Phase I Technical Narrative

Enhanced MindPrint Profiles: Integrating Cognitive and Self-Regulatory *Strategies to Improve Mathematical Outcomes* NSF-2015003)

Summary of Completed Research and Remaining Activities

In Phase I, we proposed to enhance learner profiles through two sets of activities. In one set of activities, we developed a SEL survey (*SELfe Survey*); collected cognitive (MindPrint Assessment), SEL, and math achievement data from middle school students; and analyzed the data to validate the the *SELfe Survey* and demonstrate that (1) cognitive factors alone could predict math outcomes better than the other factors, and (2) SEL factors improved the predictability of math scores above and beyond cognitive factors alone. Math achievement was measured by MAP Growth--a norm-referenced benchmark achievement assessment. We also conducted experience sampling to understand how students' self-report on the survey matched with their actual in-the-moment behaviors. Although we are still analyzing the experience sampling data, our analyses have established the reliability and validity of the *SELfe Survey*, confirmed the high predictability of math scores by the MindPrint Assessment, and improved the predictability of math scores by including SEL factors.

In a second set of activities, we proposed to group students who have similar cognitive and SEL profiles but who differ meaningfully in math achievement; use the groupings to generate recommendations for math performance improvement (e.g., so that lower-performing students might benefit from the strategies used by the higher-performing students with similar profiles); provide recommendations to students based on their profiles; and collect teacher and student feedback on the recommended strategies. By grouping students based on their profiles, we learned that those students with strong reasoning skills (i.e., verbal and abstract reasoning on MindPrint) who had lower-than-expected math scores often also had lower motivation (self-efficacy, outcome expectations, or value) to perform well in math. In other cases, other cognitive factors (weaker executive functions, attention, or visual memory) were implicated in lower math achievement, despite strong reasoning skills. Overall, we concluded that more sophisticated, theory-based statistical modeling techniques (such as structural equation modeling) would be needed to understand the different paths by which students reached, or failed to reach, their potential for math achievement.

We were also able to provide recommendations to students based on their profiles and collect teacher and student feedback on the *Boost Strategies* and *SELfe Checklists*. Anecdotal evidence suggests that students found the strategies to be "favorable, engaging, and relevant"--level 1 (*Reaction*) of Kilpatrick's training evaluation model (1959, 1976). As one student put it, "it makes me feel like I have less work, you should tell other kids about these strategies!" Based on teacher feedback, we are also continuing to modify our product to ensure the most simple, easy-to-use approach in classrooms. For example, we learned that while we had the data to effectively personalize checklists for every student, doing so would not be manageable for teachers who need to be aware of every student's performance. Instead, a better solution would be to have three checklists per class based on students' profiles. Expert teachers would have the option to modify those strategies.

In the remainder of this section, we provide more detail about the Phase I SEL survey development and data analysis. The development of the SEL survey, data collection, and data analysis supporting the feasibility of the innovation consisted of the following steps.

1. *Development of the SEL survey* was completed by adapting 46 items from research-based, validated instruments on motivation (*Motivated Strategies for Learning Questionnaire*, Pintrich & DeGroot, 1990; *STEM Career Interest Survey*, Kier et al., 2014), personality (*Big Five Inventory*; John & Srivastava, 1999), and self-regulation (*Capacity for Self-Control Scale*; Hoyle & Davisson, 2016; and *Self-Control Scale*; Tangney et al., 2004). All items were contextualized to math activities, such as

homework and studying. We also included 29 items that tapped strategy usage, drawing from MindPrint's evidence-based collection of *Boost Strategies* (with items pertaining to general math study/homework, learning new math concepts, what to do when "stuck" on a math problem, and studying for math tests). All 75 items were assessed on a 5-point Likert scale, ranging from 1 ("Never or Almost Never") to 5 ("Always or Nearly Always").

2. *Data collection* was conducted after a brief period of pilot testing and revision (with 8 students). The SEL survey was administered to students attending a public middle school in an upper-middle class suburb of a large, northeastern US city. At the time of data collection (late fall 2020), the school was using a hybrid remote/in-person model to reduce transmission of COVID-19. A total of 215 students completed the SEL survey, which was delivered online via Google Classroom. During the fall of 2020, students in this school completed the MAP Growth assessment (n = 289) and the MindPrint cognitive assessment (n = 315). A total of 174 students completed both of these assessments, of which 170 also completed the SEL survey. The combined dataset was used to confirm and extend regression results concerning the prediction of MAP Growth outcomes, as discussed below.

