computationally heavy. Thus, we applied a model where pre and post are two different groups.

Next, we conducted a second DIF analysis using only pre-data, testing for measurement invariance across first-generation status, gender, and race/ethnicity. A step-down method was used to identify and remove DIF-affected items, with a one-step fallback applied when necessary. Both uniform and non-uniform DIF were tested, with items flagged based on FDR-adjusted p-values below .05. All of the IRT and DIF analyses were done using the *mirt* package in R (Chalmers, 2012). Finally, we recalculated Cronbach's alpha using pre-data to verify internal consistency of each scale. Figure 1 shows the overview of the approach.

Figure 1. Overview of the Approach

DIF detection:
- Test: A step-down procedure (scheme = "drop_sequential") to iteratively identify and drop items showing DIF
  - A fallback one-step method (scheme = "drop") is used if the step-down failed or returned no results
- DIF is tested on General DIF (uniform and non-uniform):
  - Discrimination (a1) and thresholds (d1–d5) for graded items
  - Discrimination (a1) and difficulty (d) for binary items
- Items are flagged for DIF if:
  - The FDR-adjusted p-value was < .05

RESULTS

Fixed Mindset

The full results of psychometric property for the Fixed Mindset measure are presented in Appendix B. EFA on combined pre-post data supported a strong single-factor structure, with very high factor loadings (0.826–0.962). One item—"Your intelligence is something about you that you can't change very much"—exhibited extremely high discrimination (28.30), suggesting possible local dependence or item dominance. All items showed significant misfit, though RMSEA values remained in the moderate range (0.065–0.076). All three items exhibited significant pre-post DIF, indicating response shifts that prevent unadjusted comparisons over time. Notably, when analyzing control and treatment groups separately, DIF was only significant in the treatment group for one item, suggesting potential intervention effects. In pre-only analyses, the factor structure and item performance remained

consistent, though the standout item still showed high (but slightly reduced) discrimination. No DIF was detected by gender or first-generation status. However, DIF was observed across several race/ethnicity comparisons, particularly affecting two items ("You can learn new things, but you can't really change your basic intelligence/how smart you are", "You have a certain amount of intelligence, and you really can't do much to change it."), highlighting potential cultural or interpretive differences. Despite these issues, the scale demonstrated excellent internal consistency (α = .91), and with DIF-adjusted scoring, it remains a robust tool for measuring fixed mindset.

Learning Strategies

The full results of psychometric property for the Learning Strategies measure are presented in Appendix C. EFA on both combined and pre-only datasets supported a unidimensional structure. IRT analysis showed most items had moderate to high discrimination and a desirable spread of difficulty values, effectively capturing a range of student behaviors. Items like "I go over my lecture notes as soon as possible after lecture to rework them and note problem areas." and "I rework all of the quiz and test items that I have missed before the next class session." had high discrimination, while items such as "I regularly go to office hours or tutoring to discuss problems or questions about the homework." were easier to endorse than more effortful strategies like using mnemonics or studying well in advance. While a few items showed significant S-X² misfit, RMSEA values remained low (≤ 0.054), suggesting minimal practical concern. Pre-post DIF analysis revealed no significant DIF across time or treatment indicators, indicating strong temporal invariance. In pre-only DIF analyses, one item showed DIF by first-generation status ("I spend some time studying for this class at least five days per week (outside of class).") and another by gender ("I try to do my homework without using example problems as a guide or copying answers from my class notes or textbook."), the latter reflecting potential variation in how students interpret independent learning strategies. The most notable findings emerged from ethnicity-based DIF, where 7 of 11 items showed DIF across comparisons involving Black/African American, Asian, Latinx, and Multiracial students. This suggests possible cultural or contextual differences in strategy use or interpretation. Despite these concerns, the scale showed solid internal consistency (α = .77), supporting its use with DIF adjustment where necessary.

PERTS: Self-Efficacy and Sense of Belonging

The full results of psychometric property for the PERTS measure are presented in Appendix D. EFA on both the combined and pre-only datasets consistently supported a two-factor structure. IRT modeling confirmed strong psychometric properties: all items had very high discrimination (all > 3), indicating they sharply differentiate between students at different trait levels. However, several items had very high thresholds, particularly in the upper categories, suggesting these responses were rarely selected and possibly warrant response category revision. Despite all items showing significant statistical misfit, RMSEA values ≤ 0.060 indicated acceptable practical fit. In

pre-post DIF analysis, Sense of Belonging items were invariant across time and study conditions. However, one or more Self-Efficacy items showed DIF in both treatment and control groups, particularly those related to mastering hard topics and succeeding on difficult tests. In pre-only data, no items showed DIF by first-generation status or gender, and only one belonging item ("I feel like I belong in this class") showed DIF by race/ethnicity. Self-Efficacy items showed no racial DIF, supporting cross-group comparability. Both subscales demonstrated excellent reliability: α = .93 for Self-Efficacy and α = .88 for Belonging. Overall, the scale shows strong psychometric performance with minimal DIF concerns.

Online Self-Regulated Learning Questionnaire (OSLQ)

The OSLQ measures Goal Setting, Time Management, and Help-Seeking through 13 items rated on a 5-point Likert scale. The full results of psychometric property for the OSLQ measure are presented in Appendix E. EFA using both combined and pre-only datasets supported a three-factor structure, consistent with the theoretical model. While parallel analysis initially suggested a two-factor solution, eigenvalues and IRT model comparisons confirmed the appropriateness of three dimensions. The three-dimensional IRT model showed strong item discrimination and acceptable model fit (RMSEA ≤ 0.060), despite statistically significant S-X² p-values likely due to the large sample size (>2,500). Pre-post DIF analysis showed that Goal Setting and Help-Seeking items were invariant over time, while one Time Management item—related to scheduling regular study times—was flagged for DIF. However, this item did not show DIF when treatment and control groups were analyzed separately, suggesting the result may be an artifact of group mixing. In pre-only DIF analysis, one Help-Seeking item showed DIF by first-generation status ("If needed, I try to meet with classmates face-to-face."), and one Goal Setting item was flagged by gender ("I don't compromise quality of my work because it is online"), possibly reflecting varied interpretations of academic standards. The most DIF was observed by race/ethnicity: 8 of 13 items showed DIF across subscales, with Multiracial students appearing most frequently in flagged comparisons. This may reflect cultural or contextual differences in how self-regulation behaviors are understood or practiced. Despite these findings, the OSLQ showed strong psychometric performance. Its dimensional structure is robust, items demonstrate good sensitivity, and internal consistency was high across subscales. While race-related DIF warrants caution, the measure remains a valuable tool for assessing self-regulated learning—especially when using DIF-adjusted scores for cross-group comparisons.

