Chapter 21
Improving Learning Through Stealth Assessment of Conscientiousness

Gregory R. Moore and Valerie J. Shute

Abstract In this chapter, we describe the importance of assessing and developing conscientiousness in students and how we are approaching this challenge. After discussing the benefits conscientiousness has for learning, we describe the process we are using to create a valid stealth assessment of conscientiousness. We then discuss the current state of this work and suggest next steps and areas of future research around conscientiousness. Finally, we broaden our scope to discuss the strengths and limitations of using stealth assessment to measure noncognitive competencies, as well as give some recommendations to help others use this approach. Our hope is that this chapter will demonstrate both (a) the importance and complexity of conscientiousness measurement in educational settings, and (b) a general process for thinking about and designing assessments for noncognitive competencies.

Keywords Conscientiousness • Personality • Stealth assessment • Game-based learning • Noncognitive competencies • Learning • Assessment • Students • Educational environments • Reflection • Assessment design

To succeed in modern society, students need to develop a wide variety of competencies (Partnership for 21st Century Learning, 2015), which are the knowledge, skills, and attributes that impact life outcomes. These competencies can broadly be divided into cognitive (e.g., math and verbal proficiency, problem solving, and reasoning) and noncognitive competencies (e.g., personality factors, collaboration, motivation). While cognitive competencies receive a lot of attention in the education literature, both are important for learning. Personality is one branch of noncognitive competencies. Previous research suggests that personality impacts many different life outcomes, including academic achievement and workplace success (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). The Five-Factor Model of Personality has been a popular way to conceptualise personality for many years (McCrae & Costa, 1987). This model defines five broad categories of personality: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness. Of these personality factors, conscientiousness appears to be particularly important in education.

G.R. Moore • V.J. Shute (✉)
Florida State University, Tallahassee, FL, USA
e-mail: vshute@admin.fsu.edu

© Springer International Publishing Switzerland 2017

A. Marcus-Quinn, T. Hourigan (eds.), *Handbook on Digital Learning for K-12 Schools*, DOI 10.1007/978-3-319-33808-8_21

vshute@admin.fsu.edu

Conscientiousness and Learning

Conscientiousness is a multifaceted construct that can broadly be described as the willingness to work hard and carefully. Attempts to define the precise facets of conscientiousness have resulted in a variety of factorizations of the construct. However, certain facets are consistently found in the literature: Persistence (sometimes referred to as industriousness or perseverance), Organization (sometimes referred to as order/orderliness), Carefulness (sometimes referred to as cautiousness or self-control), and Dependability (sometimes referred to as reliability or responsibility). Furthermore, research on perfectionism suggests it is also related to conscientiousness (Parker, 1997). Indeed recent factorizations of conscientiousness have included it as facet (e.g., MacCann, Duckworth, & Roberts, 2009). We therefore conceptualise conscientiousness with five facets: the four found in many factorizations of the construct (i.e., Persistence, Organization, Carefulness, Dependability) plus Perfectionism.

Research on the specific effects of conscientiousness suggests that it has positive effects on academic achievement independent from other predictors of academic achievement, such as past performance (O'Connor & Paunonen, 2007; Poropat, 2009). In fact, conscientiousness may predict achievement as much as intelligence (Poropat, 2009). Conscientiousness is also associated with higher effort (Noftle & Robins, 2007), improved learning motivation (Colquitt & Simmering, 1998), self-regulation (Abe, 2005), higher perceived ability (Noftle & Robins, 2007), fewer behavioral problems (Abe, 2005), and achievement learning orientations (Chamorro-Premuzic, Furnham, & Lewis, 2007). The relationships between academic outcomes and conscientiousness are seen throughout the lifespan, from early childhood and adolescence (Abe, 2005; Drake & Belsky, 2014) to post-secondary education (O'Connor & Paunonen, 2007; Poropat, 2009; Trapmann, Hell, Hirn, & Schuler, 2007). There are also relationships reported between conscientiousness and various workplace outcomes (e.g., Bajor & Baltes, 2003; Bakker, Demerouti, & ten Brummelhuis, 2012; Dudley, Orvis, Lebiecki, & Cortina, 2006).