3. *Establishment of reliability and validity of the SEL survey* is critical to ensure strategy recommendations are appropriate. To determine whether SEL constructs can be reliably and validly measured based on the scales and subscales in the survey, we: 1) calculated Cronbach's alpha (α) to assess the internal consistency of items; and 2) tested the model fit of confirmatory factor analyses (CFA) to establish construct validity. Because CFA models can be used to understand the internal structure or dimensionality of constructs, some CFA models (noted below) included two or more constructs. For example, given the close conceptual relationship between self-efficacy and outcome expectations, the CFA on expectancy combines items for these constructs.

   Using these approaches, we identified eight core constructs with sufficient reliability and validity to support strategy recommendations. These are organized in Table1 below according to whether the measurement properties are "strong, " "moderate," or "poor." Strong measurement properties include high internal consistency (α >.70) AND good-to-excellent model fit (Chi-Square (χ^2) $p > .05$, CFI > .90, RMSEA > .05). Moderate measurement properties include high internal consistency OR good-to-excellent model fit. Poor measurement properties are indicated by lack of both of these characteristics.

   Overall, we found that scales for critical constructs involving motivation (expectancy, mastery), self-regulation (persistence, inhibition of non-aversive distractors), and study strategies (practice) have strong measurement properties. For three constructs (value, conscientiousness, metacognition), scales had moderate measurement properties. We note that MindPrint's cognitive tasks have been extensively validated in previous large-sample studies(e.g., Brensinger & Gur, 2010; Moore et al., 2015; Parsons, 2019; Swagerman, et al., 2016).

   In addition to these measurement models, we also explored a bifactor CFA model with five factors (*Persistence*, *Inhibition of Non-Aversive Distractors*, *Help-Seeking*, *Practice*, and *Metacognition*) for a broad construct of academic/school engagement (for other bifactor models of engagement, see Wang et al. (2016, 2019). School engagement refers to the quantity and quality of students' interactions with learning activities (Fredricks et al., 2004; Wang et al., 2019). Based on values of item-explained common variance (I-ECV) greater than .80 (Dueber, 2017; Stucky & Edelen, 2015), a unidimensional *Cognitive Engagement* Scale was constructed from the three *Practice* items and three of the *Metacognition* items. This scale had strong measurement properties: $\alpha = .778$ (n=209); CFA (n=238): $\chi^2 = 8.209$, $p = .513$, CFI = 1.000, RMSEA = .000. Bifactor models are particularly valuable for evaluating assumptions about construct dimensionality and the implications of those assumptions on test scores (Dunn & McCray, 2020; Reise et al., 2007). In this case, the bifactor analysis confirmed that it is appropriate to measure students' cognitive engagement as a unidimensional construct.

Table 1. Measurement Properties of Scales for Eight Core SEL Constructs

Constructs Sub-Constructs	Example	# of items	α	χ^2 *(p)*	CFI	RMSEA	Notes
Scales with Strong Measurement Properties							
Expectancy		8	0.824	22.272 (.220)	0.991	0.032	
Self-Efficacy	I am able to complete my math homework	5	0.763				Two-factor CFA. The fit of this model is not as good as the one-factor ("Expectancy") solution. The correlation between the factors is .98.
Outcome Expectations	Compared to other students in my math class, I expect to get a good grade.	3	0.629	28.951 (.049)	0.980	0.051	
Mastery		5	0.839	1.382 (.240)	0.999	0.040	
Growth Mindset	Even when I do poorly on a math assignment, I feel that it's within my power to improve.	2	0.819				
Mastery Goal Orientation	I prefer math problems that are challenging so I can learn new things.	3	0.790	5.849 (.321)	0.998	0.027	Two-factor CFA
Persistence	Even when math homework is long and boring, I keep working until I finish.	4	0.768	3.226 (.199)	0.994	0.051	
Inhibition of Non-Aversive Distractors	When I do math I am distracted by sports or other hobbies.	7	0.809	17.920 (.118)	0.983	0.046	
Practice	I re-solve problems I answered wrong before to ensure I understand them.	3	0.771	.317 (.574)	1.000	0.000	Item error variances were equated to over-identify the model.
Scales with Moderate Measurement Properties							
Conscientiousness	I do a thorough job on math homework assignments.	4	0.701	7.743 (.021)	0.967	0.110	
Metacognition	[When learning a new math concept] I refer back to past lessons and the textbook to understand how it connects to what I've learned before.	4	0.656	3.756 (.167)	0.938	0.072	
Value	I am interested in careers that use math.	3	0.658	1.519 (.468)	1.000	0.000	Item error variances were equated to over-identify the model.

