AIR
AMERICAN INSTITUTES FOR RESEARCH®

Measurement Issues for Rapid-Cycle Studies Designed to Support Continuous Improvement and R&D

December 2015

Ryan Williams, Jessica Heppen, John Meakin, Martyna Citkowicz, Whitney Cade, Lauren Fellers

Measuring Personal Qualities

Purpose

"Personal qualities," also known as so-called noncognitive factors (Duckworth & Yeager, 2015) include constructs such as motivation, engagement, "grit," perseverance, self-efficacy, and sense of belonging. Believed to be important precursors to learning, these constructs have been capturing more attention from educational psychologists and intervention developers attempting to positively affect student achievement. Our work related to the measurement of personal qualities has had two goals: (1) to review current methods of measuring these constructs, including traditional measurement approaches such as self-reports, as well as emerging contemporary approaches such as educational data mining, learning analytics, machine learning, and psychophysiological assessment; and (2) to consider the relative merits of these various measurement approaches for rapid-cycle experimental research, which may include short-duration randomized controlled trials (RCTs), factorial designs, or sequential assignment designs for adaptive interventions.

The primary research questions driving this task are the following:

> *How are student personal qualities measured?*
>
> *Are these methods useful for rapid-cycle experimental research?*

Overview

There is growing interest and investment in advancing research and development of programs that promote students' personal qualities. Personal qualities or noncognitive factors are defined as sets of behaviors, skills, attitudes, and strategies that are crucial to academic success but are not tapped by current academic assessments (Duckworth & Yeager, 2015; Farrington et al., 2012; U.S. Office of Educational Technology, 2013).

Recent research suggests that developing personal qualities, such as motivation and self-regulation, can improve students' academic performance as well as reduce achievement gaps and the likelihood of success in the labor market (Farrington et al., 2012; Heckman & Kautz, 2013). Students who believe that intelligence is fixed may have low motivation to put forth effort.

Ongoing work by the Carnegie Foundation has drawn increased attention to a pressing need to support the development of "productive persistence" in students, defined as having the tenacity and strategies to be academically successful (Silva & White, 2013). Drawing on research by psychologists including Carol Dweck, Claude Steele, and others, the Carnegie Foundation and the Charles A. Dana Center have been designing and iteratively testing developmental mathematics courses for community college students through networked improvement communities. The courses (Statway and Quantway) emphasize the importance of providing social-emotional supports that develop students' sense of belonging, growth mindset, and productive persistence (Silva & White, 2013). These efforts reflect an overall building sense that without attention to personal qualities, particularly in at-risk students, interventions, initiatives, and reforms intending to reduce achievement gaps and improve U.S. students' academic performance will continue to fall short (Yeager, Walton, & Cohen, 2013).

Systematic reviews and meta-analyses have investigated the relationship between personal qualities and traditional academic achievement outcomes (i.e., measures of cognitive ability). In one meta-analysis, Poropat (2009) examined the correlations between the five-factor personality model and academic performance. He found significant and positive relationships between academic performance with three of the five factors: agreeableness, conscientiousness, and openness. In another systematic review and meta-analysis, Robbins et al. (2004) found that academic goals, self-efficacy, and academic-related skills were positively and significantly predictive of college performance and persistence.

The relationship between personal qualities and traditional academic outcomes has been subject to experimental investigation as well. Recent experimental research suggests that even brief messages and tasks designed to promote and reinforce "growth mindset" (Dweck, 2006) can improve student achievement, including that of low-income and minority students (Aronson, Fried, & Good, 2002; Blackwell, Trzesniewski, & Dweck, 2007).

Clearly, personal qualities encompass a wide range of attributes. In the review by Farrington et al. (2014), the following domains of personal qualities are identified and discussed:

- *Academic skills*. These may include behaviors such as going to school on time, completing homework, organizing materials, participating in lessons, and studying.
- *Academic perseverance*. This includes grit, tenacity, delayed gratification, self-discipline, approaches to learning, and self-control.
- *Academic mindsets*. These include sense of belonging to an academic community and sense of control or ability to grow (i.e., growth mindset).
- *Learning strategies*. These include study skills, metacognitive strategies for learning, self-regulated learning, and goal setting.
- *Social skills*. These include interpersonal skills, empathy, cooperation, assertion, and responsibility.

