Evaluation of
Employment Coaching for TANF and Related Populations

Measuring Self-Regulation Skills in Evaluations of Employment Programs for Low-Income Populations:
Challenges and Recommendations

Measuring Self-Regulation Skills in Evaluations of Employment Programs for Low-Income Populations:
Challenges and Recommendations

Tim Kautz • Quinn Moore

Submitted to:
Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
330 C Street, SW
Washington, DC 20201
Project Officers: Hilary Forster and Victoria Kabak

Contract/Task Number: HHSP233201500035I / HHSP23337018T
Mathematica Reference Number: 50327

Submitted by:
Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Contents

TABLES

FIGURES

Overview

INTRODUCTION

People's ability to find, keep, and advance in a job depends on self-regulation skills in addition to education, work experience, and technical skills (Almlund et al. 2011). Self-regulation skills include the ability to finish tasks, stay organized, and intentionally control emotions and behaviors. Research has shown that these skills are important in attaining goals and in determining life outcomes, including those related to employment (Almlund et al. 2011). Research has also shown that interventions can both strengthen self-regulation skills and encourage their use (Kautz et al. 2014).

In response to this research, some employment programs, including those offered as part of the Temporary Assistance for Needy Families (TANF) program, use strategies designed to strengthen and boost participants' use of self-regulation skills (Cavadel et al. 2016; Kautz et al. 2014). To assess the effectiveness of these strategies, evaluators need a way to measure self-regulation skills. However, measuring self-regulation skills for such programs presents unique challenges.

This report discusses the challenges in measuring self-regulation skills and provides guidance on selecting measures to use in evaluations of employment programs for low-income populations. It complements a brief entitled "New Perspectives on Practice: A Guide to Measuring Self-Regulation and Goal-Related Outcomes in Employment Programs," by Cavadel et al. (2018) that focuses on providing guidance to practitioners by encouraging them to consider self-regulation outcomes and by introducing measures of goal-related and self-regulation skills. This report focuses on providing guidance to researchers on how to measure self-regulation skills in evaluation settings, highlighting ways to address unique challenges that arise in this application. This report was written by Mathematica Policy Research and funded by the Office of Planning, Research, and Evaluation (OPRE) of the Administration for Children and Families (ACF) as part of the *Evaluation of Employment Coaching for TANF and Related Populations*.

PRIMARY RESEARCH QUESTIONS

This report addresses three primary research questions:

1. What are the challenges of measuring self-regulation skills in the context of evaluations of employment programs that serve low-income populations?

2. What general criteria should evaluators use when selecting measures of self-regulation skills in this context?

3. What are the tradeoffs between different approaches for measuring self-regulation skills in this context?

PURPOSE

This report discusses issues related to measuring self-regulation skills in evaluations of employment programs for low-income populations.[1] First, it presents an overview of self-regulation skills and their importance for employment programs. Second, it introduces approaches to measuring self-regulation skills. Third, it discusses challenges when measuring self-regulation skills in evaluations of employment programs for low-income populations. Fourth, it provides criteria and recommendations for selecting measures.

KEY FINDINGS AND HIGHLIGHTS

Four challenges arise when measuring self-regulation skills in evaluations of employment programs for low-income populations. First, measures of self-regulation skills can reflect aspects of a person's situation (for example, his or her background or financial resources) in addition to his or her skills. Second, most existing measures were developed for purposes other than program evaluation, such as describing characteristics of populations generally or for diagnosing people with severe problems. Third, most existing measures were not designed for use with low-income populations. Fourth, some measures take a long time to administer or require special technology.

For use in evaluations of employment programs, we suggest that measures of self-regulation should: (1) relate to employment outcomes of interest; (2) capture skills that could be influenced by the program; (3) account for confounding factors that affect measurement but not skills, and (4) be feasible to administer in an evaluation.

To meet these criteria, we suggest using a set of both general measures of self-regulation as well as ones that are specific to the employment context, collecting information on other aspects of the participants' situations that can be affected by the program, modifying measures to fit the target population, and conducting analyses to assess the reliability and validity of selected measures.

METHODS

The report includes:

- A brief review of self-regulation skills and general approaches to measuring them

- A discussion of key challenges of measuring self-regulation skills in the context of evaluations of employment programs for low-income populations

- A critical analysis of various measurement approaches for this context

- Recommendations for measuring self-regulation skills in this context.

---

[1] For existing literature on self-regulation measures in employment programs, see "New Perspectives on Practice: A Guide to Measuring Self-Regulation and Goal-Related Outcomes in Employment Programs" by Cavadel et al. (2018), produced under the OPRE project *Goal-Oriented Adult Learning in Self-Sufficiency*. For a continuation of this work, see an upcoming brief from the *Evaluation of Employment Coaching for TANF and Related Populations* that will contribute to the literature on the validity of self-regulation measures by using data from the evaluation to test the validity of select self-regulation measures.

Executive Summary

People's ability to find, keep, and advance in a job depends on self-regulation skills in addition to education, work experience, and technical skills (Almlund et al. 2011). Self-regulation skills include the ability to finish tasks, stay organized, and intentionally control emotions and behaviors. Research has shown that these skills are important in attaining goals and in determining life outcomes, including those related to employment (Almlund et al. 2011). Research has also shown that interventions can both strengthen self-regulation skills and encourage their use (Kautz et al. 2014).

In response to this research, some employment programs, including those offered as part of the Temporary Assistance for Needy Families (TANF) program, use strategies designed to strengthen and boost participants' use of self-regulation skills (Cavadel et al. 2016; Kautz et al. 2014). To assess the effectiveness of these strategies, evaluators need a way to measure self-regulation skills.

Measuring self-regulation skills is challenging in any context because, unlike physical characteristics (such as height), self-regulation skills cannot be directly observed. Instead, self-regulation skills are always measured using *behaviors* broadly defined—how people act in various situations or perform on various tasks (Heckman and Kautz 2012). Such behaviors could include how individuals perform on a task designed to measure executive functioning, whether they control their temper at work, whether they tend to work hard, and whether they finish school.

Existing measures of self-regulation skills differ in *how* they measure skills (their mode) and *what* they measure (their content). Measurement modes include: (1) a self-report in a survey or interview that typically asks people about how they tend to behave; (2) an observer report in a survey or interview; (3) a performance task designed to capture particular self-regulation skills; and (4) administrative records about behaviors such as attendance. An important aspect of content is the extent to which measures capture skills in a specific setting through *contextualized behaviors* (for example, finishing tasks at work) or capture broader skills that apply across multiple settings through *generalized behaviors* (for example, the tendency to finish tasks in general). Important tradeoffs exist between different measurement modes and types of measurement content.

CHALLENGES OF MEASURING SELF-REGULATION SKILLS BASED ON BEHAVIORS

Because measures of self-regulation skills are based on behaviors, they can also reflect factors other than the self-regulation skill of interest (Figure ES.1). We refer to this set of other factors as a person's *situation*, broadly defined as their circumstances in life or conditions in which their skills are measured. Factors that affect a person's situation include: financial and social resources, his or her background, the setting in which the behavior is measured, and any incentives tied to exhibiting particular behaviors. Background is a key consideration for measuring skills in low-income populations because most surveys were

not developed for use with this population and might not be appropriate. For example, respondents with lower levels of education sometimes do not understand words that appear in existing surveys.