The persistence facet of conscientiousness seems to be particularly important in education. Perry, Hunter, Witt, and Harris (2010) suggested that persistence, which they called *achievement*, drives the ability of conscientiousness to predict performance. Additionally, grit—a construct that can be considered as a combination of passion and persistence—has been found to predict a variety of learning and performance outcomes, including GPA, educational attainment, student retention, and spelling bee performance (Duckworth, Peterson, Matthews, & Kelly, 2007). Grit is also independent from intelligence.

The perfectionism facet has important implications for learning and performance as well. Research suggests that there are different types of perfectionism, which differently impact learning and performance. Hamachek (1978) suggested that can perfectionism can be broken down into two types. Normal perfectionists set high, but realistic, expectations of themselves, find enjoyment in their work, and are capable of accepting less than perfection. Neurotic perfectionists, though, set unrealistic expectations of themselves, struggle to find enjoyment in their work, and struggle to

vshute@admin.fsu.edu

accept less than perfection. Throughout the years, researchers have tended to conceptualise perfectionism in a similar way, though with variations in terminology. For example, Terry-Short, Owens, Slade, and Dewey (1995) used the terms positive and negative perfectionism and Stoeber and Otto (2006) used the terms healthy and unhealthy perfectionism. In a similar vein, Parker's (1997) typology of perfectionism classified students into one of three categories: nonperfectionists (characterised by low standards and carelessness), healthy perfectionists (characterised by little fear of mistakes and good organization), and dysfunctional perfectionists (characterised by worrying about mistakes). In their review of the perfectionism literature, Stoeber and Otto (2006) suggested that healthy perfectionists achieve more, are more satisfied, have improved social skills, and are able to adapt to new situations better than unhealthy perfectionists. Furthermore, healthy perfectionists are less likely to experience anxiety, depression, and procrastination. Thus, encouraging the right type of perfectionism can help students in academic contexts and in life in general.

For all of these reasons, we want to help students develop conscientiousness, both globally and at the facet level, and provide support for students low in conscientiousness. To accomplish this goal, we need to accurately measure conscientiousness and its facets. While previous research on conscientiousness used self-report measures, these are problematic for three main reasons. First, people are often not able to accurately evaluate themselves, as it requires a level of self-knowledge that they may not have. Second, and similarly, respondents may interpret items differently. For example, if two people are rating their agreement with the item "I am tidy," they may have two different understandings of what it means to be tidy, which threatens the validity of the measure. Third, people tend to fall victim to the social desirability effect (Paulhus, 1991), presenting themselves more positively than they really are and/or more in line with what they believe the researchers wants to see. To resolve these issues and more accurately measure conscientiousness, we have been developing a stealth assessment of conscientiousness that can run invisibly in a gaming environment. The stealth assessment, and the process we are using to create it, are described next.

A Stealth Assessment of Conscientiousness

Stealth Assessment and Evidence-Centered Design

Stealth assessments are embedded in digital games such that they are unobtrusive to the learner being assessed (Shute, 2011; Shute & Ventura, 2013). These assessments use the learner's in-game actions to develop a belief about a student's level on a certain competency, and this belief about the student is updated over time to more accurately reflect her or his knowledge, skills, and attributes. Stealth assessment is based on the Evidence-Centered Design framework, or ECD (Mislevy, Steinberg, & Almond, 2003). This framework defines an approach for developing valid

vshute@admin.fsu.edu

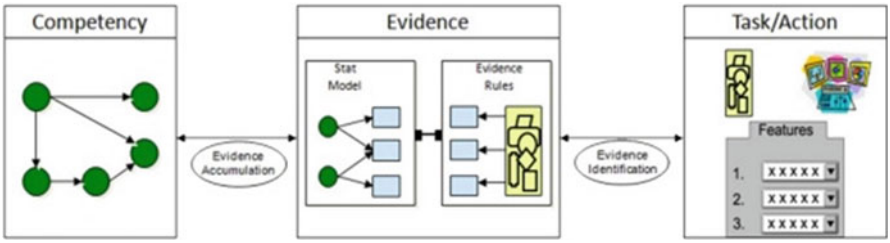

Fig. 21.1 A graphical representation of the ECD framework (From Mislevy et al., 2003)