Reliability: Cronbach's Alpha (α). Confirmatory Factor Analysis (CFA) Indices of Global Fit: Chi-Square (χ^2), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA). N for all scales was 202-209 for reliability and 238 for the CFAs.

4. *Confirmation of predictability of math scores by cognitive factors* was accomplished by replicating previous findings (reported in the Phase I proposal) that the MindPrint cognitive tasks account for 50% of the variance in math outcomes as measured by MAP Growth scores. To correct for non-normality, MAP Growth percentiles were first transformed into normal scores using Blom's formula (Bishara & Hittner, 2015; Solomon & Sawilowsky, 2009). In the first model, we regressed the normal scores of MAP Growth percentiles on the MindPrint cognitive domains (working memory, attention, abstraction/flexibility, verbal reasoning, nonverbal reasoning, spatial cognition, verbal memory, spatial memory, sensorimotor speed, processing speed) and covariates (grade, gender, race/ethnicity, SPED status, 504 Plan status, free or reduced lunch (FRL) status). This model

explained 56.7% of the variance in math outcomes. A second model, with significant predictors only (verbal reasoning, nonverbal reasoning, spatial memory, attention; indicator variables for Asian, SPED, and Gender) accounted for 52.5% of the variance in math outcomes. The Model Summary and Coefficients Table for the second model are presented below. The second model was used as the base model for the analyses with individual SEL predictors, discussed next.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.437[a]	.191	.173	.8988533	.191	10.638	3	135	.000
2	.724[b]	.525	.499	.6994808	.334	22.981	4	131	.000

a. Predictors: (Constant), Flag for SPED Status, Gender based on demographic data, Flag for Asian students

b. Predictors: (Constant), Flag for SPED Status, Gender based on demographic data, Flag for Asian students, Spatial Memory Accuracy, Verbal Reasoning Accuracy, Attention Accuracy, Nonverbal Reasoning Accuracy

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	-.163	.107		-1.531	.128			
	Gender based on demographic data	.336	.154	.169	2.175	.031	.193	.184	.168
	Flag for Asian students	1.066	.235	.355	4.540	.000	.382	.364	.351
	Flag for SPED Status	-.494	.284	-.135	-1.738	.085	-.172	-.148	-.134
2	(Constant)	-.438	.090		-4.877	.000			
	Gender based on demographic data	.303	.125	.153	2.431	.016	.193	.208	.146
	Flag for Asian students	.474	.194	.158	2.447	.016	.382	.209	.147
	Flag for SPED Status	-.475	.222	-.130	-2.136	.035	-.172	-.183	-.129
	Attention Accuracy	.204	.072	.185	2.842	.005	.308	.241	.171
	Nonverbal Reasoning Accuracy	.283	.073	.263	3.909	.000	.503	.323	.235
	Verbal Reasoning Accuracy	.355	.083	.281	4.284	.000	.484	.351	.258
	Spatial Memory Accuracy	.251	.066	.241	3.811	.000	.362	.316	.230

a. Dependent Variable: Normal Score of MAP_percentile using Blom's Formula

5. *Improving predictability of math scores by including SEL factors* was accomplished by adding SEL factors to the base regression model with covariates and MindPrint cognitive factors. Cognitive and SEL factors were combined in two sets of regression analyses. In one set of analyses, each SEL factor was tested on its own to see if it significantly increased the amount of explained variance (as measured by increase in R^2 and statistical significance of its associated F test) over and above the covariates and four cognitive factors that significantly predicted normal scores of MAP Growth percentiles in the base model. These models also included the four two-way interaction terms for the products of each focal SEL predictor and each cognitive factor. The results show that of the regression analyses on 19 target SEL predictors, all but three yielded significant main effects and/or interactions involving the target variable. The significant effects ranged from small to small-medium in effect size. The three target variables for which analyses failed to show significant effects include