Figure ES.1. Measuring self-regulation skills using behaviors
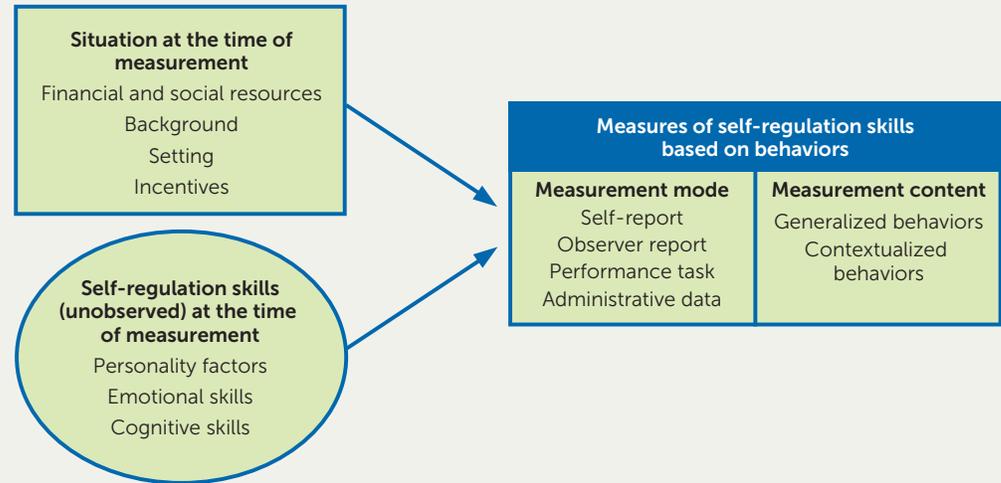
Because the measures of skills depend on the person's situation as well as his or her underlying skills, it is challenging to compare skills between people or measure how they develop in a person over time. If two people exhibit differences in a behavior at a point in time, then the difference might arise from either differences in self-regulation skills or differences in their situations. Similarly, an employment program might affect measures of self-regulation either by affecting the underlying self-regulation skills or by affecting the situation. Hence, evaluators who are interested in whether programs affect the underlying skills need to account for participants' situations.

RECOMMENDATIONS

To achieve the aim of estimating the extent to which employment programs impact employment outcomes by improving self-regulation skills, we suggest that measures of self-regulation should: (1) relate to employment outcomes of interest; (2) capture skills that could be influenced by the program; (3) account for confounding factors presented by the person's situation that affect measurement but not skills; and (4) be feasible to administer in an evaluation setting.

To meet these criteria, we suggest using a set of both general measures of self-regulation as well as ones that are specific to the employment context, collecting information on other aspects of the participants' situations that can be affected by the program, modifying measures to fit the target population, and conducting analyses to assess the reliability and validity of selected measures.

I. Introduction

People's ability to find, keep, and advance in a job depends on self-regulation skills in addition to education, work experience, and technical skills (Almlund et al. 2011). Self-regulation skills include the ability to finish tasks, stay organized, and intentionally control emotions and behaviors. Research has shown that these skills are important in attaining goals and in determining life outcomes, including those related to employment (Almlund et al. 2011). Research has also shown that interventions can both strengthen self-regulation skills and encourage their use (Kautz et al. 2014).

In response to this research, some employment programs, including those offered as part of the Temporary Assistance for Needy Families (TANF) program, use strategies designed to strengthen participants' use of existing self-regulation skills and/or help them develop skills (Cavadel et al. 2016; Kautz et al. 2014). To assess the effectiveness of these strategies, evaluators need a way to measure self-regulation skills.

Measuring self-regulation skills for employment program evaluations is challenging. In particular: (1) measures of self-regulation skills can reflect other aspects of a person's situation such as his or her background or program-provided incentives; (2) most existing measures were developed for purposes other than program evaluation, such as describing characteristics of populations generally or for diagnosing people with psychological disorders; (3) most existing measures were not designed for use with low-income populations; and (4) some measures take a long time to administer or require special technology.

This report discusses the challenges in measuring self-regulation skills and provides guidance on selecting measures to use in evaluations of employment programs for low-income populations. It complements a brief entitled "New Perspectives on Practice: A Guide to Measuring Self-Regulation and Goal-Related Outcomes in Employment Programs," by Cavadel et al. (2018) that focuses on providing guidance to practitioners by encouraging them to consider self-regulation outcomes and introducing measures of goal-related and self-regulation skills. This report focuses on providing guidance to researchers on how to measure self-regulation skills in evaluation settings, highlighting ways to address unique challenges that arise in this application. Although we focus on measurement in the context of employment programs that serve low-income populations, some of our conclusions also apply more generally to evaluations of other types of programs and to those that serve different populations. A future brief will assess the validity of self-regulation measures used in an evaluation of various coaching interventions for low-income populations (see Box 1 for an overview of this evaluation).

Box 1. Evaluation of Employment Coaching for TANF and Related Populations

To learn more about the potential of coaching to help TANF recipients and other low-income individuals reach economic security, the Office of Planning, Research, and Evaluation (OPRE) of the Administration for Children and Families (ACF) contracted with Mathematica Policy Research and Abt Associates to conduct an evaluation of employment coaching interventions. The evaluation will use an experimental research design to examine the effectiveness of coaching interventions that aim to help low-income individuals succeed in the labor market. It will also examine the implementation of the coaching interventions, the impact of coaching on self-regulation skills, and the role of self-regulation skills in generating any impacts on employment outcomes.

II. The Definition and Importance of Self-Regulation Skills

While there is no universally accepted definition of "self-regulation," following Cavadel et al. (2016), we use the term to cover the broad set of skills that allow people to intentionally control their thoughts, emotions, and behaviors. We focus on three categories of self-regulation that are most relevant to finding, keeping, and advancing in a job: (1) personality factors, (2) emotional skills, and (3) cognitive skills. Table 1 provides examples of each category. These categories and examples are neither mutually exclusive nor exhaustive but illustrate the range of self-regulation. These self-regulation skills complement each other and enable people to set, pursue, and attain goals, including those related to employment (Cavadel et al. 2016).

Table 1. Examples of self-regulation skills

Skill category	Skill	Definition
Personality factors	Motivation	The desire to start and finish tasks.
	Grit	The ability to persevere to attain long-term goals.
	Self-efficacy	The belief we have in our ability to perform at a high level.
Emotional skills	Emotion understanding	The ability to understand emotions in ourselves and others.
	Emotion regulation	The ability to alter the intensity of the emotion being experienced and the behaviors that go along with that emotion.
Cognitive skills	Executive function	A set of cognitive skills that helps us regulate and control our actions, particularly intentional action, goal setting, and goal pursuit.
	Selective attention	The ability to attend to one particular aspect of a task in the face of other thoughts, information, and actions.
	Metacognition	A skill we use to observe and evaluate how we think, sometimes referred to as "thinking about thinking."

Source: Cavadel et al. (2016).