assessments and is comprised of three models (Fig. 21.1). The competency model (CM) defines the competency or competencies of interest (e.g., algebra knowledge, problem solving skill, creativity, conscientiousness). The task model (TM) defines the features of the task environment that will elicit evidence of the competency or competencies of interest. The evidence model (EM) defines what constitutes evidence of the competency and acts as the statistical "glue" between the competency and task models. The evidence model is comprised of two parts. The evidence rules take the stream of data from the task environment (e.g., student actions) and convert it to observable variables. The statistical model specifies the relationships between these observable variables and the competency variable(s). This framework facilitates the development a valid evidence chain from the competency variables (CM) to observable variables (EM) to in-game actions (TM).

To implement the competency and evidence models for our stealth assessments, we use Bayesian networks (for more information on the mathematics of Bayesian networks and the range of applications in education, see Almond, Mislevy, Steinberg, Yan, & Williamson, 2015). These networks graphically represent the conditional dependencies among the competency variables and the observed variables. One of the benefits of Bayesian networks is that they allow us to accumulate evidence and update our beliefs over time. In our stealth assessments, as students play the game, they provide a stream of data that is analyzed in real-time. In turn, the system's beliefs about the competency variables are also updated in real-time. This accumulation allows us to obtain progressively more accurate measurements of competency variables as time goes on. Thus, stealth assessment allows for valid, real-time, unobtrusive assessments of students and avoids the weaknesses of self-report measures.

We use games as our assessment vehicles for a few reasons. First, games, and other computer-based learning environments, can automatically log student actions, which allows us to collect a lot of data at a fine-grained scale. This helps us to make valid inferences. Second, games are becoming important tools for learning and teaching. Research suggests that games, when properly applied, can improve learning outcomes (Wilson et al., 2009) and help develop twenty-first century skills, such as problem solving (Gee, 2007; Shute, Ventura, & Ke, 2015). Therefore, we believe that games can act as learning and assessment environments at the same time. Third, games are very popular and engaging, especially for young adults (Lenhart et al., 2008). People play games for their own sake and find their contexts meaningful,

vshute@admin.fsu.edu

which improves the validity of the measurements. Additionally, the engagement that games offer makes the assessment unobtrusive, which can reduce test anxiety. However, while games are excellent assessment vehicles, care still needs to be taken to select or develop a game that meets the needs of each individual assessment. We describe the game we selected for our assessment of conscientiousness, as well as the reasons for that choice, next.

Physics Playground

We elected to implement a stealth assessment for conscientiousness in the game *Physics Playground* (formally called *Newton's Playground*; Shute & Ventura, 2013; Shute, Ventura, & Kim, 2013). *Physics Playground* (Fig. 21.2) is a two-dimensional puzzle game designed to help students in middle school and high school develop their conceptual understanding of physics. In the game, players attempt to move a green ball to a red balloon, primarily by creating and using various agents of force and motion (i.e., ramps, levers, pendulums, and springboards). To create the physics agents, player must draw them on the screen using the mouse. Once drawn, these objects come to life and behave in accordance with various physics principles, such as Newton's Three Laws.

Physics Playground features 74 levels spread across seven playgrounds with increasing difficulty. Success in each level is two-tiered and based on the number of objects players create to complete the level. Players who simply beat the level get a

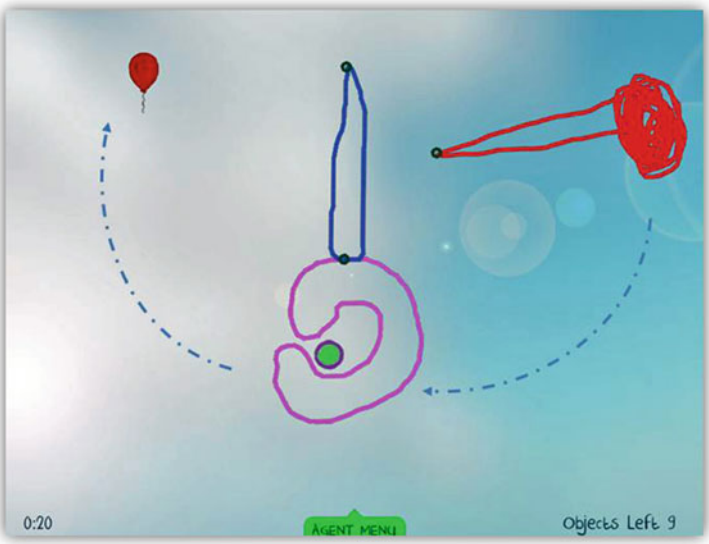

Fig. 21.2 An example of a level in *Physics Playground*. The player has drawn a pendulum object to hit a second pendulum and send the ball up to the balloon