Extrinsic Motivation/Performance Goal Orientation, Integration, and Resources. Significant interactions include those between Attainment Value and Nonverbal Reasoning, Elaboration and Attention, and Growth Mindset and Nonverbal Reasoning. These results indicate that individual SEL factors and their interactions with the cognitive factors explained an additional 1.9% to 6.5% of the variance in MAP Growth scores. Effect sizes ranged from small ($f^2 = .02$) to medium-small ($f^2 = .07$).

In a second set of regression analyses, stepwise regression methods (forwards, backwards, and stepwise per se) were used to get an overall sense of the contributions of the complete set of cognitive and SEL factors. Although these analyses should not be used in the absence of theory to draw conclusions about specific predictors, these three regression analyses—taken together—indicate that the final models explain from 60.3% to 63.9% of the variation in math outcomes. This can be compared to the 52.5% of variance explained by the base model. In other words, the addition of the SEL factors explained an additional 10% of the variance in math scores. Furthermore, these analyses revealed common predictors in the final models of two or three of the analyses, as indicated below.

Despite our small sample size, we were able to fit a full structural equation model to the data. The fit of the model was reasonable ($\chi^2 = 282.179$, $p = .000$, CFI = .943, RMSEA = .041, and RMSEA Pclose = .912). A simplified version of this model, without indicators or error terms, is presented in Figure 1 below. Only factors with significant paths ($p < .05$) are included; standardized regression weights (betas) are indicated for each regression path. The bold italicized values are squared multiple correlations, indicating how much variance in the corresponding factor is explained by all of its predictors. This model shows that the motivational factor of Expectancy has an indirect effect on MAP Growth scores (through Persistence and Cognitive Engagement); Cognitive Engagement has a direct effect on MAP Growth; and Complex Cognition has both direct and indirect effects on MAP Growth. (Complex cognition is measured by verbal reasoning, nonverbal reasoning, and spatial cognition; see Moore et al., 2105, for the CFA supporting this factor.) Although this model is tentative without cross-validation (which is planned for Phase II), it indicates that an impressive 75% of the variance in MAP Growth scores (in dark blue) is predicted by a combination of cognitive factors, SEL factors (in pink), and engagement-related factors (in blue). Note that the increase in explained variance, compared to regression, is partly due to including indirect effects, though different factors were included in the SEM analysis based on the confirmatory work.

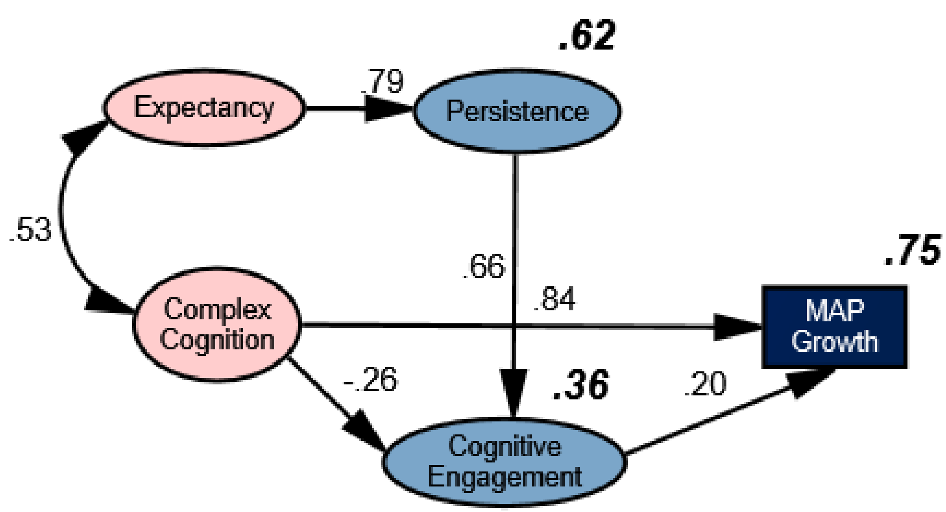

Figure 1: Full Structural Equation Model with factors differing in malleability: low (pink), moderate (blue), high (dark blue). See below for discussion.