Self-regulation skills are linked to employment outcomes. Much evidence demonstrates that success at work depends on a range of self-regulation skills. For example, numerous studies show that conscientiousness—related to motivation and grit—is particularly predictive of job performance and wages across a wide range of occupations, including those typically held by low-income populations (Nyhus and Pons 2005; Salgado 1997; Hogan and Holland 2003; Barrick and Mount 1991). Moreover, emotional skills have been shown to improve both job attendance and job search outcomes (Störmer and Fahr 2013; Gallo et al. 2003; Caliendo et al. 2015). These findings parallel the results of employer surveys and interviews that demonstrate that employers place much value on skills that fall under the umbrella of self-regulation (Secretary's Commission on Achieving Necessary Skills 1992; Holzer 1997; Barton 2006; Cunningham and Villaseñor 2016).

Self-regulation skills develop throughout life. Mounting evidence has dispelled a commonly held belief that self-regulation skills—including personality factors—are unchangeable (Almlund et al. 2011). For example, studies have shown that emotional stability and some aspects of executive functioning tend to increase well into adulthood in people from a variety of socioeconomic backgrounds (Roberts et al. 2006; Roberts and Mroczek 2008; Steinberg 2008). Importantly, recent evidence suggests that interventions—especially those that involve mentorship or aim to teach self-regulation within a particular context—can successfully strengthen self-regulation (Kautz et al. 2014). Some programs that serve low-income populations explicitly seek to address and develop self-regulation. Many other programs do not overtly target or name "self-regulation skills" but still offer services that implicitly address them. For example, by working with program participants on goal setting and pursuit, a coach or another staff member can help participants practice self-regulation. To determine whether programs improve self-regulation skills and whether any associated improvements affect employment outcomes, evaluators of employment programs have become increasingly interested in measuring these types of skills.

III. Approaches to Measuring Self-Regulation Skills Based on Behaviors

Measuring self-regulation skills is challenging because, unlike physical characteristics (such as height), self-regulation skills cannot be directly observed. Instead, self-regulation skills are always measured using *behaviors* broadly defined: how people act in various situations or perform on various tasks (Heckman and Kautz 2012). For example, grit—perseverance for long-term goals—is measured indirectly with survey items that ask about behavior such as the extent to which the respondent has "overcome setbacks to conquer an important challenge" (Duckworth and Quinn 2009). The logic is that overcoming setbacks demonstrates the use of the underlying skill. Other behaviors related to self-regulation could include how individuals perform on a task designed to measure executive functioning, whether they control their temper at work, whether they tend to work hard, and whether they finish school.

Existing measures of self-regulation skills differ in *how* they measure skills (their mode) and *what* they measure (their content).

Measurement mode. Self-regulation skills can be measured in four ways:

1. **Self-report in a survey or interview.** The most common way to measure self-regulation skills is to ask people questions about how they tend to behave across a wide range of situations. For example, the Behavior Rating Inventory of Executive Function—Adult Version (BRIEF-A) asks about nine aspects of executive functioning based on typical behavior (Roth and Gioia 2005). One item on the BRIEF-A asks about the extent to which people "get upset quickly or easily over little things." Such self-reports can be collected through surveys or interviews.

2. **Observer report in a survey or interview.** An alternative is for an observer to rate an individual's behaviors. Observers could include peers, teachers, program staff, or employers. Ideally they have observed the individual frequently. Many surveys, including the BRIEF-A, have both self-report and observer-report versions.

3. **Performance task.** Self-regulation skills can also be measured through performance on standardized tasks. For example, the "Stroop color and word test" is a task-based measure of inhibitory control (Stroop 1935). Respondents are presented with text that spells the name of one color but is written in ink of a different color. They are then instructed to name the color of the ink, rather than the color spelled out in the text. For example, they might see the word "red" written in green ink, in which case the correct response is "green." The more accurately they name the color of the ink, the better their inhibitory control. These types of tasks can be administered in a laboratory setting or through a computer platform.

4. **Administrative records of behavior.** Administrative records—collected from a program, school, or employer—provide another source of data about people's behaviors. For example, attendance records collected by schools have been used in some studies to measure self-regulation skills that relate to conscientiousness (Jackson 2016; Kautz and Zanoni 2015; West et al. 2016). Similarly, researchers could use program or employer records on participation, punctuality, or attendance to measure self-regulation skills.

Table 2 provides examples of measures in each mode.

Measurement content. Measures of self-regulation skills can be based on how people behave in a specific setting (contextualized behaviors) or across multiple settings (generalized behaviors):

- **Contextualized behaviors.** These behaviors reflect self-regulation skills in specific settings. An example is whether a person completes tasks that are specific to a particular setting, such as submitting job applications. Performance tasks are often contextualized behaviors and therefore might not apply in other settings. For example, a person's performance on the Stroop color and word test might not relate to whether they can inhibit behaviors relevant to the workplace. Few existing measures were designed for employment contexts.

- **Generalized behaviors.** These behaviors reflect self-regulation skills that may apply across many settings. An example is whether a person tends to complete tasks in general. As noted above, many existing self- and observer-reports focus on generalized behaviors. Likely because they are broadly applicable, generalized behaviors predict a wide range of outcomes in addition to employment, including educational attainment and health (Almlund et al. 2011). With some wording changes, existing measures of generalized behaviors can be adapted to the employment context.

As discussed later, important tradeoffs exist between different measurement modes and types of measurement content.

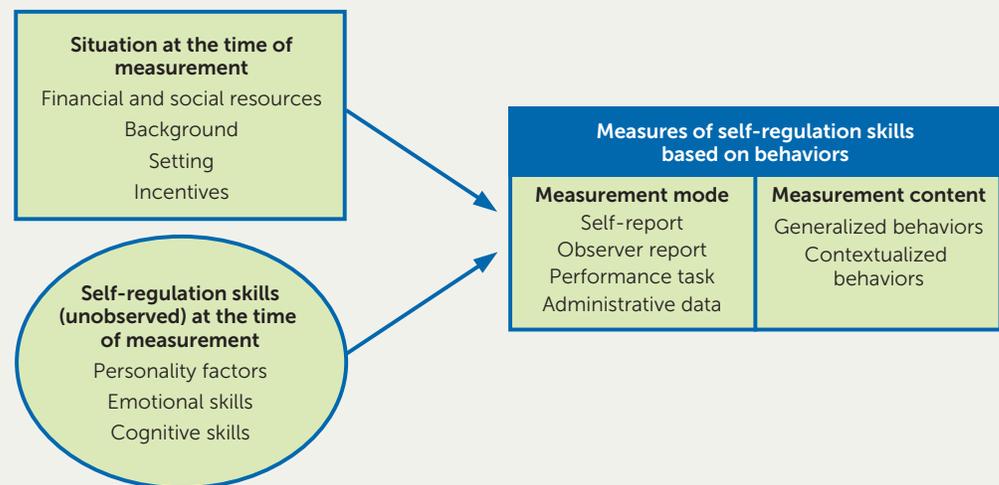**Table 2. Examples of measurement modes and measurement content**