Metacognitive Awareness Inventory (MAI)

The MAI includes 18 binary items measuring three constructs: Comprehension Monitoring, Debugging Strategies, and Evaluation. The full results of psychometric property for the OSLQ measure are presented in Appendix F. EFA on both combined and pre-only datasets supported a three-factor structure, though Monitoring and Evaluation items often loaded together. A third factor emerged reflecting strategic evaluation. However, IRT model comparisons favored a

one-dimensional structure, suggesting the items function collectively as a general metacognitive trait. Despite better theoretical alignment, the three-dimensional model showed greater item misfit and possible overfitting, especially for skewed items like "I stop and reread when I get confused," which showed very high discrimination and difficulty. Pre-post DIF analysis indicated temporal invariance, with no items flagged after FDR correction in the full sample. In subgroup analyses, DIF was limited: two items in the treatment group ("I ask others for help when I don't understand something", and I know how well I did once I finish a test.") and one in the control group showed DIF ("I ask myself if I have considered all options after I solve a problem"). In pre-only data, model fit patterns held, with the one-dimensional model again performing best. No DIF was found by first-generation status, and only one item showed DIF by gender under the one-dimensional model ("I know how well I did once I finish a test"). However, race/ethnicity-based DIF was notable: 12 unique items were flagged across groups, with comparisons involving Asian and Multiracial students most frequently affected. One item—"I pause regularly to evaluate my understanding"—was flagged in six comparisons, suggesting potential cultural differences in interpreting reflection-related items. Reliability was strong for the full scale (α = .80) but varied by subscale: Comprehension Monitoring (α = .71), Evaluation (α = .66), and Debugging showed lower consistency (α = .42–.55), indicating the need for refinement in that domain. In summary, the MAI demonstrates strong overall reliability and stability across time and most demographic groups, though ethnic DIF and weaker subscale reliability—especially in Debugging—warrant attention for future refinement or DIF-adjusted scoring in subgroup comparisons.

SUMMARY OF KEY RESULTS AND TRENDS

- **Strong structural validity across scales:** EFA and unidimensional and multidimensional IRT modeling generally supported the intended factor structures of all scales. Most measures conformed well to theoretical expectations, with fixed or multi-factor solutions aligning with original constructs. Some exceptions emerged (e.g., the MAI's overlap between Comprehension Monitoring and Evaluation), though item performance largely supported coherent latent structures.

- **High internal consistency and strong item discrimination:** Reliability was acceptable to excellent across scales, with most α values exceeding .75. Items frequently demonstrated moderate to high discrimination, with some items (especially in the Fixed Mindset and PERTS Self-Efficacy subscales) showing exceptionally high values. This indicates strong sensitivity to individual differences, though extreme discrimination in some cases (e.g., a fixed mindset item) may reflect local dependence or item dominance concerns.

- **Response category issues in polytomous scales:** Several Likert-based items showed extreme thresholds or sparsely used response categories, particularly at the upper end (e.g., PERTS Self-efficacy). This suggests ceiling effects or possible issues with the scale range, warranting consideration of response option revisions to better capture variation at higher trait levels.

- **Item misfit was common but typically not particularly concerning:** While item-level misfit (e.g., significant S-X² values) was statistically frequent—likely due to large sample sizes— global fit indices (RMSEA) were generally within acceptable or good ranges (≤ .06). This suggests practical model fit was adequate even where item-level statistics flagged issues.

- **Temporal invariance was mixed:** Pre-post DIF varied across measures. Some scales (e.g., Learning Strategies, MAI) showed strong temporal invariance, while others (e.g., Fixed Mindset, PERTS Self-Efficacy) exhibited response shifts over time, particularly in treatment groups. These findings suggest that some constructs may be especially sensitive to intervention effects, making unadjusted DIF pre-post comparisons problematic.

- **Minimal DIF by gender and first-generation status:** Across all measures, relatively few items showed differential item functioning by gender or first-generation status. Where present, gender-based DIF typically related to interpretation of academic behaviors (e.g., help-seeking, independent study), and first-generation DIF centered on effort or engagement strategies. These instances were isolated and not consistent across scales.

- **Race/ethnicity-based DIF was pervasive and patterned:** DIF by race/ethnicity was observed across nearly all scales, often affecting a substantial number of items. Items assessing behaviors or beliefs (e.g., learning strategies, self-regulation) were particularly likely to show DIF. Multiracial and Asian students appeared most frequently in flagged comparisons, and some items (e.g., those involving reflection, academic confidence, or effort norms) were flagged repeatedly. These patterns suggest cultural or contextual influences on interpretation or endorsement of items, and raise concerns about cross-group comparability if DIF is not addressed.

- **Continued utility with DIF adjustment:** Despite observed DIF and some subscale reliability concerns (e.g., MAI Debugging), all measures retain utility—especially when scoring accounts for DIF. Each scale demonstrated clear psychometric strengths and produced valid scores suitable for both cross-sectional and longitudinal analysis, when interpreted carefully.

REFERENCES

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education, 12*(1), 1–6. https://doi.org/10.1016/j.iheduc.2008.10.005

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*, 1-29.