Interestingly, the negative association between Complex Cognition and Cognitive Engagement suggests that strong reasoners may be *less* inclined to use metacognitive and practice strategies, instead relying on their strong reasoning skills to perform well in math. This outcome, while not originally predicted, is consistent with what we know to be true from research (Seo & Patall, 2020) and from near universal experience. For example, imagine Student X has well-above average complex reasoning skills and speeds through homework without a mistake, hardly studies for tests, and earns straight A's. Student X's teacher is concerned that Student X doesn't pay attention in class or hand in all her assignments, but it's hard to criticize a student with such strong grades. In contrast, Student Y has average complex reasoning skills and weak visual memory, and needs more examples and repetition to understand a new concept. Student Y often starts homework forgetting what he learned in class that day and needing to review notes in order to answer the first few problems. While Student Y is also very capable of mastery, he requires more persistence, self-regulation, and self-control to succeed in math class.

Based on our Phase I study, we can assert that the combination of cognitive and SEL data enables MindPrint to not only predict both students' scores but also drive recommendations to elevate and support both students. Student X might be doing well now, but when the work becomes more challenging, her lack of persistence and self-control might cause challenges if she fails to adapt to the more challenging work. MindPrint's strategy checklist for Student X includes methodically checking work and solving problems in multiple ways to develop persistence and a deeper understanding. Student Y's checklist centers on spaced repetition and multi-modal reinforcement so he remembers what was covered in class beyond the chapter test. Student Y might be reminded to use positive self-talk so he perseveres when topics are difficult.

As noted above, remaining activities include a full analysis of the experience-sampling data. In addition, we are writing white papers to disseminate and promote our Phase I research.

Description of Problems Encountered and Resolved

MindPrint successfully completed the most critical aspects of the Phase I research despite the COVID-19 pandemic, which limited our ability to work in-person with teachers and students. Some aspects of our proposed research were necessarily curtailed or canceled. Namely we had a smaller sample size than originally anticipated, were unable to perform in-person focus groups and user testing on the *SELfe Checklists* and *Boost Strategies,* and were not able to conduct traditional experience sampling. However, data analysis suggests the sample size was sufficient to provide conclusive results and we were able to have ongoing virtual discussions with teachers and students to collect feedback and make informed design changes on that feedback. We are still analyzing experience sampling data to determine the value in Phase I and intend to pursue this area of investigation in Phase II.

The reduction of in-person time, enabled us to delve deeper into the analysis of collected data to understand how the combination of cognitive (*MindPrint Cognitive Assessment*) and SEL (*SELfe Survey*) data enhances our ability to predict and improve student MAP Growth math achievement outcomes.

Description of Remaining Problems or Unmet Research Objectives

Due to COVID-19, the implementation of the small-scale in situ study during Phase I was infeasible. That being said, we were able to provide recommendations to students based on their profiles and collect teacher and student feedback on the *Boost Strategies* and *SELfe Checklists*. Otherwise, there are no remaining problems or unmet research objectives.

Conclusions of Phase I Findings in Support of Phase II Proposal

The innovation proposed in MindPrint's Phase I SBIR was to extend the *MindPrint Cognitive Assessment* to integrate context-specific SEL data and deliver interactive, personalized instructional strategies through proprietary algorithms. We demonstrated the feasibility of the innovation by identifying the cognitive factors, SEL factors, and study strategies that accurately predict over 60% of the variance in math outcomes.

Beyond accurately predicting variance in math outcomes, the Phase I findings also illuminated a research framework for understanding which factors are least malleable but can be supported and which factors are more malleable and the focus can be on direct improvement. Continued R&D in Phase II is anticipated to not only produce models that differentiate among factors in the following three "layers" (see Figure 1 above) but also guide which evidence-based strategies will be most effective in supporting learners.