Personal Qualities and Short-Cycle RCTs

Personal qualities may be especially appropriate for study using short-cycle experiments for two reasons. First, personal qualities may be more responsive to change within a short period of time, unlike traditional student academic outcomes such as performance and achievement. That is, certain personal qualities, such as student engagement, might be malleable enough to demonstrate meaningful treatment-response in a short period of time. Second, growing research in machine learning, educational data mining, and psychophysiological assessment continue to provide opportunities to observe unobtrusively personal qualities, especially in online learning environments (e.g., Baker, D'Mello, Rodrigo, & Graesser, 2010). Increasingly, researchers are using real-time user data (e.g., back-end or clickstream data, time spent on task, eye tracking) to build and investigate measures of complex psychological processes such as engagement, disengagement, and productive persistence. As this research advances, it is not unrealistic to imagine an intervention delivery system (e.g., an online mathematics course) that builds measures of student personal qualities such as engagement or persistence almost simultaneously, even refining the measures with new data sources as they arrive.

Personal Qualities as Mediators

Personal qualities very commonly are assumed to be affected directly or indirectly by interventions in education research. For example, Armbruster, Patel, Johnson, and Weiss (2009) investigated an active learning instructional intervention for improving student achievement in biology courses. The intervention was intended to improve student attitudes toward learning in biology courses, which in turn would improve academic performance in the course. The mediating effects of student personal qualities perhaps has been studied most prominently in the mindset literature. Recent experimental research suggest that even brief messages and tasks designed to promote and reinforce "growth mindset" (Dweck, 2006) can improve student achievement, including that of low-income and minority students (Aronson, Fried, & Good, 2002; Blackwell, Trzesniewski, & Dweck, 2007). Indeed, many of the interventions we study attempt to improve students' academic outcomes by first improving their levels of engagement with school or the content being taught (e.g. algebra), their motivation to learn, sense of belonging in school, or other attitudes and mindset-related "precursors" to learning.

Although affecting traditional short-term and long-term student academic outcomes remains a critical endeavor, the importance of further understanding of the mediating effects of student personal qualities on these traditional outcomes cannot be understated.

Measurement Approaches

Self-Report Methods

The most common strategy used to measure personal qualities of students in Grade 3 and older is to rely on the students' own self-report using surveys. Self-report instruments typically include objective items that are scaled and normed to compare children's individual responses on groups of self-report items to the responses of a normative group.

As part of our review, we uncovered a large number of instruments used to measure student personal qualities. We identified more than 200 unique self-report method instruments of constructs, including mindset, self-efficacy, academic awareness, grit, connectedness, persistence, self-control, self-regulation, locus of control, learning skills, and self-efficacy. Table

1 summarizes the measures we have identified as part of this work, which we have stored in a database along with their length, target populations, and a summary of measurement properties.

Table 1. Summary of Self-Report Measures Identified

Construct	Number of Instruments
Activation Control	1
Activity Level	1
Affiliation	1
Attainment Value	4
Attention	1
Connectedness	14
Fear	1
Frustration	1
Goal Management	1
Goal Setting	4
Growth Mindset	7
High Intensity Pleasure	1
Inhibitory Control	1
Intrinsic Value	5
Learning Skills	13
Locus Of Control	11
Meta-Cognitive Skills	1
Mindset	6
Motivation	1
Negative Mindset	14
Perceptual Sensitivity	1
Performance Awareness	11
Persistence	16
Planning	10
Pleasure Sensitivity	1
Relevance	8
Self-Competence	21
Self-Correcting	6
Self-Efficacy	24
Self-Monitoring	7
Self-Regulation	19
Shyness	1
Social Capital	9
Task Value	8
Utility Value	1

Self-report measures likely are useful in rapid-cycle RCTs. They are relatively inexpensive to administer in most cases, which makes them attractive. The key disadvantage of self-report measures are self-report bias and nonresponse, two issues that take time and resources to minimize.

Educational Data Mining and Learning Analytics

One of the most promising areas of measurement for short-cycle experimental research is in educational data mining (EDM) and learning analytics (LA). We broadly classify the key methods into three categories: machine learning, structure discovery, and relationship mining. Machine-learning techniques focus on predicting outcomes based on subsets of data, then applying these predictions to the entire data set. Structure discovery, unlike prediction, uses the entire data set and attempts to *discover* what *structures* emerge naturally from the data. Relationship mining is similar to structure discovery in that there is not a specific predicted variable of interest. The goal in relationship mining is to establish relationships between variables in a data set with a large number of variables.