Dweck, C. S., Chiu, C.-Y., & Hong, Y.-Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry, 6*(4), 267–285. https://doi.org/10.1207/s15327965pli0604_1

Hanson, J. (2017). Determination and validation of the Project for Educational Research That Scales (PERTS) survey factor structure. *Journal of Educational Issues, 3*(1), 64–82. https://doi.org/10.5296/jei.v3i1.10646

Harrison, G. M., & Vallin, L. M. (2018). Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition and Learning, 13*(1), 15-38.

Jackson, C. R. (2018). Validating and adapting the motivated strategies for learning questionnaire (MSLQ) for STEM courses at an HBCU. *AERA Open, 4*(4), 2332858418809346.

McGuire, S., McGuire, S. Y., & Angelo, T. (2015). *Teach students how to learn: Strategies you can incorporate into any course to improve student metacognition, study skills, and motivation.* Routledge.

Midkiff, M. F., Lindsey, C. R., & Meadows, E. A. (2018). The role of coping self-efficacy in emotion regulation and frequency of NSSI in young adult college students. *Cogent Psychology, 5*(1), 1520437.

Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement, 24*(1), 50-64.

Quan, Y., & Wang, C. (2025). Collapsing or not? A practical guide to handling sparse responses for polytomous items. *Methodology, 21*(1), 46-73.

Roth, A., Ogrin, S., & Schmitz, B. (2016). Assessing self-regulated learning in higher education: A systematic literature review of self-report instruments. *Educational Assessment, Evaluation and Accountability, 28*(3), 225-250.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Tsai, C. L., Wind, S., & Estrada, S. (2025). Exploring the effects of collapsing rating scale categories in polytomous item response theory analyses: An illustration and simulation study. *Measurement: Interdisciplinary Research and Perspectives, 23*(1), 66-89.

APPENDICES

Appendix A. Participant demographic characteristics

Notes: Percentages shown are valid percentages, calculated using non-missing responses for each variable, stratified by measure and study. Because the number of valid responses differs by variable, measure, and study, denominators vary accordingly.

Variable		By Measure					By Study		
		Fixed Mindset	Learning Strategies	MAI	OSLQ	PERTS	IPP	RCE	Total
First Generation Student		413 (34%)	410 (34%)	547 (37%)	426 (34%)	551 (37%)	118 (54%)	433 (34%)	551 (37%)
Gender	Female	938 (64%)	952 (65%)	1099 (62%)	959 (62%)	1113 (62%)	154 (63%)	959 (62%)	1113 (62%)
	Male	451 (31%)	451 (31%)	518 (29%)	457 (30%)	524 (29%)	67 (27%)	457 (30%)	524 (29%)
	Prefer not answer/ Unlisted	83 (6%)	61 (4%)	153 (9%)	126 (8%)	153 (9%)	25 (10%)	129 (8%)	154 (9%)
Race/ethnicity	Asian	106 (7%)	108 (7%)	120 (7%)	109 (7%)	123 (7%)	14 (6%)	109 (7%)	123 (7%)
	Black or African American	317 (22%)	318 (22%)	361 (21%)	324 (21%)	366 (21%)	42 (19%)	324 (21%)	366 (21%)
	Latinx or Hispanic	324 (22%)	325 (22%)	360 (21%)	329 (21%)	363 (21%)	34 (15%)	329 (21%)	363 (21%)
	Multiracial	104 (7%)	104 (7%)	133 (8%)	104 (7%)	135 (8%)	30 (13%)	105 (7%)	135 (8%)
	Prefer not answer/ Unlisted	135 (9%)	116 (8%)	200 (11%)	172 (11%)	200 (11%)	18 (8%)	182 (12%)	200 (11%)
	White	486 (33%)	493 (34%)	575 (33%)	496 (32%)	583 (33%)	87 (39%)	496 (32%)	583 (33%)
Time Point Survey Taken	Pre Intervention Only	562 (38%)	562 (38%)	813 (46%)	572 (38%)	815 (46%)	246 (100%)	570 (38%)	816 (46%)
	Post Intervention Only	155 (11%)	155 (11%)	187 (11%)	182 (12%)	181 (10%)	0 (0%)	181 (12%)	181 (10%)
	Both Pre and Post Intervention	758 (51%)	758 (51%)	749 (43%)	762 (50%)	765 (43%)	0 (0%)	765 (50%)	765 (43%)

Appendix B. Psychometric Properties of <u>Fixed Mindset</u> Measure

Exploratory Factor Analysis

Item	Factor Loading	
	Combined Pre-post data	Pretest data only
You can learn new things, but you can't really change your basic intelligence/how smart you are.	.826	.796
Your intelligence is something about you that you can't change very much.	.962	.959
You have a certain amount of intelligence, and you really can't do much to change it.	.899	.880
Eigenvalue	2.42	2.32
Total Variance	80.5%	77.6%

Dimensionality and Item Statistics

Item	Combined Pre-post data						Pretest data only						
	a	d1-d5	S-X²	df	RMSEA	p-value	a	d1-d5	S-X²	df	RMSEA	p-value	Alpha if Item deleted
You can learn new things, but you can't really change your basic intelligence/how smart you are.	3.31	5.73, 3.50, 1.68, -0.01, -3.16	224.65	16	.076	< .001	2.94	5.52, 3.37, 1.68, 0.05, -2.92	136.34	16	.075	< .001	0.91
Your intelligence is something about you that you can't change very much.	28.30	45.85, 30.28, 20.20, 6.99, -17.00	83.87	8	.065	< .001	15.22	26.57, 18.07, 12.15, 4.44, -8.82	27.46	8	.043	.001	0.82
You have a certain amount of intelligence, and you really can't do much to change it.	4.78	8.67, 6.06, 4.06, 1.91, -2.53	158.69	13	.070	< .001	4.35	8.68, 6.15, 4.24, 2.14, -2.18	86.49	11	.071	< .001	0.86

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

11

Differential Item Functioning

Notes: Only p-values adjusted for False Discovery Rate (FDR) are reported. A dash (–) indicates p > .05. For Race/Ethnicity, a "Prefer not to answer/Unlisted" category was included in the DIF analysis; however, the results for this category are not reported in this technical report.