> 1) *Least malleable (but not unmalleable) "raw ingredients" that are inputs into learning processes* (e.g., the factors in pink above). This layer includes personality characteristics, cognitive aptitudes/capacities (i.e., MindPrint cognitive factors), and motivation (expectancy, value, mindset, and goal orientation). Inputs would also include covariates (race/ethnicity, IEP/504 status, gender, age/grade, and prior math knowledge). Although some of these factors cannot be easily changed (e.g., we don't anticipate growth in cognitive factors over time), they can be supported so students learn strategies to compensate for their weaker skills.
>
> 2) *More malleable "engagement processes" that recruit or are constrained by raw ingredients and prior knowledge* (e.g., in blue above); and 2) directly impact the quantity and quality of learning. These processes fall under the multidimensional construct of academic/school engagement (Fredricks et al., 2004; Wang et al., 2019; Wang & Eccles, 2013) and include behavioral processes (persistence, effort, focus) and cognitive processes (study and practice strategies, including metacognitive and self-regulated learning approaches). These factors will be positively influenced by factors in layer 1 but also can be directly influenced and improved by strategies such as using concrete examples, defining purpose, and more traditional SEL strategies around improving student self-efficacy and self-management.
>
> 3) *Most malleable achievement metrics* (in dark blue above). In Phase II, we anticipate using multiple indicators of achievement, including MAP Growth scores, end-of-year state math exam scores, and math class grades. We propose that these outcomes will reflect the direct effects of raw ingredients (especially cognitive factors) on performance as well as indirect effects of these ingredients on engagement processes, which in turn affect performance. In short, to the extent that we can support weaker less malleable skills and influence more malleable skills, we expect that achievement will also increase.

Opportunities for Improvement in Phase II Research

We have identified three major opportunities as a result of the Phase I research:

Revision and Extension of the SEL Survey. Based on the measurement properties of the *SELfe Survey* scales discussed above, additional R&D can improve the measurement of value, conscientiousness, impulsivity, and metacognition and provide a more thorough measurement of factors related to school engagement. In addition, as part of our Phase II work plan, we intend to incorporate experience sampling to validate results from the survey and understand how those differences might improve our ability to collect more predictive SEL data in the future.

Refinement and Expansion of an SEL Course*.* Importantly, administrators and teachers interviewed through iCorps unanimously indicated that students' actual "minds-on" engagement in learning activities is a top priority in their work; at the same time, 70% also indicated that they didn't have an effective solution for how to engage all students and 50% indicated concerns about students' self-management skills. This is not surprising given the critical importance of engagement to learning (e.g., Dotterer & Lowe, 2011; Wang & Holcombe, 2010), including the learning of SEL competencies. In a large-scale study (25,896

K-12 students) of the associations between SEL and student engagement, Yang et al. (2018) found that teaching SEL competencies was significantly associated with behavioral-cognitive engagement at the student and school levels. Our customer discovery determined that a student course to develop their self-awareness on the cognitive and SEL competencies and providing them with the *SELfe Checklists* to develop these skills would be the most efficient and effective approach. Our Phase II R&D will include delivering individualized strategies to students through a modular SEL course and monitoring their growth on the *SELfe Survey*, and, over time, math outcomes.

Efficacy Study of the SEL Course. Although implementation of the small-scale in situ study during Phase I was infeasible, the Phase II project affords us the opportunity to study the impact of the newly developed *SEL Course* using a longitudinal randomized control trial (RCT) design that supports causal inference and mediation analysis, and in a larger sample that permits a test of measurement invariance across groups differing in socio-economic status (SES), race/ethnicity, and gender. In Phase II, we propose to shift our efforts from predictive modeling to explanatory modeling. That is, instead of just using cognitive and SEL factors to predict math outcomes, we will explore how these outcomes result from the *combination of these factors* in particular forms of engagement or learning. Based on this understanding, we will implement and refine the *SELfe* Course — adapted to students' individual differences on our cognitive and SEL measures —that leads to increases in the quantity and quality of engagement/learning processes, which in turn leads to increases in math achievement. In Part 2, we elaborate on the technical objectives, approach, and work plan that we will enable the MindPrint innovation to mature as a market solution.