vshute@admin.fsu.edu

silver trophy, while players who solve the level with fewer than the par number of objects get a gold trophy. Gold trophies indicate mastery of the agent(s) the player used in the level (e.g., mastery of pendulums, which in turn allows for inferences about mastery of relevant physics concepts). Silver trophies indicate that the player may not have mastery of the agent(s). Before starting the game, players view tutorial videos that teach them how to draw and use each of the agents of force and motion. These videos can be viewed again at any time.

There are two main reasons that we chose to use this game to develop our conscientiousness assessment. First, it was a game that we developed internally. Therefore, we could implement the stealth assessments and make changes to the game without restriction, which makes the process much easier. In particular, this allowed us to implement a log file system that facilitates data collection and organization. These log files record each action as well as a variety of information associated with each action, such as the time of the action, the number of restarts of the level, the number of objects used, and the coordinates and trajectory of the ball. These pieces of data are all used to update the assessment's beliefs about the player's conceptual physics understanding, as well as other competencies.

Second, players have the ability to exhibit most of the facets of conscientiousness in *Physics Playground*. Players can demonstrate tenacity on particularly challenging levels (persistence), make plans before redoing a failed level (organization), think carefully about each game action (carefulness), and go for gold trophies (perfectionism). This allows us to collect data from the game that accurately represents the student's level of conscientiousness. However, it is worth noting that the dependability facet does not lend itself well to a game-based stealth assessment of conscientiousness. Dependability refers to actions such as doing your work on time and keeping your promises. These are actions that either are not measurable automatically in a gaming environment or are confounded by other competencies. For example, completing work on time is confounded with ability in *Physics Playground*. This said, we believe that our stealth assessment can accurately measure conscientiousness with indicators of the four other facets. Our working competency model of conscientiousness is described next.

Competency Model of Conscientiousness

As discussed previously, dependability does not lend itself to a game-based assessment of conscientiousness. Thus, we did not include it in the competency model for our stealth assessment. We started by working with a four-factor model of conscientiousness: persistence, organization, carefulness, and perfectionism. However, the facets of conscientiousness are not directly observable. We therefore needed to develop a variety of in-game indicators for each of the facets. To determine these indicators, we went back to the literature on each of the facets. For example, in the persistence literature, persistence has often been measured by time spent on and attempts to solve very hard or impossible tasks (Eisenberger & Leonard, 1980; Feather, 1961). Thus, we developed the following indicators for persistence: *Time*

vshute@admin.fsu.edu

Spent on Unsolved Problems, *Number of Level Restarts*, and *Number of Level Revisits*. We also explored existing, validated measures of each of the facets, such as items in the International Personality Item Pool (Goldberg, 1999). While these previous measures were all self-reports, they gave us ideas about what types of indicators would measure the facet of interest and be appropriate for the game.

However, while defining the indicators, we discovered that the indicators of organization and the indicators of carefulness tended to be the same. For instance, the average number of objects drawn per level (reverse-coded) was an indicator of organization (i.e., organised players have a plan and don't need to draw many extra objects) and of carefulness (i.e., careful players think carefully about when and where to draw an object, and draw fewer objects as a result). Therefore, we decided to combine the organization and carefulness facets into a single facet, called *carefulness*, for the competency model of the conscientiousness stealth assessment. This three-factor competency model, with the complete list of indicators for all facets, can be seen in Table 21.1.