Item	Pre-post	First Generation Status	Gender	Race/Ethnicity
You can learn new things, but you can't really change your basic intelligence/how smart you are.	.021	.946	.676	White VS Black/African American < .001 White Vs and Multiracial .009
Your intelligence is something about you that you can't change very much.	.040	.946	.440	–
You have a certain amount of intelligence, and you really can't do much to change it.	.015	.946	.440	Asian vs Black/African American .002 White VS Black/African American < .001

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

12

Appendix C. Psychometric Properties of <u>Learning Strategies</u> Measure

Exploratory Factor Analysis

Item	Factor Loading	
	Combined Pre-post data	**Pretest data only**
I realize that I can still do well in this class even if I have done poorly on the quizzes and tests up to this point.	.224	.203
I make diagrams or draw mental pictures of the concepts discussed in class.	.439	.420
I try to do my homework without using example problems as a guide or copying answers from my class notes or textbook.	.332	.302
I make mnemonics for myself to help me remember facts and equations.	.493	.479
I regularly go to office hours or tutoring to discuss problems or questions about the homework.	.576	.582
I always preview the material that will be discussed before I go to class.	.488	.486
I go over my lecture notes as soon as possible after lecture to rework them and note problem areas.	.598	.608
I rework all of the homework problems and questions before the test or quiz.	.530	.528
I rework all of the quiz and test items that I have missed before the next class session.	.559	.561
I spend some time studying for this class at least five days per week (outside of class).	.563	.552
I participate in a study group where we do homework and quiz ourselves on the material.	.580	.583
Eigenvalue	3.49	4.43
Total Variance	25.2%	24.7%

Dimensionality and Item Statistics

Item	Combined Pre-post data						Pretest data only						Alpha if Item deleted
	a	d	S-X²	df	RMSEA	p-value	a	d	S-X²	df	RMSEA	p-value	
I realize that I can still do well in this class even if I have done poorly on the quizzes and tests up to this point.	0.658	1.692	6.03	8	.000	.644	0.591	1.684	3.717	8	.000	.882	0.78
I make diagrams or draw mental pictures of the concepts discussed in class.	1.212	1.042	8.14	8	.003	.420	1.101	0.925	4.830	8	.000	.776	0.76
I try to do my homework without using example problems as a guide or copying answers from my class notes or textbook.	0.861	1.034	4.14	8	.000	.844	0.747	0.881	8.131	8	.004	.421	0.77
I make mnemonics for myself to help me remember facts and equations.	1.406	0.835	7.01	8	.000	.535	1.317	0.708	8.686	8	.008	.370	0.75
I regularly go to office hours or tutoring to discuss problems or questions about the homework.	1.937	-0.991	18.73	7	.027	.009	2.026	-1.228	12.047	6	.028	.061	0.74
I always preview the material that will be discussed before I go to class.	1.405	1.013	2.83	8	.000	.944	1.393	0.989	10.639	8	.016	.223	0.75
I go over my lecture notes as soon as possible after lecture to rework them and note problem areas.	2.193	1.433	28.71	8	.034	.000	2.236	1.264	16.923	7	.033	.018	0.74
I rework all of the homework problems and questions before the test or quiz.	1.805	1.561	20.72	8	.027	.008	1.747	1.411	14.341	8	.024	.073	0.75
I rework all of the quiz and test items that I have missed before the next class session.	1.873	1.303	6.02	8	.000	.645	1.886	1.271	7.579	7	.008	.371	0.74
I spend some time studying for this class at least five days per week (outside of class).	1.709	0.665	10.51	8	.012	.231	1.610	0.531	2.516	7	.000	.926	0.74
I participate in a study group where we do homework and quiz ourselves on the material.	2.137	-1.339	44.98	6	.054	.000	2.268	-1.692	28.518	6	.053	.000	0.74

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

14

Differential Item Functioning

Notes: Only p-values adjusted for False Discovery Rate (FDR) are reported. A dash (–) indicates p > .05. For Race/Ethnicity, a "Prefer not to answer/Unlisted" category was included in the DIF analysis; however, the results for this category are not reported in this technical report.

Item	Pre-post	First Generation Status	Gender	Race/Ethnicity
I realize that I can still do well in this class even if I have done poorly on the quizzes and tests up to this point.	.500	–	–	White VS Black/African American .008
I make diagrams or draw mental pictures of the concepts discussed in class.	.892	–	–	–
I try to do my homework without using example problems as a guide or copying answers from my class notes or textbook.	.321	–	.004	Asian VS White .006 Asian VS Multiracial .022
I make mnemonics for myself to help me remember facts and equations.	.982	–	–	–
I regularly go to office hours or tutoring to discuss problems or questions about the homework.	.406	–	–	–
I always preview the material that will be discussed before I go to class.	.382	–	–	–
I go over my lecture notes as soon as possible after lecture to rework them and note problem areas.	.500	–	–	White VS Black/African American .008 White VS Latinx .004 White VS Multiracial .015
I rework all of the homework problems and questions before the test or quiz.	.982	–	–	White VS Hispanic .031 White VS Multiracial .015 Asian VS Multiracial .011
I rework all of the quiz and test items that I have missed before the next class session.	.321	–	–	Multiracial vs Asian .022 Multiracial vs White .015
I spend some time studying for this class at least five days per week (outside of class).	.982	.006	–	White VS Black/African American .018 White VS Multiracial .029
I participate in a study group where we do homework and quiz ourselves on the material.	.321	–	–	–

