

5-2019

# The Importance of Contextualized, Facet-Level Personality Measures

Joseph Ligato

Clemson University, josephligato@gmail.com

Follow this and additional works at: [https://tigerprints.clemson.edu/all\\_theses](https://tigerprints.clemson.edu/all_theses)

---

## Recommended Citation

Ligato, Joseph, "The Importance of Contextualized, Facet-Level Personality Measures" (2019). *All Theses*. 3053.  
[https://tigerprints.clemson.edu/all\\_theses/3053](https://tigerprints.clemson.edu/all_theses/3053)

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact [kokeefe@clemson.edu](mailto:kokeefe@clemson.edu).

THE IMPORTANCE OF CONTEXTUALIZED, FACET-LEVEL  
PERSONALITY MEASURES

---

A Thesis  
Presented to  
the Graduate School of  
Clemson University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Science  
Applied Psychology

---

by  
Joseph Ligato  
May 2019

---

Accepted by:  
Dr. Fred Switzer III, Committee Chair  
Dr. Mary Anne Taylor  
Dr. Patrick Rosopa

## ABSTRACT

A problem that has arisen in the field of personality psychology is that while personality traits are related to outcome variables, the predictive validity of these associations is low to medium (Rosenthal, 1994; Rosnow & Rosenthal, 1989). One of the reasons for this is because personality has traditionally been defined as something generalizable across situations and time. This generalizability across situations and time is called the invariance of personality (Mischel, 2004). We argue that personality is stable at a different level of analysis, and that level of analysis is the specific context, but not stable across different situations. The current study looked at a fully contextualized personality measure and compared it to a non-contextualized measure of the same personality trait/facets to assess whether incremental validity can be gained by targeting specific situations. Results show that despite the presence of nuisance factors for both general and academic conscientiousness that the contextualized measure showed incremental validity.

*Keywords:* personality, contextual, situational, facet, generalized, person-situation debate, personality measurement, cross-situational consistency

## TABLE OF CONTENTS

	Page
<b>TITLE PAGE</b> .....	<b>i</b>
<b>ABSTRACT</b> .....	<b>ii</b>
<b>LIST OF TABLES</b> .....	<b>v</b>
<b>LIST OF FIGURES</b> .....	<b>vi</b>
<b>INTRODUCTION</b> .....	<b>1</b>
<b>PERSONALITY: THE BEGINNING</b> .....	<b>1</b>
<b>PERSON-SITUATION DEBATE</b> .....	<b>1</b>
<b>ELIMINATE CONTEXT</b> .....	<b>3</b>
<b>INCORPORATE SITUATION</b> .....	<b>3</b>
<b>PERSON-SITUATION INTERACTION</b> .....	<b>4</b>
<b>BIG FIVE</b> .....	<b>5</b>
<b>PERSONALITY AND ACADEMIC SUCCESS</b> .....	<b>6</b>
<b>CONTEXTUALIZED PERSONALITY</b> .....	<b>7</b>
<b>FRAME OF REFERENCE</b> .....	<b>8</b>
<b>FACET LEVEL PERSONALITY</b> .....	<b>9</b>
<b>MODULAR FRAMEWORK</b> .....	<b>12</b>
<b>METHODS</b> .....	<b>16</b>
<b>OVERVIEW</b> .....	<b>16</b>
<b>PARTICIPANT’S DEMOGRAPHICS</b> .....	<b>16</b>
<b>VARIABLES</b> .....	<b>17</b>
<b>PROCEDURE</b> .....	<b>19</b>
<b>SOFTWARE</b> .....	<b>19</b>
<b>ANALYSES</b> .....	<b>20</b>
<b>RESULTS</b> .....	<b>24</b>
<b>OUTLIER ANALYSIS</b> .....	<b>24</b>
<b>SIX HYPOTHEZED FACETS: RELIABILITIES</b> .....	<b>25</b>
<b>SIX HYPOTHEZED FACETS: DESCRIPTIVE STATISTICS AND CORRELATIONS</b> .....	<b>30</b>
<b>FOUR FACTOR ANALYSIS DERIVED VARIABLES: RELIABILITIES</b> .....	<b>36</b>
<b>FOUR FACTOR ANALYSIS DERIVED VARIABLES: DESCRIPTIVES AND CORRELATIONS</b> .....	<b>37</b>
<b>GENERAL CONSCIENTIOUSNESS CONFIRMATORY FACTOR ANALYSIS: MODELS AND INTERPRETATIONS</b> .....	<b>39</b>
<b>ACADEMIC CONSCIENTIOUSNESS CONFIRMATORY FACTOR ANALYSIS: MODELS AND INTERPRETATIONS</b> .....	<b>44</b>
<b>ACADEMIC CONSCIENTIOUSNESS: SUMMARY FINDINGS FOR ITEM-LEVEL SCALE CHANGES</b> .....	<b>52</b>
<b>CONSCIENTIOUSNESS SCALES: SUMMARY FINDINGS FOR EFA AND CFA</b> .....	<b>53</b>
<b>ACADEMIC CONSCIENTIOUSNESS: ITEM-LEVEL INCREMENTAL VALIDITY AND FACTOR STRUCTURE INTERPRETATION</b> .....	<b>54</b>
<b>RESULTS FOR ORIGINAL HYPOTHESIS:</b> .....	<b>56</b>
<b>DISCUSSION</b> .....	<b>61</b>