Measurement mode	Example measure	Description of measurement content	Example item
Self-report on surveys or interviews	Behavior Rating Inventory of Executive Function—Adult Version (BRIEF-A), self-report (Roth and Gioia 2005)	A 75-item self-reported survey that measures aspects of executive functioning, including the ability to inhibit, self-monitor, plan/organize, shift attention, initiate tasks, monitor tasks, control emotions, use working memory, and organize materials.	"I get upset quickly or easily over little things."
	Grit scale (Duckworth and Quinn 2009)	A 10-item survey that measures grit.	"I have overcome setbacks to conquer an important challenge."
	Big Five Inventory (John et al. 1991)	A 44-item survey that measures personality factors including openness to experience, conscientiousness, extraversion, agreeableness, and emotional stability.	"I see myself as someone who tends to be organized."
	Self-regulation Skills Interview (Ownsworth et al. 2000)	A semi-structured interview that assesses awareness, readiness to change, and strategic behavior for people who report facing a difficulty related to everyday living, labeled their "main difficulty."	"How motivated are you to learn some different strategies to help overcome [main difficulty]?"
Observer report on surveys or interviews	BRIEF-A, informant report (Roth and Gioia 2005)	A 75-item observer-reported survey that measures the same aspects of executive functioning as the BRIEF-A self-report.	"Makes careless errors when completing tasks."
Performance task	Marshmallow test (Mischel and Ebbesen 1970)	A measure of self-control in which participants are given one marshmallow and are rewarded with a second marshmallow if they refrain from eating the first marshmallow within an allotted time. People who do not eat the first marshmallow are viewed as having higher self-control.	
	Digit span task (Wechsler 1949)	A measure of working memory in which respondents are given a series of numbers to memorize and asked to recall that series in reverse order. The better that they can recall the numbers, the better their working memory.	
	Stroop color and word test (Stroop 1935)	A measure of inhibitory control in which respondents must inhibit natural responses. Respondents are presented with text that spells the name of one color but is written in the ink of a different color. They are instructed to name the color of the ink, rather than the color spelled out in the text. For example, they might see the word "red" written in green ink.	
Administrative records	Attendance records (Jackson 2016; Kautz and Zanoni 2015)	Attendance or participation in a school, program, or other activity can measure self-regulation. This information is often tracked in administrative records.	

IV. Challenges of Measuring Self-Regulation Skills Based on Behaviors

Because measures of self-regulation skills are based on behaviors, they can reflect factors other than the self-regulation skill of interest. All measures of self-regulation skills are based on behaviors. But as illustrated in Figure 1, factors other than self-regulation skills may affect behaviors at the time of measurement. We refer to this set of other factors as a person's *situation*. A person's situation includes:

- **Financial and social resources.** This category includes all the resources available to a person such as savings, social support, access to childcare through relatives, and access to transportation. As discussed further below, these resources can affect behaviors. For example, someone might be late to an appointment not because they lack self-regulation skills but because they do not have adequate transportation or childcare.

- **Background.** A person's background is his or her demographic and socio-economic status, including level of education and peer groups. As discussed in more detail below, background can affect how people respond to a survey.

- **Setting.** The setting refers to the location in which someone exhibits a behavior. For example, people could behave differently at home versus at work.

- **Incentives.** Incentives are rewards that people receive for exhibiting specific behaviors. For example, employment programs might reward participants with money for regular attendance or achieving specific goals, or an employer might give an employee a raise if they perform well.

Figure 1. Measuring self-regulation skills using behaviors

Situation at the time of measurement
Financial and social resources
Background
Setting
Incentives

Self-regulation skills (unobserved) at the time of measurement
Personality factors
Emotional skills
Cognitive skills

Measures of self-regulation skills based on behaviors

Measurement mode
Self-report
Observer report
Performance task
Administrative data

Measurement content
Generalized behaviors
Contextualized behaviors

If two people exhibit differences in a behavior, then the difference might arise either from differences in self-regulation skills, differences in their situations, or some combination of both. Importantly, this discussion focuses on measurement at a point in time. At the time of measurement, skills and situations are fixed, so the situation does not affect the skills (or vice-versa). In this way, we view skills as attributes that are tied to a person so that, at a point in time, the person has the same set of skills, regardless of their situation. For example, on a given morning, a person would have the same set of skills regardless of whether they decide to attend work or not; they take their skills with them to either situation. In contrast, a person's behavior depends on their situation. For example, on a given morning, someone might behave differently depending on whether they are at work or at home, even though they would have the same set of skills at the time they enter either situation. Situations, however, could affect the underlying skills over time. For example, someone who attends work regularly might develop additional self-regulation skills as result of their gained experience. We consider skill development further in the next section. Examples of differences in situations that can affect measures of self-regulation even though the underlying skill is the same include:

- **Changes in financial and social resources can affect behaviors used to measure self-regulation.** Considerable evidence points to the importance of financial and social resources in the effective use of self-regulation skills (Mullainathan and Shafir 2014). When these resources are scarce, it is harder to use self-regulation skills. Hence, changes in available resources can lead to changes in behavior, even if there has been no change in the underlying strength of the self-regulation skill. This possibility calls into question the interpretation of some measures. For example, an item from the BRIEF-A asks respondents the extent to which they "make careless errors when completing tasks." However, all people have a tendency to focus on an unmet need, which can cause them to become distracted from other tasks (Mullainathan and Shafir 2014). For example, individuals who lack financial resources might worry about paying rent, which distracts them from their work and causes them to make more errors. Similarly, someone who receives additional financial resources might immediately make fewer mistakes and perform better on *measures* of self-regulation skills, even though his or her underlying skills did not change.

- **Background can affect how people understand the phrasing of survey questions.** Most self-reported measures of self-regulation skills were not developed specifically for low-income populations, so are not always appropriate for people in this situation. Respondents with lower levels of education might not understand some words. For example, during a pretest of measures for the evaluation of coaching interventions described in Box 1, a respondent indicated that he did not understand the word "prioritizing" when asked if he has trouble prioritizing activities. The final survey rephrased the question to ask if respondents have trouble "deciding which activities to get done first."

> Most self-reported measures of self-regulation skills were not developed specifically for low-income populations, so are not always appropriate for people in this situation.

- **Reference points can depend on background.** Recent evidence suggests that people respond to surveys based on their own reference points (Kyllonen and Bertling 2013; Primi et al. 2016). People rate their own skills relative to people they know, rather than the population as a whole. In this way, a person's reference point might depend on his or her background. *Reference bias* arises when people have different reference points when responding to survey questions. It is mostly likely to occur with questions in which response options are subjective, such as whether some-one exhibits a behavior "often," "sometimes," or "never." For example, an individual's interpretation of whether they "often" react emotionally depends on the extent to which people they know tend to react emotionally. Comparing survey responses of people with different reference points can be misleading.

- **Interventions might affect reference points.** Impact evaluations are susceptible to reference bias in a unique way: the intervention could shift the reference point of the participant group but not the comparison group. For example, a program that attempts to develop self-regulation skills might first teach about various types of self-regulation skills and help participants understand which skills they lack relative to the general population, essentially shifting their reference point. This shift could cause people to respond differently on a survey than they did previously, even if their behavior or underlying skill has not changed. Alternatively, program participants might rate themselves relative to other participants in the program, which might mask an improvement in skills if all participants' skills improved as a result of participation.

> Acknowledging that all self-regulation skills are measured using behaviors and that behaviors are influenced by a person's situation can help guide the selection of self-regulation measures to meet the goals of an evaluation.