Appendix D. Psychometric Properties of PERTS Measure

Exploratory Factor Analysis

Item	Factor Loading			
	Combined Pre-post data		Pretest data only	
	Factor 1	Factor 2	Factor 1	Factor 2
I can earn an A or top grade in this class		.854		.857
I can meet all the learning goals my teacher set		.767		.781
I can master the hardest topics in my class		.874		.834
I can do well on tests even when they're difficult		.829		.768
I feel like I belong in this class.	.737		.722	
I feel like I can be myself in this class	.921		.908	
I feel comfortable in this class	.916		.896	
I feel respected in this class	.939		.924	
Variance	39%	35%	37.7%	33.4%

Dimensionality and Item Statistics

Data	Model	Log-likelihood	Estimated parameters	AIC	BIC	SABIC
Combined Pre and Post	One-dimensional Model	-19430.09	40	38940.18	39173.29	39046.20
	Two-dimensional Model	-17886.55	40	35853.11	36086.21	35959.12
Pre-only	One-dimensional Model	-12404.79	40	24889.57	25104.26	24977.19
	Two-dimensional Model	-11487.54	40	23055.07	23269.76	23142.69

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

16

Item	Combined Pre-post data							Pretest data only							Alpha if Item deleted
	a1	a2	d1-d4	S-X²	df	RMSEA	p-value	a1	a2	d1-d4	S-X²	df	RMSEA	p-value	
I can earn an A or top grade in this class	3.68		7.07, 4.94, 0.34, -3.10	376.07	66	.043	<.001	3.63		7.31, 5.19, 0.49, -3.15	202.08	30	.060	<.001	0.84
I can meet all the learning goals my teacher set	3.27		7.83, 5.90, 1.28, -2.49	418.61	67	.046	<.001	3.34		8.52, 6.21, 1.48, -2.55	235.10	54	.046	<.001	0.85
I can master the hardest topics in my class	3.86		7.93, 5.37, -0.73, -4.50	360.60	63	.043	<.001	3.37		7.36, 4.82, -0.76, -4.19	201.11	57	.040	<.001	0.84
I can do well on tests even when they're difficult	3.10		6.61, 4.59, -0.09, -3.58	341.97	65	.041	<.001	2.54		5.88, 4.11, -0.16, -3.33	217.84	58	.042	<.001	0.86
I feel like I belong in this class.		3.19	6.92, 5.70, 2.19, -1.38	456.90	63	.050	<.001		3.01	6.84, 5.61, 2.02, -1.47	237.05	55	.046	<.001	0.92
I feel like I can be myself in this class		5.03	10.40, 9.11, 4.21, -1.04	324.02	60	.042	<.001		4.82	10.18, 8.79, 3.91, -1.31	164.44	44	.042	<.001	0.90
I feel comfortable in this class		6.66	13.01, 11.01, 5.56, -0.81	433.69	59	.050	<.001		5.91	11.71, 10.06, 4.84, -1.05	201.00	48	.045	<.001	0.89
I feel respected in this class		4.93	10.49, 9.63, 5.25, -0.28	288.49	57	.040	<.001		4.73	10.43, 9.61, 4.96, -0.44	153.99	40	.042	<.001	0.90

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

17

Differential Item Functioning

Notes: Only p-values adjusted for False Discovery Rate (FDR) are reported. A dash (–) indicates p > .05. For Race/Ethnicity, a "Prefer not to answer/Unlisted" category was included in the DIF analysis; however, the results for this category are not reported in this technical report.

Item	Pre-post	First Generation Status	Gender	Race/Ethnicity
I can earn an A or top grade in this class	–	–	–	–
I can meet all the learning goals my teacher set	–	–	–	–
I can master the hardest topics in my class	< .001	–	–	–
I can do well on tests even when they're difficult	< .001	–	–	–
I feel like I belong in this class.	–	–	–	–
I feel like I can be myself in this class	–	–	–	Asian VS Black/African American .042
I feel comfortable in this class	–	–	–	–
I feel respected in this class	–	–	–	–

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

18

Appendix E. Psychometric Properties of <u>OSLQ</u> Measure

Exploratory Factor Analysis

Item	Factor Loading					
	Combined Pre-post data			Pretest data only		
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
I find someone who is knowledgeable to consult with when I need help		.631			.625	
If needed, I try to meet with classmates face-to-face		.794			.751	
I am persistent in getting help from the instructor through email		.360			.336	
I share my problems with my classmates online so we know what we are struggling with and how to solve our problems.		.898			.901	
Although we don't have to attend daily classes, I still try to distribute my studying time evenly across days.			.925			.909
I allocate extra studying time for my online courses because I know it is time-demanding.			.537			.561
I try to schedule the same time every day or every week to study for my online courses, and I observe the schedule.			.866			.838
Although we don't have to attend daily classes, I still try to study for my online courses across days.			.540			.598
I don't compromise quality of my work because it is online	.562			.493		
I keep a high standard for my learning in my online courses	.930			.934		
I set goals to help me manage studying time for my online courses	.726			.647		
I set standards for my assignments in online courses	.830			.806		
I set short-term (daily/weekly) goals and as long-term goals (monthly/semester)	.774			.694		
Variance	24.2%	17.2%	15.2%	21.3%	17.6%	14.6%

Dimensionality and Item Statistics

Data	Model	Log-likelihood	Estimated parameters	AIC	BIC	SABIC
Combined Pre and Post	One-dimensional Model	-33966.78	65	68063.56	68442.2	68235.68
	Three-dimensional Model	-33318.12	65	66766.23	67144.87	66938.35
Pre-only	One-dimensional Model	-21210.96	65	42551.92	42900.62	42694.13
	Three-dimensional Model	-20865.45	65	41860.89	42209.59	42003.10