Table of Contents (Continued)	Page
<b>IMPLICATIONS FOR THEORY, RESEARCH, AND PRACTICE</b> .....	61
<b>FURTHER INTERPRETATION OF NEGATIVE FACTORS</b> .....	65
<b>LIMITATIONS</b> .....	68
<b>FUTURE RESEARCH</b> .....	69
<b>APPENDICES</b> .....	<b>74</b>
<b>EXAMPLE QUESTIONNAIRE ITEMS</b> .....	74
<b>GENERALIZED CONSCIENTIOUSNESS ITEMS</b> .....	81
<b>ACADEMIC CONSCIENTIOUSNESS ITEMS</b> .....	82
<b>LIKERT SCALE POINTS</b> .....	84
<b>TABLES</b> .....	84
<b>FIGURES</b> .....	106
<b>REFERENCES</b> .....	<b>108</b>

## LIST OF TABLES

Table	Page
1. RELIABILITIES FOR ORIGINAL SIX HYPOTHESIZED FACETS: .....	26
2. DESCRIPTIVE STATISTICS AND CORRELATIONS OF ORIGINAL HYPOTHESIZED VARIABLES .....	30
3. FACTOR ANALYSIS GENERAL CONSCIENTIOUSNESS: TWO FACTOR .....	32
4. FACTOR ANALYSIS ACADEMIC CONSCIENTIOUSNESS: TWO FACTOR .....	34
5. RELIABILITIES FOR FACTOR ANALYSIS DERIVED VARIABLES: .....	37
6. DESCRIPTIVE STATISTICS AND CORRELATIONS OF GPA AND FACTOR ANALYSIS DERIVED VARIABLES...	38
7. CHI-SQUARE OF ALL ESTIMATED GENERAL CONSCIENTIOUSNESS MODELS .....	40
8. MODEL FIT OF ALL ESTIMATED GENERAL CONSCIENTIOUSNESS MODELS .....	41
9. LMTTEST FOR ADDING PARAMETERS TO GENERAL CONSCIENTIOUSNESS SCALE .....	43
10. CHI-SQUARE OF ALL ESTIMATED ACADEMIC CONSCIENTIOUSNESS MODELS (GCM) .....	45
11. MODEL FIT OF ALL ESTIMATED ACADEMIC CONSCIENTIOUSNESS MODELS (ACM).....	46
12. LMTTEST FOR ADDING PARAMETERS TO ACADEMIC CONSCIENTIOUSNESS SCALE .....	48
13. MULTIPLE REGRESSION FOR ACADEMIC CONSCIENTIOUSNESS ITEMS: PART 1 .....	49
14. MULTIPLE REGRESSION FOR ACADEMIC CONSCIENTIOUSNESS ITEMS: PART 2 .....	50
15. MULTIPLE REGRESSION FOR ACADEMIC CONSCIENTIOUSNESS ITEMS: PART 3 .....	51
16. MULTIPLE REGRESSION FOR ACADEMIC CONSCIENTIOUSNESS ITEMS: PART 4 .....	52
17. STEPWISE REGRESSION FOR ACADEMIC CONSCIENTIOUSNESS ITEMS .....	55
21. PATH CONSTRAINTS FOR ACADEMIC AND GENERAL CONSCIENTIOUSNESS RELATIONSHIP WITH COLLEGE GPA .....	56
18. MULTIPLE REGRESSION OF ACADEMIC CONSCIENTIOUSNESS .....	57
19. MEDIATION ANALYSIS OF ACADEMIC CONSCIENTIOUSNESS: PART 1.....	57
20. MEDIATION ANALYSIS OF ACADEMIC CONSCIENTIOUSNESS: PART 2.....	59

## LIST OF FIGURES

Figure	Page
1. SCREE PLOT FOR GENERALIZED CONSCIENTIOUSNESS EXPLORATORY FACTOR ANALYSIS .....	33
2. SCREE PLOT FOR ACADEMIC CONSCIENTIOUSNESS EXPLORATORY FACTOR ANALYSIS .....	36
3. ACADEMIC CONSCIENTIOUSNESS, 1 FACTOR SOLUTION WITH NO NEGATIVELY WORDED ITEMS, V26 AND V45 REMOVED, COVARIANCE (E44, E27) AND (E28, E27) .....	59
4. GENERAL CONSCIENTIOUSNESS, 1 FACTOR SOLUTION WITH NO NEGATIVELY WORDED ITEMS, 2 COVARIANCES (E11, 10) AND (E15, 14).....	60

## INTRODUCTION

### **Personality: The Beginning**

In 1928, Hartshorne & May conducted a study on conscientiousness in schoolchildren that led to the assumption that personality would have cross-situational consistency (Mischel, 2004, p. 2). Theodore Newcomb conducted a very important study in 1929 that unsuccessfully worked towards refuting the notion of cross-situational consistency. In this landmark study, 51 boys were assessed in 21 situations at a summer camp, and it was found that depending on the situation the results of introversion-extroversion scores differed significantly. The average correlation coefficient across situations was only about 0.14 (Mischel, 2004, pp. 2–3). This is extremely low, to the point where it would be considered a poor effect size if it was a correlation between personality and an outcome variable, let alone as a correlation between a personality trait in different contexts.