V. Selecting Measures of Self-Regulation Skills in an Employment Evaluation

Acknowledging that all self-regulation skills are measured using behaviors and that behaviors are influenced by a person's situation can help guide the selection of self-regulation measures to meet the goals of an evaluation. Figure 2 expands Figure 1 by placing skill measurement in the context of evaluating an employment program that targets self-regulation skills. As in Figure 1, the blue shapes and arrows depict how the situation and self-regulation skills affect measures of self-regulation skills at the time of measurement. The gray arrows indicate processes that take place over time. For example, Arrows 1 and 2 show that over time an employment program could affect self-regulation skills but also could affect a participant's situation. Similarly, over time changes both in the situation and in self-regulation skills can affect employment outcomes. Conversely, changes in employment can affect both the situation and skills. For example, employment could affect a person's situation by providing additional financial resources, which in turn could allow him or her to focus more on work-related tasks. Research has also shown that employment can directly affect self-regulation skills through gained experience (Gottschalk 2005).

Figure 2.
Measurement pathways for evaluations of employment programs
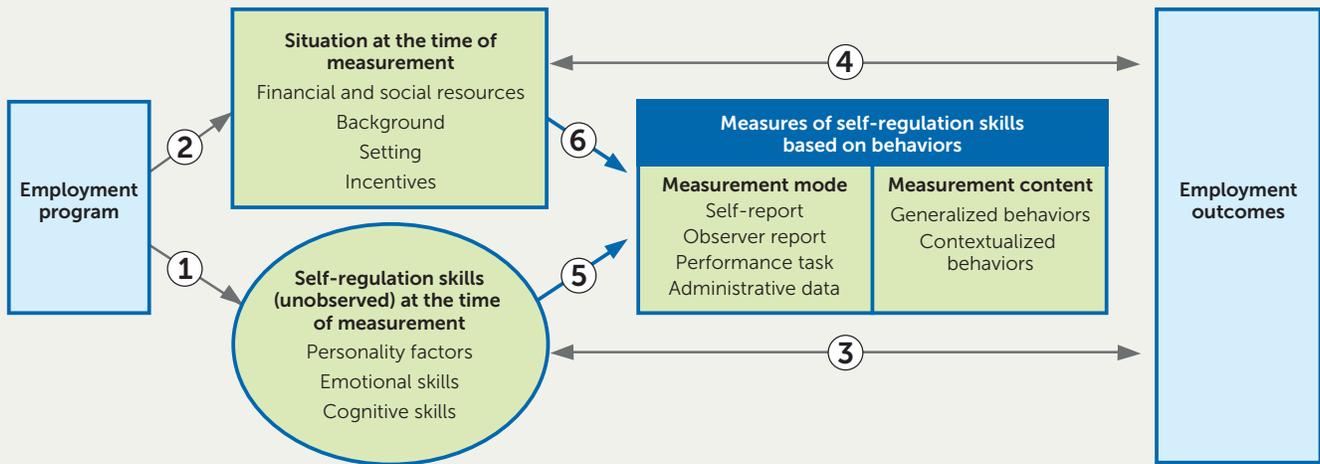
The rectangular boxes of Figure 2 indicate directly observable aspects of the ways a program may affect employment outcomes. The oval contains aspects that we cannot observe. We can observe program participation, some aspects of a person's situation, measures of self-regulation skills, and employment outcomes. Self-regulation skills are not observed directly but are captured indirectly using the measures of behaviors.

The numbered arrows indicate the pathways by which programs can affect employment outcomes. For example, Arrow 1 indicates that the program could affect self-regulation skills. Because the act of measuring behaviors does not affect employment outcomes, even if the underlying skills do, there is no arrow between the *measure* of self-regulation skills and outcomes.

The goal of an evaluation is to determine whether a program works as intended, which involves estimating some or all of the relationships depicted by the arrows in Figure 2, including:

- The impact of the program on self-regulation skills at the time of measurement (Arrow 1). Arrow 1 captures the many ways that the program could impact self-regulation skills over time, including by changing the situation. For example, a program might incentivize the use of self-regulation skills, which could cause participants to practice and develop those skills over time.

- The impact of the program on aspects of a person's situation at the time of measurement (for example, access to transportation or childcare) (Arrow 2).

- The extent to which a program affects employment outcomes by impacting self-regulation skills (Arrows 1 and 3) or by impacting a person's situation (Arrows 2 and 4).

By estimating these impacts, not only can an evaluation shed light on whether an employment program impacts key outcomes, but also whether it impacts them through the mechanism of self-regulation skills.

The following considerations illustrated by Figure 2 should be taken into account when attempting to estimate these impacts:

- **Programs can affect *measures* of self-regulation skills through their impacts on either underlying self-regulation skills or the situation.** The program could impact measures of self-regulation skills by impacting the situation at the time of measurement (Arrow 6), the underlying skills at the time of measurement, or both (Arrow 5). For example, if a program helps participants access childcare and transportation, then participants might face lower levels of stress and perform better on some measures of emotion regulation even though their underlying skills did not change. Or employment programs that offer incentives for completing job applications may lead to changes in behaviors (such as, submitting more job applications or updating resumes) because of the incentives. Alternatively, a program might help participants develop long-lasting skills that allow them to better regulate their emotions. The distinction between skills and behaviors matters because permanent changes in skills provide people with lasting benefits, whereas temporary changes in behavior due to temporary changes in situation may not (for example, if a program provides incentives or childcare support during the program but not after).[1] Thus evaluators are likely to be more interested in changes in underlying skills than temporary changes in a person's situation. An exception is if the program attempts to permanently change aspects of a person's situation. It is also possible that a program could have a lasting impact on skills by impacting behavior through the situation. For example, incentives might cause participants to develop a lasting skill through practice.

---

[1] It is also possible that a situation could boost a person's skill in the short-term, but that they could lose that skill over time once they leave the situation. However, this possibility is difficult to examine in evaluation settings.

- **Employment outcomes can affect the situation and self-regulation skills.** As discussed above, while the situation and self-regulation skills can affect employment, employment can also affect the situation and self-regulation skills. Arrows 3 and 4 are bidirectional. This bidirectional relationship makes it difficult to disentangle the direct effect of the program on skills (Arrow 1) from the impact of employment on skills (Arrow 3). The program may have improved skills which in turn increased employment, or it may have increased employment by some other mechanism (providing job leads, for example), which in turn improved skills, or some combination of both pathways. As discussed later, multiple measurements of self-regulation skills and employment outcomes over time can help to distinguish these pathways.

VI. Criteria for Selecting Measures of Self-Regulation Skills

The goals of evaluations of employment programs (as discussed above) suggest that measures of self-regulation skills should meet the following four criteria:

1. **Relate to employment outcomes.** Employment programs work to develop self-regulation skills in order to boost employment outcomes. If the measured self-regulation skills do not relate to the targeted employment outcomes (Arrow 3 in Figure 2), then the evaluation will not provide evidence on the mechanisms through which the program affects employment outcomes. For example, of the Big Five personality factors (Table 2), conscientiousness is the most correlated with job performance, whereas openness to experience is only weakly correlated with job performance (Barrick and Mount 1991), suggesting that evaluators of employment programs should prioritize measuring conscientiousness over openness to experience.