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

19

Item	Combined Pre-post data								Pretest data only								Alpha if Item deleted
	a1	a2	a3	d1-d4	S-X²	df	RMSEA	p-value	a1	a2	a3	d1-d4	S-X²	df	RMSEA	p-value	
I find someone who is knowledgeable to consult with when I need help		1.81		3.73, 1.88, 0.19, -2.35	301.65	99	.030	<.001		1.69		3.59, 1.79, 0.10, -2.35	204.75	93	.030	<.001	.70
If needed, I try to meet with classmates face-to-face		2.53		2.65, 0.11, -2.00, -4.40	503.84	88	.046	<.001		2.26		2.35, -0.02, -2.07, -4.40	306.86	86	.044	<.001	.68
I am persistent in getting help from the instructor through email		1.09		3.58, 2.07, 0.32, -1.74	244.59	105	.024	<.001		1.05		3.51, 2.02, 0.30, -1.80	160.92	91	.024	<.001	.77
I share my problems with my classmates online so we know what we are struggling with and how to solve our problems.		3.75		4.47, 0.81, -1.85, -5.24	477.55	89	.044	<.001		3.49		4.10, 0.49, -2.23, -5.35	285.31	82	.043	<.001	.64
Although we don't have to attend daily classes, I still try to distribute my studying time evenly across days.			3.95	8.16, 4.81, 2.60, -2.21	450.48	95	.041	<.001			3.43	7.64, 4.44, 2.42, -1.95	248.59	78	.041	<.001	.74
I allocate extra studying time for my online courses because I know it is time-demanding.		2.20		5.77, 3.95, 1.81, -1.62	335.91	94	.034	<.001		1.96		5.39, 3.74, 1.70, -1.58	208.12	82	.034	<.001	.78

Item	Combined Pre-post data								Pretest data only								
	a1	a2	a3	d1-d4	S-X²	df	RMSEA	p-value	a1	a2	a3	d1-d4	S-X²	df	RMSEA	p-value	Alpha if Item deleted
I try to schedule the same time every day or every week to study for my online courses, and I observe the schedule.			5.44	11.64, 7.60, 4.72, -2.58	414.94	92	.039	<.001			4.97	11.37, 7.22, 4.54, -2.52	228.61	72	.040	<.001	.73
Although we don't have to attend daily classes, I still try to study for my online courses across days.			1.99	5.08, 2.77, 1.26, -1.80	319.04	101	.031	<.001			1.86	4.87, 2.52, 1.17, -1.83	191.87	85	.031	<.001	.80
I don't compromise quality of my work because it is online	1.37			4.24, 2.78, 1.48, -0.89	422.81	105	.037	<.001	1.23			4.05, 2.57, 1.35, -0.87	211.35	87	.033	<.001	.87
I keep a high standard for my learning in my online courses	4.12			11.92, 8.59, 4.71, -1.63	284.10	71	.037	<.001	4.15			11.89, 8.25, 4.75, -1.82	171.75	67	.034	<.001	.75
I set goals to help me manage studying time for my online courses	3.31			9.20, 6.45, 3.85, -1.82	420.79	77	.045	<.001	3.10			8.92, 5.77, 3.54, -1.80	239.06	66	.044	<.001	.76
I set standards for my assignments in online courses	2.93			8.84, 6.92, 4.17, -1.36	285.21	76	.035	<.001	2.82			8.78, 6.69, 3.99, -1.45	169.07	64	.035	<.001	.77
I set short-term (daily/weekly) goals and as long-term goals (monthly/semester)	2.79			7.76, 5.42, 3.04, -1.75	323.27	85	.035	<.001	2.63			7.41, 4.93, 2.81, -1.74	178.85	70	.034	<.001	.77

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

21

Differential Item Functioning

Notes: Only p-values adjusted for False Discovery Rate (FDR) are reported. A dash (–) indicates p > .05. For Race/Ethnicity, a "Prefer not to answer/Unlisted" category was included in the DIF analysis; however, the results for this category are not reported in this technical report.

Item	Pre-post	First Generation Status	Gender	Race/Ethnicity
I find someone who is knowledgeable to consult with when I need help	–	–	–	–
If needed, I try to meet with classmates face-to-face	–	.035	–	Latinx/Hispanic VS Multiracial .005
I am persistent in getting help from the instructor through email	–	–	–	–
I share my problems with my classmates online so we know what we are struggling with and how to solve our problems.	–	–	–	–
Although we don't have to attend daily classes, I still try to distribute my studying time evenly across days.	–	–	–	Black/African American VS Latinx/Hispanic .050
I allocate extra studying time for my online courses because I know it is time-demanding.	–	–	–	–
I try to schedule the same time every day or every week to study for my online courses, and I observe the schedule.	.002	–	–	Black/African American VS White .014 Latinx/Hispanic VS White.020
Although we don't have to attend daily classes, I still try to study for my online courses across days.	–	–	.001	–
I don't compromise quality of my work because it is online	–	–	–	Asian VS Multiracial .011 Black/African American VS Multiracial .003 Black/African American VS White .005 Latinx/Hispanic VS Multiracial .002 Latinx/Hispanic VS White <.001
I keep a high standard for my learning in my online courses	–	–	–	–
I set goals to help me manage studying time for my online courses	–	–	–	Black/African American VS Multiracial .001 Latinx/Hispanic VS Multiracial .005
I set standards for my assignments in online courses	–	–	–	–
I set short-term (daily/weekly) goals and as long-term goals (monthly/semester)	–	–	–	–