### **Person-Situation Debate**

The idea that personality might be situationally specific is not a new argument and has, in fact, been made in numerous contexts, but it has gone under different names. B. F. Skinner for example, a prominent behaviorist believed that “behavior is situationally specific, driven by external cues and stimuli” (Mastroianni, 2011, p. 2). A second example can be found in Stanley Milgram’s shock experiment, which showed that participants are very likely to obey authority even in extreme circumstances (1963). A third prominent example can be found in Zimbardo’s work with people pretending to be either prisoners or prison guards, and where people randomly assigned to be prison



guards started to commit psychological torture on the participants pretending to be prisoners. Zimbardo calls situations that push people to act in bad ways “good apples in bad barrels” (Mastroianni, 2011, p. 7).

Personality psychologists want to find a high degree of cross-situational stability in order to legitimize and generalize the usage of personality measures (Furr & Funder, 2004, p. 433). Social psychologists seek to strengthen recognition of situations, since they believe situations are highly important. Personality psychologists would say that social psychologists want to strengthen the importance of situations at the expense of personality traits (Nisbett & Ross, 1985; Ross & Nisbett, 1991). This gets into a subject known as the Person-Situation debate that has raged on between personality and social psychologists for years.

This scenario mirrors the Nature-Nurture debate but without the resolution. The Nature-Nurture debate eventually decided that each side contributed significantly to a person’s development and that there were interaction terms that also occurred between the two (Fleeson & Nofhle, 2009, p. 152; Funder, 2006, p. 32). An example of an interaction in this debate is epigenetics, which is how genes (nature) express themselves based on circumstances (nurture). However, given the predominant view in personality psychology, the dichotomy between the two camps formed without coming up with interaction terms or any compromise (Fleeson, 2004, p. 86). Personality as a field seems to consider any personality psychologist who studies context to be trying to ruin their field (Mischel, 2004, pp. 3–4).

## **Eliminate Context**

One-way to try to resolve this dilemma is to simply aggregate across situations. This increases ease of measure creation, but at the expense of much of personality's predictive validity. This refers to aggregation, which is used to find an overall "true score" (Mischel, 2004, p. 3; Mischel & Shoda, 1995, pp. 1–2). This aggregation supports the predominant view espoused in personality psychology, which is that the "basic qualities of the person are assumed to be independent of, and unconnected with, situations" (Mischel, 2004, p. 3). The argument on eliminating context is never directly made; it is only implied based on the notion that personality is generalizable across cultures, and given that, it should not be situationally specific (Heine & Buchtel, 2009, pp. 377–379).

## **Incorporate Situation**

Another camp in this debate is that context should be considered if researchers studying personality ever want to increase the predictive validity of personality measures and reach a resolution over the lack of coherence in the study of personality measurement. The belief underlying this camp is that people view situations differently and will have different feelings and thoughts about the situation, even if they have the same true score of a given personality trait. If situations lead to different affects and cognitions, then we can expect that the individual will behave in a way that takes that differing context into account (Mischel, 2004, p. 5).

Various studies have also shown the benefit of considering roles/situations for their ability to account for what was previously perceived as error variance (Baird &

Lucas, 2011, p. 1103). Some researchers contend that the magnitude of the correlations between personality and specific behaviors is misunderstood, since personality variables are still capable of predicting behaviors correctly about 66% of the time when compared to random chance between two choices (50%). The original reason for considering situations might be partially based on a misunderstanding about the strength of personalities' validity coefficients, but the authors of this critique still believe situation is extremely important to consider (Guillaume, Kumagai, Kawamoto, Sato, & others, 2012, p. 5).

In one study, it was found that general and work-specific measures of personality were highly consistent, but there was support for the notion that the work-specific measure had incremental validity (Bowling & Burns, 2010, p. 6). In other literatures, such as self-efficacy, we can also see that the field is moving towards more context-specific measures when looking at studies on task-specific self-efficacy as compared to general self-efficacy (Dickerson & Taylor, 2000, pp. 195–196). Wang, Bowling, and Eschleman (2006) conducted a meta-analysis on work and general locus of control showing that work locus of control is best to use in a work-related context.

### **Person-Situation Interaction**

The necessity of taking interaction into account started primarily in the 1970s when personality as a beneficial measure for outcome variables was under fire for low predictive validity and perceived unfairness (Griffo & Randall Colvin, 2009, pp. 243–244). An example where the interaction between situation and personality is important comes from Fleeson & Leicht (2006, pp. 17–18) where they talk about the interaction

between relationship type and state trust. Results can also be found in evolutionary literature, which shows that person\*situation interactions have an evolutionary benefit (Buss, 2009, p. 241).