2. **Capture skills that could be influenced by the program.** Even if a skill relates to employment outcomes, it might not be sensitive to a particular program. To test whether an employment program is working as hypothesized by improving self-regulation skills (Arrow 1 in Figure 2), evaluators should select measures that capture skills that align with the goals of the program. If the program targets emotional skills but not personality factors, then measures should focus on emotion regulation rather than grit, for example.

3. **Account for aspects of the situation that could also affect the outcome measures.** As illustrated by Arrow 6 in Figure 2, the program could impact a situational factor that affects both a measure of a self-regulation skill and an employment outcome. With only the measure of the self-regulation skill, the program might appear to improve skills when it only has a short-term impact on behaviors because it just affected a participant's situation.

4. **Be feasible to administer in an evaluation.** The measures must be feasible to implement in an evaluation. Long measures or measures that require additional technology or set up may be infeasible. Short self-reported surveys are relatively easy and quick to administer compared to some types of performance tasks.

VII. Recommendations

Different types of measurement modes and content meet the four key measure selection criteria to different degrees. Table 3 outlines the pros and cons of different types of measurement modes and content, as well as some additional considerations. In particular, it highlights trade-offs between practical considerations (for example, instrument length) and technical considerations (for example, reference bias). It also shows that different approaches can complement each other, suggesting the benefits of adopting multiple approaches within a single evaluation (for example, including measures with both contextualized and generalized content). It is also important to remember that the best approach depends on the nature of the evaluation. For example, some programs might not collect administrative data or lack staff that could accurately fill out an observer report for both the program participants and a comparison group, making those modes infeasible.

Based on the tradeoffs displayed in Table 3, we offer some recommendations for selecting measures that satisfy the first three criteria listed above while considering the feasibility of measurement (the fourth criterion).

Selecting measures in the context of employment evaluations. Measures selected for an employment evaluation must be aligned with the program's goals (Criterion #1) and be relevant to employment outcomes (Criterion #2), in addition to being feasible within the context of the evaluation (Criterion #4). We recommend the following:

> A program's logic model can serve as a guide to identify which self-regulation skills or behaviors could be influenced by the program, as well as which measures are theoretically linked to outcomes.

- **Use a program's logic model to guide the selection of measures.** A program's logic model can serve as a guide to identify which self-regulation skills or behaviors could be influenced by the program, as well as which measures are theoretically linked to outcomes. One challenge is that programs that implicitly target self-regulation skills often use different language to describe them. Understanding which measures to include might require in-depth conversations with program operators. It is important to note that the empirical links between some measures of self-regulation skills and employment outcomes have not been well studied. This gap in the research requires evaluators to draw in part on theoretical considerations based on a program's logic model.

- **Consider including measures that span both generalized and contextualized measurement content.** Generalized and contextualized behaviors complement each other, as reflected in Table 3. Employment programs often focus on teaching self-regulation skills in the context of employment (for example, how to set and pursue a goal of attaining a job). Contextualized behaviors, therefore, might be more sensitive to an employment program and more directly linked to specific employment outcomes, such as securing work, suggesting some key reasons to include them in an evaluation. On the other hand, while generalized skills might be less sensitive to an employment program, they are applicable to a broader set of outcomes so might confer greater benefits in more aspects of the participant's life.

Table 3.
Trade-offs between types of measurement modes and measurement content

Approach	Examples	Pros	Cons	Other considerations
Trade-offs between measurement modes				
Self-report	Behavior Rating Inventory of Executive Function—Adult Version (BRIEF-A), self-report; Big Five Inventory (BFI)	• Existing measures linked to employment outcomes • Existing measures are relatively short and easy to administer • People might have more information about themselves than a third-party	• Potentially susceptible to reference bias • Existing measures might be too broad to be impacted by programs	• Pretest can help align the questions to target population • Surveys can be modified to address reference bias or align more closely with the intervention
Observer report	BRIEF-A, informant report; Big Five Inventory (BFI)	• Existing measures linked to employment outcomes • Can reduce reference bias relative to a self-report if raters are well trained or rate a variety of people	• High burden if program staff must administer • Third-party observers might have less information than the individuals themselves • Potentially susceptible to reference bias • Existing measures might be too broad to be impacted by programs	• Surveys can be modified to address reference bias or align more closely with the intervention
Performance tasks	Marshmallow test; digit span task; Stroop test	• Less susceptible to reference bias	• Unlikely to reflect the self-regulation skills targeted by employment programs • Can be complicated and time-consuming to administer • Little evidence on links to employment outcomes	
Administrative data	Attendance records; completion of program activities	• Existing measures linked to employment outcomes • Relatively low burden to collect • Often objective in nature • Likely to align with the nature of the intervention if collected by the program	• Might require additional validation to demonstrate how records link to existing measures of self-regulation skills	
Trade-offs between measurement content				
Generalized behaviors	• Whether someone tends to complete all tasks • Whether someone controls emotions	• Potentially applies to a wide variety of situations	• Less chance of being impacted by an intervention	
Contextualized behaviors	• Whether someone completes job applications • Whether someone controls emotions while at work	• Aligns more closely with the intervention than a generalized behavior • More directly tied to outcomes of immediate interest	• Might apply to a narrower set of outcomes	

- **Prioritize other measurement modes over performance tasks.** Although performance tasks are appealing because they offer an objective measure, they tend to: (1) be challenging to administer, (2) not have been validated using real-world employment outcomes, and (3) capture specific skills that are not targeted by existing employment programs. For these reasons, they are unlikely to be as useful in an evaluation of an employment program.

- **Conduct analyses to assess the reliability and validity of measures.** Assessing reliability and validity helps to provide evidence that measures relate to key outcomes and could be impacted by the intervention. Although many measures of self-regulation skills have been validated, few have been validated specifically for the purpose of evaluating employment programs targeted to low-income populations. This gap is potentially problematic because the reliability and validity of self-regulation measures can differ for different populations (Schmitt et al. 2007). For this reason, we suggest exploring the indicators of validity and reliability outlined in Box 2.

Box 2. Approaches to assessing reliability and validity

Predictive validity: The correlation between measures of self-regulation skills and employment outcomes. A high predictive validity indicates that the measures are linked to key outcomes.

Discriminant validity: The extent to which measures designed to capture different self-regulation skills are empirically unrelated to each other, as measured by the correlation across conceptually different measures.

Internal consistency (reliability): The extent to which items designed to capture the same self-regulation skill are consistent with each other, often summarized with a statistic like Cronbach's Alpha (Cronbach 1951). Measures with low reliability are measured with more measurement error (the component of a measure that does not reflect the underlying skill). If measures are too noisy, then evaluations might not find significant impacts in cases where the program does improve the underlying skills.

Accounting for participants' situations. Different approaches to measuring self-regulation skills account for the situation in different ways (Criterion #3). Table 4 outlines trade-offs between some of these approaches. Addressing this issue may involve lengthening the survey or interview instruments or collecting data from additional sources. We recommend:

- **Collecting some data on aspects of the situation that could be influenced by the program.** One approach to accounting for the situation is collecting data on aspects of participants' situations that could be affected by the program and hence complicate the interpretation of impact findings. For example, one potential measure of a person's self-regulation skills is whether he or she consistently shows up to job-related appointments. Showing up, however, might also depend on aspects of the situation influenced by the program, such as access to transportation or childcare. In this case, collecting data on access to transportation and childcare can help with the interpretation of any impacts on whether someone shows up to job-related appointments. If the program does not impact access to transportation or childcare but does impact whether someone shows up to appointments, it is more likely that the program improved underlying self-regulation skills.