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

22

Appendix F. Psychometric Properties of <u>MAI</u> Measure

Exploratory Factor Analysis

Item	Factor Loading					
	Combined Pre-post data			**Pretest data only**		
	Factor 1	**Factor 2**	**Factor 3**	**Factor 1**	**Factor 2**	**Factor 3**
I ask myself if I have considered all options when solving a problem			.749			.843
I find myself analyzing the usefulness of strategies while I study	.489			.485		
I ask myself periodically if I am meeting my goals	.562			.552		
I ask myself questions about how well I am doing while learning something new	.668			.637		
I consider several alternatives to a problem before I answer			.590			.536
I find myself pausing regularly to check my comprehension	.525			.433		
I periodically review to help me understand important relationships	.483			.497		
I ask others for help when I don't understand something	.222			.217		
I change strategies when I fail to understand		.262			.266	
I re-evaluate my assumptions when I get confused		.385			.326	
I stop and reread when I get confused		.791			.757	
I stop and go back over new information that is not clear		.745			.711	
I ask myself if I have considered all options after I solve a problem	.229		.513	.375		.350
I ask myself how well I accomplish my goals once I'm finished	.656			.713		
I ask myself if there was an easier way to do things after I finish a task	.210			.243		
I know how well I did once I finish a test	.175			.231		
I ask myself if I learned as much as I could have once I finish a task	.565			.617		
I summarize what I've learned after I finish	.599			.649		
Variance	15.7%	8%	6.8%	17%	7.2%	6.4%

Dimensionality and Item Statistics

Data	Model	Log-likelihood	Estimated parameters	AIC	BIC	SABIC
Combined Pre and Post	One-dimensional Model	-14720.79	36	29513.58	29723.1	29608.72
	EFA-based Three-dimensional Model	-14967.77	36	30007.53	30217.05	30102.67
	Theory-based Three-dimensional Model	-15324.34	36	30720.67	30930.2	30815.82
Pre-only	One-dimensional Model	-9435.345	36	18942.69	19135.52	19021.15
	EFA-based Three-dimensional Model	-9641.307	36	19354.61	19547.44	19433.08
	Theory-based Three-dimensional Model	--9842.26	36	19756.52	19949.35	19834.98

One-Dimensional Model

Item	Combined Pre-post data						Pretest data only						
	a	d	S-X²	df	RMSEA	p-value	a	d	S-X²	df	RMSEA	p-value	Alpha if Item deleted
I ask myself if I have considered all options when solving a problem	2.07	3.23	20.29	11	0.018	0.042	2.101	3.12	15.31	10	.018	.121	.68
I find myself analyzing the usefulness of strategies while I study	1.825	2.55	11.18	10	0.007	0.344	1.77	2.42	9.36	10	.000	.498	.68
I ask myself periodically if I am meeting my goals	2.02	2.71	15.08	10	.014	.129	2.12	2.64	6.97	10	.000	.728	.79
I ask myself questions about how well I am doing while learning something new	2.38	3.22	12.67	10	.010	.243	2.18	2.96	9.55	10	.000	.481	.79
I consider several alternatives to a problem before I answer	1.77	3.00	14.38	11	.011	.213	1.77	2.97	7.27	11	.000	.777	.78
I find myself pausing regularly to check my comprehension	1.77	2.72	8.34	11	.000	.682	1.64	2.65	13.05	11	.011	.290	.78
I periodically review to help me understand important relationships	2.00	3.02	15.71	10	.015	.108	1.96	3.02	11.77	10	.011	.300	.79
I ask others for help when I don't understand something	0.78	1.45	25.97	12	.022	.011	0.73	1.46	22.14	12	.023	.036	.79
I change strategies when I fail to understand	1.29	3.46	28.59	12	.024	.005	1.22	3.30	17.88	12	.018	.119	.79
I re-evaluate my assumptions when I get confused	1.70	4.17	21.34	12	.018	.046	1.58	4.09	21.42	12	.022	.045	.80

Item	Combined Pre-post data						Pretest data only						
	a	d	S-X²	df	RMSEA	p-value	a	d	S-X²	df	RMSEA	p-value	Alpha if Item deleted
I stop and reread when I get confused	1.13	4.28	14.21	11	.011	.222	1.12	4.52	11.96	10	.011	.288	.79
I stop and go back over new information that is not clear	1.38	4.49	24.33	11	.022	.011	1.36	4.68	12.98	10	.014	.225	.79
I ask myself if I have considered all options after I solve a problem	2.33	2.89	10.20	10	.003	.423	2.37	2.75	7.86	9	.000	.549	.80
I ask myself how well I accomplish my goals once I'm finished	2.42	2.75	17.88	9	.020	.037	2.43	2.62	17.28	9	.024	.044	.80
I ask myself if there was an easier way to do things after I finish a task	0.93	1.89	11.98	12	.000	.447	1.00	1.82	5.13	11	.000	.925	.78
I know how well I did once I finish a test	0.49	1.20	25.56	12	.021	.012	0.48	1.09	19.87	12	.020	.070	.78
I ask myself if I learned as much as I could have once I finish a task	2.23	2.51	15.69	10	.015	.109	2.34	2.50	6.57	9	.000	.682	.80
I summarize what I've learned after I finish	2.01	1.72	11.67	9	.011	.233	2.04	1.61	10.70	9	.011	.297	.80

Theory-based Three-dimensional Model

Item	Combined Pre-post data								Pretest data only								
	a1	a2	a3	d	S-X²	df	RMSEA	p-value	a1	a2	a3	d	S-X²	df	RMSEA	p-value	Alpha if Item deleted
I ask myself if I have considered all options when solving a problem	1.80			2.99	32.27	11	.028	.001	1.96			2.99	22.17	10	.028	.014	.68
I find myself analyzing the usefulness of strategies while I study	1.83			2.54	18.85	11	.017	.064	1.80			2.43	16.20	10	.020	.094	.68
I ask myself periodically if I am meeting my goals	2.00			2.67	24.53	10	.024	.006	2.12			2.62	16.74	10	.021	.080	.67
I ask myself questions about how well I am doing while learning something new	2.27			3.11	27.31	10	.026	.002	1.96			2.76	18.80	10	.024	.043	.68
I consider several alternatives to a problem before I answer	1.60			2.87	31.09	11	.027	.001	1.76			2.96	20.14	11	.023	.043	.69

Technical Report: Psychometric Properties of Self-Directed Learning Survey Measures in Broad Access Institution Contexts