The literature is unclear when talking about person\*situation interactions. Sometimes when talking about interactions they mean contextualized measures; other times they actually mean how general personality measures interact with some type of context measure. While we will primarily focus on the former, the latter has some examples that need mentioning. Specific models for person\*situation interactions have been created, such as the Latent State-Trait models, which randomly sample situations (Geiser et al., 2015, pp. 166–168). Person\*situation interactions have been tested under a variety of conditions to see whether they explain performance variables better than personality. Results from the interpersonal trust literature show that within-person differences can exist as both contextual measures and actual interactions between general measures and context measures i.e. situational cues, which complicates discussions of contextual effects (Geukes, Mesagno, Hanrahan, & Kellmann, 2012, pp. 247–249).

### **Big Five**

The Five Factor Model (FFM) is considered to be one of the most supported models of overall personality that exists in modern times (Petska, 2006, p. ii). The FFM's factors are openness, conscientiousness, extraversion, agreeableness, and neuroticism (O.C.E.A.N.). The construct we will be focusing on in this study is conscientiousness. Conscientiousness can be described as: “socially prescribed impulse control that facilitates task and goal-directed behaviors, such as thinking before acting, delaying

gratification, following norms and rules, planning, organizing, and prioritizing tasks” (Petska, 2006, p. 15)

The literature shows that while the Big Five is a measure of personality that has cross-cultural validity, it does not necessarily mean that it is the right level of analysis in order to predict outcome variables. Since the Big Five are so generalized, various researchers have come up with ideas about facet-level variables to look at, which might have more theoretical justification for studying with regards to some outcome variables. One researcher named John Johnson, a prominent psychologist from Pennsylvania State University, believes that the Big Five should have 30 facets, so he created scales to differentiate between facets. For Conscientiousness, he broke it down into six facets and these are: Self-Efficacy, Orderliness, Dutifulness, Achievement-Striving, Self-Discipline, and Cautiousness (2014a, pp. 82–83). These are the scales we will examine in our study.

### **Personality and Academic Success**

The relationship between conscientiousness and academic success has been studied in multiple contexts using many different measures of academic success including high school GPA, college GPA, ACT/SAT scores, etc. Petska (2006) study of ACT and Conscientiousness together accounted for 14% of the variance in an individual’s college GPA.

Poropat (2009) did a meta-analysis of studies on the relationship between academic success and personality using the FFM. Most of the research studied was from higher education with a lesser amount from secondary and primary school. He found that

Conscientiousness was correlated with academic performance overall. The extent that conscientiousness was correlated with academic success was equal to intelligence for college performance, when controlling for secondary education.

### **Contextualized Personality**

Contextualized personality has rarely been studied and a lot of studies have not used completely contextualized personality but have instead just tagged on specific situations at the end. This reduces the cognitive demands on the participants filling out the questionnaire, but also may not activate situation-specific memories as much as more fully contextualized questions (Robie & Risavy, 2016). It is the difference between “I complete tasks successfully at school” and “I complete my homework successfully”; the prior is not fully contextualized, while the latter is fully contextualized.

An example of a study that did use fully contextualized personality looked at “college conscientiousness”. Two facts emerged from this study:

On the one hand, just as Newcomb did, we also found that behaviors were highly variable across different situations. An individual might be higher than most people in a trait in some situations but also distinctively lower than most in other situations. On the other hand, individuals also showed temporal stability in their behavior within particular situations that were highly similar and formed a type, or “functional equivalence class,” of situations. It was noteworthy that their perceptions of their own trait consistency were strongly related to that temporal stability, and unrelated to the variability of their behavior from one type of situation to another (Mischel, 2004, pp. 4–5).

What this means is that there does seem to be coherence in personality, but it seems to be at a lower level than the aggregate of the trait. It seems that personality shows coherence in similar situations that likely elicit the same feelings and thoughts within the individual.

Postlethwaite and Shaffer (2012) in their meta-analysis found that context is important to consider. Their main hypothesis was that considering context would have incremental validity over generalized scales. This hypothesis was supported, which provides initial evidence to the importance of further creating and studying contextualized scales. A study by Dunlop (2015) found that considering life goals and narratives may help to further the understanding of contextualized personality, such that inter-contextual variability might be a product of differing life goals and narratives. This aligns with Walter Mischel and Yuichi Shoda's work on Cognitive-Affective Personality Systems theory (CAPS), which discusses how the factors that constitute a situation are dependent on the perceiver's constructs and subjective experiences (1995, p. 252).

### **Frame of Reference**

A way to look at contextualized personality is to use specific frames of reference (Schmit, Ryan, Stierwalt, & Powell, 1995, pp. 608–610). A way of doing this that is low on cognitive demand for participants is to just tag the situation at the end of the manifest variable phrase (Robie & Risavy, 2016). A more nuanced way that is more cognitively demanding for participants is to fully contextualize the manifest variable phrase by adding in specifics of the situation in question. The latter is expected to offer incremental validity over non-contextualized personality traits and that is what is being studied in this research proposal.