- **Considering survey or experimental methods that explicitly account for the situation.** Surveys can explicitly account for reference bias or other aspects of the situation in several ways, including adding context to the question; using objective questions; measuring reference points by supplementing questions with anchoring vignettes; or using situational judgement tests that ask respondents how they would respond in hypothetical situations (see Table 4 for descriptions of each of these approaches). Adding context or anchoring vignettes to the questions, however, increases the time it takes to complete the survey, thereby adding more burden to participants. Similarly, some experimental designs explicitly account for the situation by randomly assigning different situations to different participants. For example, an experiment might randomly assign some participants to a coaching intervention with no incentives, others to a coaching intervention with incentives for exhibiting self-regulation behaviors, and others to a control group. This design allows evaluators to estimate the impact of the incentives (part of the situation) on measures of self-regulation, but it comes at a cost of adding an additional treatment condition.

- **Pretesting survey items that were not designed for low-income populations.** Pretesting helps ensure that survey material is appropriate for the target population's background. A pretest involves administering the survey to a small group of people who are within or similar to that target population in advance of the evaluation and then conducting in-depth interviews to identify questions that respondents did not understand or that they think could be worded more clearly. This approach helps reduce the impact that the person's background can have on responding to survey questions.

> Different measurement modes are susceptible to different types of situational biases. Using multiple modes can help to provide a more complete and robust measurement of a particular skill of interest.

- **Using multiple measurement modes when possible.** Different measurement modes are susceptible to different types of situational biases. Using multiple modes can help to provide a more complete and robust measurement of a particular skill of interest. For example, self-reports can be beneficial because people have more information about themselves than other sources. On the other hand, administrative records might be more objective and less subject to reference bias even if based on a more limited set of behaviors. Using multiple measurement modes provides confidence that the results do not reflect a bias arising from any one approach.

- **Carefully interpreting short-term impacts and collecting follow-up data on skills.** Even the most rigorous methods do not account for all aspects of a person's situation that could confound the measurement of self-regulation skills. For this reason, we suggest acknowledging this possibility when interpreting impacts of the program on self-regulation skills. In addition, collecting both short-term and long-term data can help in two ways. First, it can establish whether the program had a lasting impact on measures. Second, it can help in separating out the two-way relationship between self-regulation skills and employment outcomes illustrated by Arrow 3 in Figure 2. For example, collecting measures of self-regulation skills before and after someone is employed allows the use of statistical models to estimate the direct impact of the program on self-regulation skills and the indirect impact of the program on self-regulation skills through employment (Heckman et al. 2016). These approaches estimate how employment at a point in time affects skills at a later point in time, accounting for the possibility that employment depends on program participation and previous levels of the skills.

Table 4. Approaches to account for the situation in measuring self-regulation skills

Approach	Pros	Cons
Add context to survey questions. Adding descriptions of context to a survey question can help control for the possibility that people face different situations (for example, asking questions about how people behave in workplace settings). *Situational judgment tests* place respondents in a hypothetical situation and ask how they would behave—a way to account for situation (Kyllonen and Bertling 2013).	• Increases the predictive power for outcomes in similar context • Reduces the likelihood of reference bias	• Potentially limits the predictive power for outcomes in other contexts • Can require additional survey time
Use objective questions. Likert-type scales based on subjective responses are particularly susceptible to reference bias. Converting the question to a more objective behavior, such as the number of times the individual completed a particular action, can reduce the role of reference bias.	• Reduces the likelihood of reference bias • Places the results in more interpretable terms	• Can be challenging to accomplish for general types of behaviors (for example, whether someone completes tasks)
Pretest surveys. Testing the survey on the target population can help ensure that the content of the instrument matches the background of the target population.	• Helps ensure the survey matches the background of the respondents	• Requires additional resources to administer and additional modifications to an instrument
Use statistical methods that control for the situation. Measures of the situation can be explicitly incorporated into statistical measurement models. For example, if the behavior is attending job-related appointments, the measurement model could control for whether the individual had access to transportation.	• Reduces the role of the situation in measurement	• Requires collecting data on situations
Measure reference points. *Anchoring vignettes* are supplemental questions that provide a way to adjust the respondents' answers to questions about their own skills by "anchoring" them in their own assessment of a hypothetical person's behavior described in the survey (Kyllonen and Bertling 2013; Primi et al 2016). They can also be used to gauge whether reference points changed after the intervention.	• Can address reference bias	• Do not exist for most surveys so need to be developed • Requires additional survey time for administration
Experimentally vary the situation. One way to estimate the impact of the situation on measures of self-regulation is to experimentally vary parts of the intervention that could impact the measures of self-regulation skills independently of the underlying skills themselves. For example, an experimental evaluation could include two treatment arms, one that includes incentives for exhibiting particular behaviors and one that does not.	• Provides the strongest basis for disentangling the role of a certain aspect or aspects of the situation in an evaluation • Sheds light on which aspects of a program are effective	• Requires an additional treatment group in the experiment, necessitating a larger sample size to detect the same level of impact

VIII. Conclusions

Success in employment depends on self-regulation skills (Almlund et al. 2011). For this reason, self-regulation skills have become a key focus of some employment programs and evaluations of these programs. Several challenges arise when measuring self-regulation skills in the context of employment evaluations for low-income populations: (1) measures of self-regulation skills can reflect aspects of a person's situation in addition to his or her skills; (2) most existing measures were developed for purposes other than program evaluation; (3) most existing measures were not designed for use with low-income populations; and (4) some current measures might place too much burden on programs and/or participants. To address these challenges, we suggest using a set of both general measures of self-regulation as well as ones that are specific to the employment context, collecting information on other aspects of the participants' situations that can be affected by the program, modifying measures to fit the target population, and conducting analyses to assess the reliability and validity of selected measures.

References

Almlund, M., A. Duckworth, J.J. Heckman, and T. Kautz. "Personality Psychology and Economics." In *Handbook of the Economics of Education*, edited by E. A. Hanushek, S. Machin, and L. Woessmann (Vol. 4, pp. 1–181). Amsterdam: Elsevier, 2011.

Barrick, M.R., and M.K. Mount. "The Big Five Personality Dimensions and Job Performance: A Meta-Analysis." *Personnel Psychology*, vol. 44, no. 1, 1999, pp. 1–26.

Barton, P.E. (2006). "High School Reform and Work: Facing Labor Market Realities." Princeton, NJ: Policy Evaluation and Research Center, Policy Information Center, Educational Testing Service, 2006.

Caliendo, M., D.A. Cobb-Clark, and A. Uhlendorff. "Locus of Control and Job Search Strategies." *Review of Economics and Statistics*, vol. 97, no. 1, 2015, pp. 88–103.

Cavadel, E.W., J.F. Kauff, M.A. Anderson, S. McConnell, and M. Derr. "Self-Regulation and Goal Attainment: A New Perspective for Employment Programs." OPRE Report #2017-12, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2016.

Cavadel, E.W., J.F. Kauff, A. Person, and T.K. Kravis. "Perspectives on Practice: A Guide to Measuring Self-Regulation and Goal-Related Outcomes in Employment Programs." OPRE Report #2018-37. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2018.