25

Item	Combined Pre-post data								Pretest data only								Alpha if Item deleted
	a1	a2	a3	d	S-X²	df	RMSEA	p-value	a1	a2	a3	d	S-X²	df	RMSEA	p-value	
I find myself pausing regularly to check my comprehension	1.85			2.78	15.29	11	.013	.169	1.72			2.71	16.19	11	.017	.134	.69
I periodically review to help me understand important relationships	1.87			2.90	36.93	11	.031	<.001	1.84			2.91	27.54	11	.031	.004	.69
I ask others for help when I don't understand something		0.70		1.42	35.46	12	.028	<.001		0.62		1.42	28.53	11	.032	.003	.60
I change strategies when I fail to understand		1.71		3.91	38.60	12	.030	<.001		1.88		3.98	27.58	12	.029	.006	.50
I re-evaluate my assumptions when I get confused		2.42		5.15	33.89	12	.027	.001		2.27		4.99	31.57	12	.032	.002	.48
I stop and reread when I get confused		5.07		10.42	32.65	12	.026	.001		3.31		7.79	20.43	11	.023	.040	.45
I stop and go back over new information that is not clear		4.76		9.85	29.49	12	.024	.003		3.14		7.39	19.06	11	.022	.060	.45
I ask myself if I have considered all options after I solve a problem			2.34	2.83	29.72	11	.026	.002			2.52	2.79	22.04	10	.028	.015	.58
I ask myself how well I accomplish my goals once I'm finished			2.69	2.88	42.06	11	.034	<.001			2.71	2.74	28.11	10	.034	.002	.57
I ask myself if there was an easier way to do things after I finish a task			0.87	1.85	25.08	12	.021	.014			0.91	1.76	12.03	12	.001	.443	.66
I know how well I did once I finish a test			0.48	1.19	94.93	12	.053	<.001			0.50	1.09	68.70	12	.055	<.001	.69
I ask myself if I learned as much as I could have once I finish a task			2.80	2.85	27.30	11	.024	.004			2.84	2.77	17.83	10	.022	.058	.57
I summarize what I've learned after I finish			1.81	1.56	38.03	11	.031	<.001			1.86	1.47	25.89	10	.032	.004	.59

Differential Item Functioning

Notes: Only p-values adjusted for False Discovery Rate (FDR) are reported. A dash (–) indicates p > .05. For Race/Ethnicity, a "Prefer not to answer/Unlisted" category was included in the DIF analysis; however, the results for this category are not reported in this technical report.

Item	One-dimensional Model				Theory-based Three-dimensional Model			
	Pre-post	First Generation Status	Gender	Race/Ethnicity	Pre-post	First Generation Status	Gender	Race/Ethnicity
I ask myself if I have considered all options when solving a problem	–	–	–	–	–	–	–	–
I find myself analyzing the usefulness of strategies while I study	–	–	–	–	–	–	–	–
I ask myself periodically if I am meeting my goals	–	–	–	–	–	–	–	–
I ask myself questions about how well I am doing while learning something new	–	–	–	–	–	–	–	–
I consider several alternatives to a problem before I answer	–	–	–	–	–	–	–	–
I find myself pausing regularly to check my comprehension	–	–	–	Asian VS Black/African American .009 Asian VS Latinx/Hispanic .044 Asian VS Multiracial .004 Asian VS White .002	–	–	–	Asian VS Multiracial .016
I periodically review to help me understand important relationships	–	–	–	–	–	–	–	White VS Multiracial .048
I ask others for help when I don't understand something	–	–	.001	White VS Latinx/Hispanic .013	–	–	–	–
I change strategies when I fail to understand	–	–	–	–	–	–	–	–
I re-evaluate my assumptions when I get confused	–	–	–	–	–	–	–	–
I stop and reread when I get confused	–	–	–	Asian VS Multiracial .032				

Item	One-dimensional Model				Theory-based Three-dimensional Model			
	Pre-post	First Generation Status	Gender	Race/Ethnicity	Pre-post	First Generation Status	Gender	Race/Ethnicity
Black/African American VS Multiracial .020, Multiracial, Black/African American VS White .022	–	–	–	–				
I stop and go back over new information that is not clear	–	–	–	Asian VS Multiracial .048 Black/African American VS Multiracial .020	–	–	–	–
I ask myself if I have considered all options after I solve a problem	–	–	–	–	–	–	–	White VS Multiracial .047
I ask myself how well I accomplish my goals once I'm finished	–	–	–	–	–	–	–	–
I ask myself if there was an easier way to do things after I finish a task	–	–	–	–	–	–	–	–
I know how well I did once I finish a test	–	–	.004	–	–	–	–	–
I ask myself if I learned as much as I could have once I finish a task	–	–	–	–	–	–	–	White VS Multiracial .028
I summarize what I've learned after I finish	–	–	–	–	–	–	–	–

SRI

SRI Education, a division of SRI International, is helping federal and state agencies, school districts, major foundations, nonprofit organizations, and international and commercial clients tackle some of the most complex issues in education to help students succeed. Our mission is **to reduce barriers, optimize outcomes, and ensure educational equity for all children, youth, and families**. We do this by conducting high-quality research, supporting use of data and evidence, helping to strengthen state and local systems, and developing tools that improve teaching and accelerate and deepen learning. Our work covers a range of topics: early learning and development, disability and inclusion, supporting multilingual learners, student behavior and well-being, teaching quality, digital learning, STEM and computer science, and literacy and language arts, and college and career pathways. **We believe diversity in our organization and project teams leads to better and more equitable research and technical assistance, resulting in improved outcomes for all.**

SRI is a nonprofit research institute whose innovations have created new industries, extraordinary marketplace value, and lasting benefits to society.

Silicon Valley
(SRI Headquarters)
333 Ravenswood Avenue
Menlo Park, CA 94025
+1.650.859.2000
education@sri.com

Washington, D.C.
1100 Wilson Boulevard, Suite 2800
Arlington, VA 22209
+1.703.524.2053

www.sri.com/education-learning/

STAY CONNECTED