## **Facet Level Personality**

A meta-analysis conducted on facets (narrow traits) of conscientiousness provides evidence for using facets for better theory-building (Dudley, Orvis, Lebiecki, & Cortina, 2006). The principle of compatibility would dictate that specific predictors should be used to predict specific outcome variables, and general predictors should be used for general outcome variables. This notion is contentious since, from a theory-building standpoint, specific predictors allow for more nuanced interpretations no matter the specificity of the outcome variable; whereas a general predictor loses clarity.

Contextualized personality measures are another example of more specific predictors. Therefore, combining contextualized and facet-level measures provides an even more fine-grained prediction than either one separately. Facet-level variables are one way to fix deficiency in our personality predictors. Contextualized personality could be interpreted as a way to fix contamination of our personality predictors. Therefore, facet-level contextualized personality would be one way to fix both contamination and deficiency in our predictors.

### **Achievement-Striving Facet.**

Achievement-striving taps into the desire of the individual to advance in life. People high in this facet want to be successful, so they persist at their goals and give their best effort to tasks relevant to their goals. An example of a generalized achievement-striving question is “I do more than what’s expected of me” (Johnson, 2014b, p. 83). An example of an academic achievement-striving question is “I aim to get a perfect grade on assignments.” It is self-explanatory why having a desire for the best possible grades is



important for having a high college GPA. A student is essentially deciding the upper limit to how well he/she will do in a class. This is one of the facets considered likely to have a strong, theoretically meaningful relationship with our outcome variable of college GPA.

### **Self-Discipline Facet.**

Self-discipline is a facet that deals with staying on task, avoiding distractions, practicing self-regulation, and delaying gratification. Given this definition, studies have indirectly shown that self-discipline likely has a relationship with college GPA, at least partially through the delayed gratification component of self-discipline (Herndon, 2011). An example of general self-discipline is “I carry out my plans” (Johnson, 2014b, p. 83). An example of an academic self-discipline question is “I follow a schedule to work on assignments.” This facet is likely important because it gives the student the ability to adjust the rest of his/her schedule to set the amount of time aside that is needed to learn the material. This facet is also important for understanding the relationship between conscientiousness and college GPA.

### **Self-Efficacy Facet.**

Self-efficacy, also sometimes called competence, is a person’s belief that they can complete a task. An example of a general self-efficacy question is “I complete tasks successfully” (Johnson, 2014b, p. 82). An example of an academic self-efficacy question is “I complete my school assignments successfully.” The researchers believe that a student’s feeling of confidence in completing a task likely leads him/her to starting a task and persisting through to its completion, which is important for being successful in college and likely also has a relationship with having a high college GPA.

### **Cautiousness Facet.**

Cautiousness is a facet that deals with how quickly one makes decisions. An example of a general cautiousness question is “I rush into things” (Johnson, 2014b, p. 83). A negatively worded version of an academic cautiousness question is “I act without thinking at school.” It is thought that acting quickly is unnecessary for achieving high levels of success in college and therefore will not be related to college GPA.

### **Orderliness Facet.**

Orderliness deals with keeping life neat and organized. An example of a general orderliness question is “I like to tidy up” (Johnson, 2014b, p. 82). An example of an academic orderliness question is “I take organized notes in class.” This facet is not thought to have much of a relationship with college GPA. Academic versions of being orderly were very difficult to come up with, which means that if there is a relationship between this facet and the outcome variable, it is likely because of a common source.

### **Dutifulness Facet.**

Dutifulness as a facet has some problems. Usually when one thinks of dutifulness, they think of being responsible, but the general questions normally associated with it seem to tap into more of an honesty facet. An example of a dutifulness question is “I keep my promises” (Johnson, 2014b, p. 83). A negatively worded version of an academic dutifulness question is “I cheat on tests.” This facet likely has very little to do with academic success, mostly because it is thought that cheating has minimal benefits, if any, with regards to college GPA, usually depending more on whether the person is caught or not.

## Modular Framework

The modular framework is “often adopted when products have become established on the market and have evolved further in their life cycle” (Christensen, 2001; Schilling, 2000). The product of personality measures have been established to a large extent after the widespread acceptance of the Five Factor Model; now it is time to make specific comparisons about the component parts of personality measures to see where, when, why, how, and potentially with whom these personality measures work best.

Personality measurement is at a critical point where it needs to be broken down further. According to Lievens & Sackett (2017, p. 44):

First, a modular approach allows breaking down a large and complex system into smaller more manageable parts. Whereas the functioning of the system as a whole remains typically a black box, a modular approach enables gaining better insight into the workings of the different components. Superordinate constructs, such as the Big Five, can be viewed as “black boxes” where the functioning of the construct is a mystery because we do not know which parts are actually causing the prediction. Facets allow us to gain better insight into the inner workings of black boxes. The modular framework then is the perspective that acknowledges the importance of specific predictors.