Cunningham, W. V., and Villaseñor, P. (2016). "Employer voices, employer demands, and implications for public skills development policy connecting the labor and education sectors." *The World Bank Research Observer*, vol. 31, no. 1, 2016, pp. 102–134.

Cronbach, L.J. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika*, vol. 16, no 3, 1951, pp. 297–334.

Duckworth, A.L., and D.S. Yeager. "Measurement Matters: Assessing Personal Qualities Other than Cognitive Ability for Educational Purposes." *Educational Researcher*, vol. 44, no. 4, 2015, pp. 237–251.

Duckworth, A.L., and P.D. Quinn. "Development and validation of the Short Grit Scale (GRIT–S)." *Journal of Personality Assessment*, vol. 91, no. 2, 2009, pp. 166–174.

Gallo, W.T., J. Endrass, E.H. Bradley, D. Hell, and S.V. Kasl. "The Influence of Internal Control on the Employment Status of German Workers." *Schmollers Jahrbuch*, vol. 123, no. 1, 2003, pp. 71–81.

Gottschalk, P. "Can Work Alter Welfare Recipients' Beliefs?" *Journal of Policy Analysis and Management*, vol. 24, no. 3, 2005, pp. 485–498.

Grant, A.M., and M.J. Cavanagh. "Coaching and Positive Psychology" in *Designing Positive Psychology: Taking Stock and Moving Forward*, edited by K.M. Sheldon, T. Kashdan, and M. F. Steger. New York: Oxford University Press, 2011.

Heckman, J.J., and T. Kautz. "Hard Evidence on Soft Skills." *Labour Economics*, vol. 19, no. 4, 2012, pp. 451–464.

Heckman, J. J., Humphries, J. E., and Veramendi, G. Dynamic treatment effects. Journal of econometrics, vol. 191, no. 2, 2016, pp. 276–292.

Holzer, H. "Is There a Gap Between Employer Skill Needs and the Skills of the Work Force?" In *Transitions in Work and Learning: Implications for Assessment*, edited by A. Lesgold, M. Feuer, and A. Black. Washington, DC: National Academy Press, 1997.

Hogan, J., and B. Holland. "Using Theory to Evaluate Personality and Job-Performance Relations: A Socioanalytic Perspective." *Journal of Applied Psychology*, vol. 88, no. 1, 2003, pp. 100–112.

Jackson, C.K. "What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes." NBER Working Paper No. 22226. Cambridge, MA: National Bureau of Economic Research, 2016.

John, O.P., E.M. Donahue, and R.L. Kentle. *The Big Five Inventory—Versions 4a and 54*. Berkeley, CA: University of California Berkeley, 1991.

Kautz, T., J.J. Heckman, R. Diris, B. ter Weel, and L. Borghans. (2014). "Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success." OECD Education Working Papers No. 110. 2014. Available at https://www.oecd.org/edu/ceri/Fostering-and-Measuring-Skills-Improving-Cognitive-and-Non-Cognitive-Skills-to-Promote-Lifetime-Success.pdf. Accessed November 1, 2017.

Kautz, T., and W. Zanoni. "Measuring and Fostering Non-Cognitive Skills in Adolescents: Evidence from Chicago Public Schools and the OneGoal Program." Chicago: University of Chicago, Department of Economics, unpublished manuscript, 2015.

Kyllonen, P.C., and J.P. Bertling. "Innovative Questionnaire Assessment Methods to Increase Cross-Country Comparability." In *Handbook of International Large–Scale Assessment: Background, Technical Issues, and Methods of Data Analysis*, edited by L. Rutkowski, M. von Davier, and D. Rutkowski. Boca Raton, FL: CRC Press, 2013.

Mischel, W., and E.B. Ebbesen. "Attention in Delay of Gratification." *Journal of Personality and Social Psychology*, vol. 16, no. 2, 1970, pp. 329–337.

Mullainathan, S., and Shafir, E. "Freeing up intelligence." *Scientific American Mind*, vol. 21, no. 1, 2014, pp. 58–63.

Nyhus, E.K., and E. Pons. "The Effects of Personality on Earnings." *Journal of Economic Psychology*, vol. 26, no. 3, 2005, pp. 363–384.

Ownsworth, T. L., McFarland, K., and Young, R. M. Development and standardization of the Self-regulation Skills Interview (SRSI): A new clinical assessment tool for acquired brain injury. *The Clinical Neuropsychologist*, vol. 14, no. 1, 2000, 76–92.

Primi, R., C. Zanon, D. Santos, F. De Fruyt, and O.P. John. "Anchoring Vignettes." *European Journal of Psychological Assessment*, vol. 32, no. 1, 2016, pp. 39–51.

Roberts, B., N.R. Kuncel, R. Shiner, A. Caspi, and L.R. Goldberg. "The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes." *Perspectives on Psychological Science*, vol. 2, no. 4, 2007, pp. 313–345.

Roberts, B.W., K.E. Walton, and W. Viechtbauer. "Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies." *Psychological Bulletin*, vol. 132, no. 1, 2006, p. 1–25.

Roberts, B.W., and D. Mroczek. "Personality Trait Change in Adulthood." *Current Directions in Psychological Science*, vol. 17, no. 1, 2008, pp. 31–35.

Roth, R.M., and G.A. Gioia. *Behavior Rating Inventory of Executive Function—Adult Version*. Lutz, FL: Psychological Assessment Resources, 2005.

Salgado, J.F. (1997). "The Five Factor Model of Personality and Job Performance in the European Community." *Journal of Applied Psychology*, vol. 82, no. 1, 1997, pp. 30–43.

Schmidt, F.L., and J. Hunter. "General Mental Ability in the World of Work: Occupational Attainment and Job Performance." *Journal of Personality and Social Psychology*, vol. 86, no. 1, 2004, pp. 162–173.

Schmitt D.P., J. Allik, R.R. McCrae, V. Benet-Martínez, L. Alcalay, and L. Ault. "The Geographic Distribution of Big Five Personality Traits: Patterns and Profiles of Human Self Description Across 56 Nations." *Journal of Cross-Cultural Psychology*, vol. 38, 2007, pp. 173–212.

Secretary's Commission on Achieving Necessary Skills. "Learning a Living: A Blueprint for High Performance: A SCANS Report for America 2000." Washington, DC: United States Government Printing Office, 1992.

Steinberg, L. "A Social Neuroscience Perspective on Adolescent Risk-Taking." *Developmental Review*, vol. 28, no. 1, 2008, pp. 78–106.

Störmer, S., and R. Fahr. "Individual Determinants of Work Attendance: Evidence on the Role of Personality." *Applied Economics*, vol. 45, no. 19, 2013, pp. 2863–2875.

Stroop, J.R. "Studies of Interference in Serial Verbal Reactions." *Journal of Experimental Psychology*, vol. 18, no. 6, 1935, p. 643.

Wechsler, D. *Wechsler Intelligence Scale for Children*. 1949.

West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F., and Gabrieli, J. D. "Promise and paradox: Measuring students' non-cognitive skills and the impact of schooling." *Educational Evaluation and Policy Analysis*, vol. 38, no. 1, 2016, pp. 148–170.