Next Steps

Our next steps in this project are to implement the conscientiousness assessment into Physics Playground, to validate the assessment, and to adjust the assessment as needed. Implementing the assessment will require us to build the Bayesian networks (with conditional probability tables) and embed them into the game. Then, we will need to conduct a pilot study to test the validity of our stealth assessment. To do this, we will have students play Physics Playground with the stealth assessment embedded and complete an external measure of conscientiousness. We can calculate the correlation between the stealth assessment and the external measure to

Table 21.1 The competency model of conscientiousness used for our stealth assessment where [R] refers to indicators that are reverse-coded

Competency	Facets	Indicators
Conscientiousness	Persistence	Time on unsolved problems
		Number of restarts on unsolved problems
		Number of revisits to unsolved problems
	Carefulness	Number of object limits reached in a problem [R]
		Average number of objects drawn per level [R]
		Average time, in seconds, spent drawing per object
		Average time, in seconds, between actions
		Average number of seconds before making an action on the first attempt
	Perfectionism	Ratio of gold trophy solutions to silver trophy solutions
		The number of revisits to levels with a silver trophy
		The number of restarts of a level in less than 3 actions

vshute@admin.fsu.edu

determine validity. However, measures of conscientiousness tend to be self-reports, so we will likely be comparing the stealth assessment to self-report measures. Due to the problems with self-report, mentioned earlier, we expect that there will be small to moderate relationships between the measures. This validity study will also allow us to determine whether our competency model or Bayesian networks need to be adjusted (e.g., altering the difficulty or discrimination parameters in the Bayes nets based on pilot data). After demonstrating the validity of the stealth assessment of conscientiousness, we will be able to measure conscientiousness accurately and, in turn, conduct meaningful research on it. Our ideas for future research on conscientiousness are described next.

Future Directions for Conscientiousness Research

There are three main directions in which research on conscientiousness needs to go. In this section, we discuss these different directions and how the implementation of a game-based stealth assessment can benefit them. We start by discussing how we might develop conscientiousness in students. We then describe how instruction might be designed to adapt to different levels of conscientiousness. Finally, we focus on the persistence facet of conscientiousness to explore whether or not it is a state or a trait and how that distinction impacts education practice.

Developing Conscientiousness

While personality traits, such as conscientiousness, are often considered to be relatively stable, previous research suggests that people can, in fact, learn to become more conscientious over time. The work of Eisenberger (1992) suggests that persistence, which he calls *industriousness*, is learned over time based on how a person is rewarded for their effort. That is, persistence may become generalised, as effort rewarded in one environment often impacts persistence in other environments. Thus, beneficial personality traits can be developed, and unhealthy aspects can potentially be attenuated.

Since conscientiousness predicts a variety of academic outcomes, helping students to develop conscientiousness should be a focus of future research. This research needs to examine the particular types of interventions and/or contexts that will help students to efficiently and effectively develop their conscientiousness. For example, perhaps requiring middle school students to use agendas will help them develop conscientiousness. Or perhaps certain teaching methods develop conscientiousness better than others. This type of research will likely need to be longitudinal, checking in on the sample of students periodically to measure conscientiousness over time. Our proposed stealth assessment of conscientiousness may be used to conduct these periodic assessments.

vshute@admin.fsu.edu

Adapting to Different Levels of Conscientiousness

Another branch of future conscientiousness research concerns the ability to adapt instruction based on the learner's estimated level of conscientiousness. This is particularly important if developing conscientiousness through direct instruction isn't possible or practical. Some prior research suggests that conscientiousness-based adaptation is useful. For example, Cheramie and Simmering (2010) found that people lacking in conscientiousness need more accountability in their learning. It therefore seems likely that people with high and low conscientiousness levels will thrive under different situations and with different forms of support. Future research should clarify what situations and types of support specifically help low conscientiousness students and what situations and types of support specifically help high conscientiousness students.

The stealth assessment proposed here can measure conscientiousness, which can in turn be used by educators to tailor instruction to individuals' needs. More interestingly, the conscientiousness stealth assessment can also be used to drive real-time adaptivity in a computer learning environment, such as a game. Since the stealth assessment runs in real-time, the learning environment can change on the fly based on the system's current beliefs about the student. For instance, if the stealth assessment determines that the student is too low in conscientiousness (based, perhaps, on some cut-score), the game might add in specific goals that the player needs to complete to increase accountability. The feedback and support can be applied as soon as possible. This contrasts with nonadaptive educational environments, where feedback and support only appear after the fact. For more on adaptivity in educational contexts, see Shute and Zapata-Rivera (2012).