The modular framework also has another benefit of allowing us to see whether multiple facets are capable of predicting the same variance. Lievens & Sackett state (2017, p. 44):

As a second conceptual advantage, a modular approach to selection procedures promotes identifying and exploiting communalities among selection procedures. That is,

it may show that the same components underlie superficially different selection procedures and that they produce similar effects across them.

This is important because to a large extent, studying variables that do not have any reasonable association with an outcome variable can lead to confusion about how to proceed in training for it or feelings of unfairness among applicants if selecting for it. An example being if orderliness is related to College GPA, even though there are very few instances where college students can be orderly in a college environment that would not already fall under another more relevant facet of conscientiousness, such as self-discipline. In this instance, using self-discipline would make more sense because of its relevance and face validity.

The literature in preceding decades seems to have focused on supporting the usage of these constructs as a whole but as the field of personality has felt more justified in using predictors, they have decided to look at personality at a more specific level. This gets into the idea of “looking under the hood” at what is really going on in these relationships between variables (Lievens & Sackett, 2017). By looking at facet-level versions of variables or looking at contextual versions, or even better looking at both, researchers get a sense for the specifics of the relationship between variables.

This allows the field to make progress in terms of not only predictive validity but also the relevant personality theories. In the past, multifaceted variables like conscientiousness were just thrown at the wall and researchers accepted that some parts of the variable “stuck”, but were not able to have a sense for which parts of conscientiousness were actually responsible for the relationship. Moving forward, the

field needs to look at how to properly contextualize variables, as well as get a sense for how to properly breakdown superordinate constructs. Various researchers have their own ideas about how to break down something like conscientiousness. The main trend is looking at things at a more specific level to try to get closer to understanding the causal mechanisms for the relationship.

The primary gap in the literature is that we have separate literatures on context and facets of superordinate constructs, but as Shaffer and Postlethwaite (2012) mention at the end of their paper when discussing future directions, we need to combine the two approaches: to look at contextualized variables at a fine-grained facet level.

**Hypothesis 1:** Academic Conscientiousness as a whole will have *incremental validity over and above* Generalized Conscientiousness in relation to College GPA. Both will be statistically significant.

**Hypothesis 2a:** Academic Achievement-Striving as a facet of Academic Conscientiousness will have *higher predictive validity* than Achievement-Striving as a facet of Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 2b:** Academic Achievement-Striving as a facet of Academic Conscientiousness will have *higher predictive validity* than Academic Conscientiousness as a whole.

**Hypothesis 2c:** Achievement-Striving as a facet of Generalized Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness as a whole.

**Hypothesis 2d:** Academic Achievement-Striving as a facet of Academic Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 3a:** Academic Self-Discipline as a facet of Academic Conscientiousness will have *higher predictive validity* than Self-Discipline as a facet of Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 3b:** Academic Self-Discipline as a facet of Academic Conscientiousness will have *higher predictive validity* than Academic Conscientiousness as a whole.

**Hypothesis 3c:** Self-Discipline as a facet of Generalized Conscientiousness will have *higher predictive validity* than Conscientiousness as a whole.

**Hypothesis 3d:** Academic Self-Discipline as a facet of Academic Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 4a:** Academic Self-Efficacy as a facet of Academic Conscientiousness will have *higher predictive validity* than Self-Efficacy as a facet of Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 4b:** Academic Self-Efficacy as a facet of Academic Conscientiousness will have *higher predictive validity* than Academic Conscientiousness as a whole.

**Hypothesis 4c:** Self-Efficacy as a facet of Generalized Conscientiousness will have *higher predictive validity* than Conscientiousness as a whole.

**Hypothesis 4d:** Academic Self-Efficacy as a facet of Academic Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness. Both will be statistically significant.

## METHODS

### Overview

We used a variety of software depending on their capabilities: SPSS, Excel, R, and EQS. Our primary tool for data clean up (fixing missing data, performing outlier analysis, creating composites, reverse scoring, etc.) was SPSS. Some graphing was done in Excel. R and EQS were used for structural equation modeling.

### Participant's Demographics

Our sample size for individuals who filled out the survey and were not removed during outlier analysis was 358. Due to the fact that not all participants filled out undergraduate college GPA, the sample size of that portion of analysis was 280. Participants range in ages from 18 to 71 ( $SD = 9.75$ ). Our participants were 54.7 male and 45.3 female. We had no one write in a text entry for the other category of gender. This is fairly representative of college graduate gender breakdowns, especially considering that those not writing in an undergraduate GPA score were more likely to be male, which means the analysis relevant to our academic outcome variable has more females than the overall sample (“America’s College Students Infographic | Postsecondary Success,” 2014).

We broke race/ethnicity into a multiple-choice “pick all that apply” question as well as a yes/no question on whether they self-identify as Hispanic/Latino or not. This means that the number of people identifying as each, if added up, exceeded the total number of participants because people could select multiple categories. Results show that 10.6% of participants identified as Hispanic or Latino. This is under the national average of college graduates but because our sample is representative of a wide age range, it is likely that our sample is representative of a breakdown of Hispanic/Latino students graduating over the time period of our participants graduating (“America’s College Students Infographic | Postsecondary Success,” 2014).