EDM and LA are some of the most promising frontiers in measurement and likely are well suited to certain types of rapid-cycle experimental research. Especially in studies of online learning programs (e.g., online algebra), archived back-end data may be used for constructing real-time measures of nonacademic constructs such as engagement and disengagement. One of the major disadvantages to these approaches is that they have high upfront costs in building algorithms that effectively extract measures of the target constructs. In the long term, however, data collection could be automated fully. Such measures avoid the administration and self-report biases inherent in traditional measurement approaches. Some of the foundational work in this area was developed by Ryan Baker, Sidney D'Mello, Art Graesser, Kalina Yacef, Neil Heffernan, Ivon Arroyo, Ken Koedinger, and Carolyn Rosé.

Psychophysiological Assessment

Psychophysiological assessment is another rapidly developing area in contemporary measurement. Over the last decade, researchers have experimented with physiological and psychophysiological methods to measure student outcomes. These include the use of computers, cameras, pressure and posture sensors, heart rate, electroencephalography (EEG), saliva, and galvanic skin response (GSR). We provide a brief review of some of the studies that have utilized these methods in the subsections that follow in order to provide a snapshot of the ongoing research in this growing field. Table 2 provides a cross-tabulation of these methods with the outcomes they have been used to measure.[1]

---

[1] Note that Table 2 is not meant to include a comprehensive review of the literature; rather, it provides a summary of the measures and outcomes included in this paper.

Table 2. Cross-Tabulation of New Psychophysiological Measures With Noncademic Student Outcomes

Psychophysiological Measures	Constructs				
	Affect	Engagement or Interest	Self-Efficacy	Stress	Cognitive Load
Cortisol				X	
IgA				X	
CgA				X	
GSR	X	X	X		
Heart Rate		X	X		
Pressure Sensor	X	X			
Posture Sensor	X	X			
Facial Detection	X	X			X
EEG	X	X			X

Note. IgA is the antibody secretory immunoglobulin A. CgA is the secretory protein chromogranin A. GSR denotes galvanic skin response. Facial expression includes eye tracking and any muscle movement detection in the face. EEG denotes all electroencephalography technology.

Like measures derived from EDM and LA, psychophysiological assessment provides unique and promising measurement opportunities for purposes of rapid-cycle experimental research. In fact, Baker and his colleagues have used these two approaches concurrently, building measures of engagement and disengagement. To that end, psychophysiological measures provide an important opportunity for measurement validation research. Although these approaches obviate self-report biases in traditional measures, unlike EDM and LA approaches, they may be intrusive and not as amenable to real-time data collection in large-scale field experiments. As technology improves, we are likely to see hardware that is increasingly mobile and less intrusive (e.g., step trackers), which may expand the utility of measurement approaches.

Conclusions and Next Steps

We have studied three approaches to measuring student qualities other than academic achievement: self-report methods, EDM and LA, and psychophysiological assessment. All three approaches have cost and efficiency advantages and disadvantages for purposes of rapid-cycle experimental research. Two areas of research require further attention, by American Institutes for Research and others. First, psychometric research on the utility, reliability, and validity of the contemporary measurement approaches described here is underdeveloped. This work includes building EDM- and LA-derived measures of student qualities and validating them against psychophysiological measures as well as traditional self-report measures. Second, empirical investigations on the malleability of student qualities are needed. The malleability of student qualities is discussed at length in the literature, primarily from a theoretical perspective (e.g., Farrington et al., 2014), but there are no systematic empirical investigations. Given that, a systematic review and meta-analysis would provide important empirical information about these measures' responsiveness to change. Pursuit of research in these two areas will advance knowledge about the measurement of student "precursors" to learning in the context of rapid-cycle experimental research.

References

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE-Life Sciences Education*, *8*(3), 203–213.

Aronson, J., Fried, C. B., & Good, C. (2002). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *Journal of Experimental Social Psychology*, *38*(2), 113–125.

Baker, R. S. J. d., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, *68*(4), 223–241.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, *44*(4), 237–251.

Dweck, C. (2006). *Mindset: The new psychology of success*. New York: Random House.

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

Heckman, J. J., & Kautz, T. (2013). *Fostering and Measuring Skills: Interventions That Improve Character and Cognition* (Working Paper No. 19656). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w19656

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261–288.

Silva, E., & White, T. (2013). *Pathways to improvement: Using psychological strategies to help community college students master developmental math.* Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved from http://www.carnegiefoundation.org/sites/default/files/pathways_to_improvement.pdf

U.S. Office of Educational Technology. (2013). *Promoting grit, tenacity, and perseverance: critical factors for success in the 21st century*. Damascus, MD: Penny Hill Press.

Yeager, D., Walton, G., & Cohen, G. L. (2013). Addressing achievement gaps with psychological interventions. *Phi Delta Kappan*, *94*(5), 62–65.