Persistence: State or Trait?

It is often assumed that some people are more persistent than others, and that this is reflective of differences in personality. However, research suggests that external factors also matter when it comes to persistence. For example, both Feather (1961) and Eisenberger and Leonard (1980) found that a person's expectations of success influenced how long they would stick with a task. Thus, whether or not someone persists at a task is not just a function of their personality. It is also a function of the nature of the task. This makes intuitive sense, as people may persist longer on tasks they find worthwhile, interesting, and/or feasible.

However, this begs the question of whether or not we should conceptualise persistence as a state, a trait, or both. Future research should explore in what ways persistence is a state and in what ways it is a trait. In particular, this research may need to examine when it is useful to think about persistence as a state and when it is useful to think about it as a trait. In this regard, it might make sense to distinguish between persistent behavior (a state) and a persistent disposition (a personality

vshute@admin.fsu.edu

trait). This conceptualization matters because it impacts how we approach helping students. If we think about persistence as a state, then we might focus on designing our educational environments to encourage persistence. On the other hand, if we think about persistence as a trait, then we might focus on training students to be conscientious. Of course, it is possible that we need to think of persistence in both ways. For instance, perhaps we should design environments that reward effort to encourage persistent behavior (state), which in turn can develop a persistent disposition (trait), as Eisenberger (1992) suggests. By validly measuring persistence with our conscientiousness stealth assessment, researchers can more fully explore the nature of persistence.

Stealth Assessment and Other Noncognitive Competencies

Throughout this chapter, we have discussed the importance of conscientiousness, how we are measuring it, and what still needs to be learned about it. However, we also wanted our work on conscientiousness to serve as an exemplar of how to develop a valid assessment of a noncognitive competency. We conclude this chapter with a discussion of the strengths and limitations of the stealth assessment of noncognitive competencies, as well as some specific suggestions for how to use stealth assessment, so that other researchers can effectively use it to meet their needs.

Strengths and Limitations of Stealth Assessment

The strengths of using stealth assessments to measure noncognitive competencies have been described throughout this chapter. For clarity, we briefly state them again here. First, stealth assessments are unobtrusive such that the learner does not know that they are being assessed. The distinction between learning and assessment is completely blurred. This reduces the saliency of the assessment, minimizing test anxiety and consequently improving validity. Second, stealth assessments are based on a student's actions, which counters the aforementioned issues inherent in self-reports (e.g., social desirability effects). Third, stealth assessments run continuously while students play the game. This facilitates the examination of noncognitive competencies over time and allows the system to adapt to the student on the fly. Fourth, stealth assessments are embedded in educationally relevant environments. Thus, noncognitive competencies can be assessed while students are learning. In summary, stealth assessments are efficient, valid, and do not disrupt learning.

However, it is also important to acknowledge the limitations of stealth assessment. For one, stealth assessment can be a difficult and time consuming process. It requires researchers to engage in a variety of processes, which include thoroughly reviewing the literature, developing competency, evidence, and task models, embedding the stealth assessment into a gaming environment, and refining the assessment through validity and other pilot studies. It is also requires a variety of skill sets, including

vshute@admin.fsu.edu

those of educators, psychometricians, and computer scientists. Therefore, before deciding to use stealth assessment, researchers must ensure that they have the time, resources, and skills to successfully complete the process.

Moreover, stealth assessment may not be appropriate for measuring all noncognitive competencies. For example, while we consider dependability to be a facet of conscientiousness, we could not find a valid way to measure it in Physics Playground. Therefore, it ended up not being a factor in our implemented assessment. Thus, researchers looking to use stealth assessment need to carefully think about whether or not it is appropriate for the construct they are studying. If researchers decide that stealth assessment is appropriate, then they need to carefully select the right learning environment/game for the construct they are studying. There needs to be alignment between the environment and what is being measured. For example, while Physics Playground is a great environment for assessing physics knowledge, problem solving, and conscientiousness, it would not be suitable to measure noncognitive competencies such as leadership and communication, at least without modification.