For race/ethnicity, we found 313 (87.4%) participants identify as White. We had 26 (7.3%) participants identify as Black or African American. We had 12 (3.4%) participants identify as American Indian or Alaska Native. We had 19 (5.3%) participants identify as Asian. We had 1 (.3%) participant identify as Native Hawaiian or Pacific Islander. As you can see, 3.7% of participants identified as multiracial. Our results are supportive of the growing diversity in higher education, though still more work needs to be done in future studies to make sure we research a more diverse population in order to ensure the generalizability of our results (“America’s College Students Infographic | Postsecondary Success,” 2014).

## **Variables**

### **Non-contextualized conscientiousness NEO PI-R scale.**

This is one of five components of the NEO PI-R that make up the five major domains of personality. This scale has six facets to it. Conscientiousness has been shown



to be a personality trait to have one of the strongest positive relationships with measures of academic success such as College GPA, HS GPA, and SAT/ACT scores (Petska, 2006). Cronbach's alpha for this scale is based on a sample size of 619,150. All are within .67 to .88 which are considered good alphas for scales. A five-point response scale was used to indicate respondents' degree of agreement with each statement. Not Accurate At All (1), Slightly Accurate (2), Moderately Accurate (3), Very Accurate (4), and Extremely Accurate (5) (see Appendix B).

#### **Fully contextualized conscientiousness NEO PI-R scale.**

This is a version of the NEO PI-R that was created specifically for academic environments. It does not have a sample to find Cronbach's alpha yet. It was created as a potential example of how to improve the predictive validity of personality measures. It was predicted before the start of this study that this measure would do better than the non-contextualized conscientiousness scale in terms of the strength of its relationship with College GPA. A five-point response scale was used to indicate respondents' degree of agreement with each statement. Not Accurate At All (1), Slightly Accurate (2), Moderately Accurate (3), Very Accurate (4), and Extremely Accurate (5) (see Appendix B).

#### **College GPA.**

College GPA was used to assess whether fully contextualized personality traits, such as conscientiousness, have incremental validity over non-contextualized conscientiousness. College GPA can range from 0 to 4.0, with scores below 2.0 being

rare because of academic dismissal. Scores usually have decimal places going out 2 or 3 places, such as a College GPA of 3.875.

### **Procedure**

The researchers used Qualtrics to create the survey and MTurk to get participants. Students were asked if they consent to participating and were instructed on the ways to correctly fill out the questionnaire as well as their freedom to not participate or to quit at any time. Confidentiality was assured because we did not ask for any distinct identifying characteristics. Once the required amount had filled out the questionnaires, data was analyzed. Data analysis did not occur before all data had been collected, unless preliminary data needed to be collected to assess the internal consistency of our fully contextualized conscientiousness measure. Data analysis before all data is collected is a questionable research practice that leads to fishing for significant results through either collecting more data than agreed upon before the study or stopping collection early if statistically significant results have already been found (Krishna & Peter, 2018).

### **Software**

A variety of software was used for the analyses in order to meet all the demands of conducting research ethically and thoroughly as well as conveying the research in the most understandable way possible. Because the researchers had not learned R yet, a variety of software was needed to meet all these demands. Excel was used for some calculations as well as graphs. SPSS, which is a statistical analyses software package was used for univariate and multivariate outlier analysis as well as testing reliability of our

measures. EQS a Structural Equation Modeling (SEM) software was used for the analyses of the factor structure of both measures.

## **Analyses**

### **Univariate Outlier Analysis.**

Data screening occurred for our sample in SPSS before any further analyses were conducted. Descriptive statistics such as minimum, maximum, means, and standard deviations were looked at to determine if there were any overt instances of error outliers. Things like scatterplot, box plot, and histogram were used for univariate outlier analysis. Standard deviation at the item level was found to be extremely useful for detecting fake responses that were not detected with other more thorough outlier analysis like Mahalanobis distance. There is a current problem with automated programs (bots) and outsourcing of surveys that are leading to low quality data from MTurk and Qualtrics (Dreyfuss, 2018; Harris, 2018; TurkPrime, 2018).

We used standard deviations and difference scores between positively and negatively worded items to decide whether to look at individual surveys for questionable data. We ended up finding quite a few surveys that filled out all the multiple-choice questions using only 1 or 2 survey responses. We also looked at things like duration of time to complete survey to see if people completed the survey faster than reasonably possible. We also looked at whether people who reported having a college GPA even reported having a college education to decide whether they were giving us faulty data or not.

Direct attention checks were not used because of some research showing attention checks can affect results for data collected after the attention check occurs. The interpretation is that people get lazy after the attention check because they think they have made it past anything the researcher can use to check for attention (Vannette, 2016).