References

Bishara, A. J., & Hittner, J. B. (2015). Reducing Bias and Error in the Correlation Coefficient Due to Nonnormality. Educational and Psychological Measurement, 75(5), 785–804. https://doi.org/10.1177/0013164414557639

Brensinger, C., & Gur, R. E. (2010). A cognitive neuroscience based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscientific Methods*, *187*(2), 254–262. https://doi.org/10.4135/9781452224749.n8

Dotterer, A. M., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. *Journal of youth and adolescence*, *40*(12), 1649-1660.

Dueber, D. M. (2017). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. https://doi.org/10.13023/edp.tool.01

Dunn, K. J., & McCray, G. (2020). The place of the bifactor model in confirmatory factor analysis investigations into construct dimensionality in language testing. Frontiers in psychology, 11, 1357.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.3102/00346543074001059.

Hoyle, R. H., & Davisson, E. K. (2016). Varieties of self-control and their personality correlates. *Handbook of self-regulation: Research, theory, and applications*, 396-413.

John, O. P., & Srivastava, S. (1999). *The Big-Five trait taxonomy: History, measurement, and theoretical perspectives* (Vol. 2, pp. 102-138). Berkeley: University of California.

Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2014). The Development of the STEM Career Interest Survey (STEM-CIS). *Research in Science Education*, *44*(3), 461–481. https://doi.org/10.1007/s11165-013-9389-3

Kilpatric, D. L. (1976). Evaluation of Training. In Craig, R. L. (Ed.), Training and Development Handbook (second edition). New York, NY: McGraw-Hill Book Company.

Kirkpatric, D. L. (1959). Techniques for evaluating training programmes. Journal of American Society of Training Directors, 11, 3-9.

Moore, T. M., Reise, S. P., Gur, R. E., Hakonarson, H., & Gur, R. C. (2015). Psychometric Properties of the Penn Computerized Neurocognitive Battery. *Neuropsychology*, *29*(2), 235–246. https://doi.org/10.1002/cncr.27633.Percutaneous

Parsons, T. D. (2019). *Technologically Enhanced Neuropsychological Assessments Review and Update. March*.

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33–40. https://doi.org/10.1037//0022-0663.82.1.33

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 16 Suppl 1, 19–31. https://doi.org/10.1007/s11136-007-9183-7

Seo, E., & Patall, E. A. (2020). Feeling proud today may lead people to coast tomorrow: Daily intraindividual associations between emotion and effort in academic goal striving. Emotion.

Solomon, S.R. & Sawilowsky, Shlomo. (2009). Impact of Rank-Based Normalizing Transformations on the Accuracy of Test Scores. Journal of Modern Applied Statistical Methods. 8. 448-462. 10.22237/jmasm/1257034080.

Stucky, B. D., & Edelen, M. O. (2015). *Using hierarchical IRT models to create unidimensional measures from multidimensional data.* In S. P. Reise & D. A. Revicki (Eds.), *Multivariate applications series. Handbook of item response theory modeling: Applications to typical performance assessment* (p. 183–206). Routledge/Taylor & Francis Group.

Swagerman, S. C., De Geus, E. J. C., Kan, K. J., Van Bergen, E., Nieuwboer, H. A., Koenis, M. M. G., HulshoffPol, H. E., Gur, R. E., Gur, R. C., & Boomsma, D. I. (2016). The computerized neurocognitive battery: Validation, aging effects, and heritability across cognitive domains. *Neuropsychology*, *30*(1), 53–64. https://doi.org/10.1037/neu0000248

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality*, *72*(2), 271-324.

Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American educational research journal*, *47*(3), 633-662.

Wang, M. T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The Math and Science Engagement Scales: Scale development, Validation, And psychometric properties. *Learning and Instruction*, *43*, 16–26. https://doi.org/10.1016/j.learninstruc.2016.01.008

Wang, M. T., Fredricks, J., Ye, F., Hofkens, T., & Linn, J. S. (2019). Conceptualization and assessment of adolescents’ engagement and disengagement in school: A multidimensional school engagement scale. *European Journal of Psychological Assessment*, *35*(4), 592–606.

Wang, M.-T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28,* 12–23. https://doi.org/10.1016/j.learninstruc.2013.04.002

Yang, C., Bear, G. G., & May, H. (2018). Multilevel associations between school-wide social–emotional learning approach and student engagement across elementary, middle, and high schools. *School Psychology Review*, *47*(1), 45-61.