Additionally, researchers need to choose a game that they can modify, whether it is a game made in house or obtained from a third party. The reason for this is that stealth assessment requires the researcher to embed data collection directly into the source code of the game. If this access cannot be achieved, stealth assessment cannot be used. We have achieved some success using our own game (Shute et al., 2013) and through partnerships with game developers (Shute et al., 2015) in past work, so this an obstacle that can be overcome with good planning and preparation.

We present these limitations here to give readers a better understanding of when stealth assessment might be appropriate, what types of collaborations are necessary, and problems they may encounter. Despite these limitations, we believe that stealth assessment is useful in a wide variety of situations. We next present some practical tips for those looking to use stealth assessments in their work.

Stealth Assessment Advice

First, researchers should take care when creating log files for a game or adapting log files from an existing game. The log files need to be simple and easy to parse. If they are not, creating an organised assessment can become very challenging. This a particular concern when adapting a third party's log files. Typically, these log files were never intended for assessment, so they will likely contain extraneous information and can be difficult for the researcher to understand. Researchers will need to work with the third party to adapt the logging system to their needs, or potentially create a new logging system.

Second, and as briefly discussed above, demonstrating validity is an important part of developing a stealth assessment. This can be challenging because external measures that address the same construct as the stealth assessment may or may not have good alignment. For example, conscientiousness is typically measured with self-report measures. However, the stealth assessment described in this chapter is expected to be an improvement over self-report measures. Thus, these measures

vshute@admin.fsu.edu

have poor alignment and we do not expect (nor do we especially desire) that our stealth assessment measures will be highly correlated with the self-report measures. We expect small to moderate correlations at best. When possible, though, it is important to make sure that there is alignment the external measure and the stealth assessment. For example, in our previous work on persistence (Ventura & Shute, 2013), we used a performance-based measure of persistence—the amount of time spent on an impossible task—in addition to self-report measures to validate our persistence stealth assessment. This allowed us to be confident that our assessment was valid. For more detailed stealth assessment recommendations, see Wang, Shute, and Moore (2015).

Conclusion

In this chapter, we demonstrate the need to measure conscientiousness in educational settings and discussed the complexity inherent in this task. We hope that this shows (a) educators the importance of conscientiousness, and (b) researchers the need for more thorough research on conscientiousness. The stealth assessment presented here acts as a jumping off point for future research on conscientiousness. However, we also hope that this chapter encourages researchers to examine stealth assessment as a potential means of collecting valid information on a variety of noncognitive competencies and gives them ideas about how to get started. Stealth assessment allows for more valid data collection than the self-report measures used for many noncognitive competencies. With more valid data, we can develop a deeper understanding of the skills and traits that impact learning and, in turn, help students of all ages improve their learning outcomes.

References

Abe, J. A. (2005). The predictive validity of the Five Factor Model of personality with preschool age children: A nine year follow-up study. *Journal of Research in Personality, 39*, 423–442. doi:10.1016/j.jrp.2004.05.002.

Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian networks in educational assessment*. New York, NY: Springer.

Bajor, J. K., & Baltes, B. B. (2003). The relationship between selection optimization with compensation, conscientiousness, motivation, and performance. *Journal of Vocational Behavior, 63*, 347–367. doi:10.1016/S0001-8791(02)00035-0.

Bakker, A. B., Demerouti, E., & ten Brummelhuis, L. L. (2012). Work engagement, performance, and active learning: The role of conscientiousness. *Journal of Vocational Behavior, 80*, 555–564. doi:10.1016/j.jvb.2011.08.008.

Chamorro-Premuzic, T., Furnham, A., & Lewis, M. (2007). Personality and approaches to learning predict preference for different teaching methods. *Learning and Individual Differences, 17*, 241–250. doi:10.1016/j.lindif.2006.12.001.

vshute@admin.fsu.edu

Cheramie, R. A., & Simmering, M. J. (2010). Improving individual learning for trainees with low conscientiousness. *Journal of Managerial Psychology, 25*(1), 44–57. doi:10.1108/02683941011013867.

Colquitt, J. A., & Simmering, M. J. (1998). Conscientiousness, goal orientation, and motivation to learn during the learning process: A longitudinal study. *Journal of Applied Psychology, 83*(4), 654–665.

Drake, K., & Belsky, J. (2014). From early attachment to engagement with learning in school: The role of self-regulation and persistence. *Developmental Psychology, 50*(5), 1350–1361. doi:10.1037/a0032779.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. doi:10.1037/0022-3514.92.6.1087.

Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*(1), 40–57. doi:10.1037/0021-9010.91.1.40.

Eisenberger, R. (1992). Learned industriousness. *Psychological Review, 99*(2), 248–267.

Eisenberger, R., & Leonard, J. M. (1980). Effects of conceptual task difficulty on generalized persistence. *American Journal of Psychology, 93*(2), 285–298.

Feather, N. T. (1961). The relationship of persistence at a task to expectation of success and achievement related motives. *Journal of Abnormal and Social Psychology, 63*(3), 552–561.

Gee, J. P. (2007). *What video games have to teach use about learning and literacy* (2nd ed.). New York, NY: Palgrave/Macmillan.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg: Tilburg University Press.

Hamachek, D. E. (1978). Psychodynamics of normal and neurotic perfectionism. *Psychology: A Journal of Human Behavior, 15*(1), 27–33.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A. R., Evans, C., & Vitak, J. (2008). *Teens' gaming experiences are diverse and include significant social interaction and civic engagement.* Retrieved from http://www.pewinternet.org/Reports/2008/Teens-Video- Games-and-Civics.aspx

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of Conscientiousness. *Learning and Individual Differences, 19*, 451–458. doi:10.1016/j.lindif.2009.03.007.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81–90.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspective, 1*(1), 3–62.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116–130. doi:10.1037/0022-3514.93.1.116.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*, 971–990. doi:10.1016/j.paid.2007.03.017.

Parker, W. D. (1997). An empirical typology of perfectionism in academically talented children. *American Educational Research Journal, 34*(3), 545–562.

Partnership for 21st Century Learning. (2015). Retrieved from http://www.p21.org

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic.

Perry, S. J., Hunter, E. M., Witt, L. A., & Harris, K. J. (2010). P = f (Conscientiousness x Ability): Examining the facets of conscientiousness. *Human Performance, 23*, 343–360. doi:10.1080/08959285.2010.501045.

vshute@admin.fsu.edu

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338. doi:10.1037/a0014996.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313–345.

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Charlotte, NC: Information Age Publishers.

Shute, V. J., & Ventura, M. (2013). *Measuring and supporting learning in games: Stealth assessment*. Cambridge, MA: The MIT Press.

Shute, V. J., Ventura, M., & Ke, F. (2015). The power of play: The effects of Portal 2 and Lumosity on cognitive and noncognitive skills. *Computers & Education, 80*, 58–67. doi:10.1016/j.compedu.2014.08.013.

Shute, V. J., Ventura, M., & Kim, Y. J. (2013). Assessment and learning of qualitative physics in Newton's Playground. *The Journal of Educational Research, 106*, 423–430. doi:10.1080/00220671.2013.832970.

Shute, V. J., & Zapata-Rivera, D. (2012). Adaptive educational systems. In P. Durlach (Ed.), *Adaptive technologies for training and education* (pp. 7–27). New York, NY: Cambridge University Press.

Stoeber, J., & Otto, K. (2006). Positive conceptions of perfectionism: Approaches, evidence, challenges. *Personality and Social Psychology Review, 10*(4), 295–319.

Terry-Short, L. A., Owens, R. G., Slade, P. D., & Dewey, M. E. (1995). Positive and negative perfectionism. *Personality and Individual Differences, 18*(5), 663–668.

Trapmann, S., Hell, B., Hirn, J. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Journal of Psychology, 215*(2), 132–151. doi:10.1027/0044-3409.215.2.132.

Ventura, M., & Shute, V. (2013). The validity of a game-based assessment of persistence. *Computers in Human Behavior, 29*, 2568–2572. doi:10.1016/j.chb.2013.06.033.

Wang, L., Shute, V. J., & Moore, G. (2015). Lessons learned and best practices of stealth assessments. *International Journal of Gaming and Computer Mediated Simulations, 74*(4), 66–87. doi:10.4018/IJGCMS.2015100104.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., … Conkey, C. (2009). Relationships between game attributes and learning outcomes. *Simulation & Gaming, 40*(2), 217–266. doi: 10.1177/1046878108321866

vshute@admin.fsu.edu