### **Multivariate Outlier Analysis.**

Multivariate outliers were looked at next in SPSS. We looked at things such as Mahalanobis distance, Studentized Deleted Residuals, and Cook's d, to determine if there were any extreme cases that might need to be removed. Any outliers, whether univariate or multivariate, were decided on a case-by-case basis whether to exclude. We were planning to list the reasons for exclusions for each case, based on prior samples that have shown minimal cases needing to be excluded, unfortunately because of the recent problem with MTurk, we had to remove dozens of cases of clearly low-quality data and went over a few representative cases.

Some cases that were considered outliers were kept in the data for reasoning regarding the fact that questions could have multiple interpretations. Despite the fact that we could have made our scales look more reliable by getting rid of some of these outlying participants using standard cutoff scores

### **Reliability Analysis.**

Internal consistency of the scales were looked at, particularly for the fully contextualized conscientiousness scale because of its recent creation and lack of validation as a legitimate measure. Cronbach's Alpha has a lot of problems with it though and considering the large amount of questions we reviewed in this phase of scale

creation, it was expected that Cronbach's Alpha would have been an unhelpful measure of reliability. Test-retest reliability as well as potentially more robust measures of reliability such as Coefficient Omega were used once the scale was refined.

### **Hypothesis 1 Analysis.**

**Hypothesis 1:** Academic Conscientiousness as a whole *will have incremental validity over and above* Generalized Conscientiousness in relation to College GPA. Both will be statistically significant.

We will be creating a composite of both Academic and Generalized Conscientiousness. The paths to College GPA will be restrained to be equal to see if there is significant chi square differences, meaning that they are statistically significantly different from each other.

### **Hypotheses 2a/b/c/d Analysis.**

**Hypothesis 2a:** Academic Achievement-Striving as a facet of Academic Conscientiousness will have *higher predictive validity* than Achievement-Striving as a facet of Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 2b:** Academic Achievement-Striving as a facet of Academic Conscientiousness will have *higher predictive validity* than Academic Conscientiousness as a whole.

**Hypothesis 2c:** Achievement-Striving as a facet of Generalized Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness as a whole.

**Hypothesis 2d:** Academic Achievement-Striving as a facet of Academic Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness. Both will be statistically significant.

For this set of hypotheses, we will be comparing Academic and Generalized Achievement-Striving facet with the superordinate construct of Academic and Generalized Conscientiousness. We will do this by constraining the paths to be equal and see if there is a statistically significant difference.

### Hypotheses 3ab/c/d Analysis.

**Hypothesis 3a:** Academic Self-Discipline as a facet of Academic Conscientiousness will have *higher predictive validity* than Self-Discipline as a facet of Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 3b:** Academic Self-Discipline as a facet of Academic Conscientiousness will have *higher predictive validity* than Academic Conscientiousness as a whole.

**Hypothesis 3c:** Self-Discipline as a facet of Generalized Conscientiousness will have *higher predictive validity* than Conscientiousness as a whole.

**Hypothesis 3d:** Academic Self-Discipline as a facet of Academic Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness. Both will be statistically significant.

For this set of hypotheses, we will be comparing Academic and Generalized Self-Discipline facet with the superordinate construct of Academic and Generalized Conscientiousness. We will do this by constraining the paths to be equal and see if there is a statistically significant difference.

### Hypotheses 4a/b/c/d Analysis.

**Hypothesis 4a:** Academic Self-Efficacy as a facet of Academic Conscientiousness will have *higher predictive validity* than Self-Efficacy as a facet of Generalized Conscientiousness. Both will be statistically significant.

**Hypothesis 4b:** Academic Self-Efficacy as a facet of Academic Conscientiousness will have *higher predictive validity* than Academic Conscientiousness as a whole.

**Hypothesis 4c:** Self-Efficacy as a facet of Generalized Conscientiousness will have *higher predictive validity* than Conscientiousness as a whole.

**Hypothesis 4d:** Academic Self-Efficacy as a facet of Academic Conscientiousness will have *higher predictive validity* than Generalized Conscientiousness. Both will be statistically significant.

For this set of hypotheses, we will be comparing Academic and Generalized Self-Efficacy facet with the superordinate construct of Academic and Generalized

Conscientiousness. We will do this by constraining the paths to be equal and see if there is a statistically significant difference.

## RESULTS

### Outlier Analysis

#### Univariate Outlier Results.

Univariate outlier analysis showed some errors occurred in exporting data as well as some participant errors. One participant managed to circumvent the 0-4 numerical restriction on undergraduate GPA scores by accidentally typing in “1,7” which led Qualtrics to export it as an undergraduate GPA score of 17. Numerous individuals were filtered out using indirect attention checks.

In my results, I ask individuals what their highest level of education completed is, with options for “some undergrad currently attending” or “some undergrad did not finish and not currently attending.” This means that if they put “less than high school” or “high school graduate only”, I can filter those results out as indirect attention checks. There is no way for them to have college GPA scores if they say they have not attended college.

#### Multivariate Outlier Results.

One respondent was an interesting case, despite being considered a multivariate outlier in a couple of Mahalanobis’ distance ways, I decided to keep this person’s survey in the analysis. They answered 5, Extremely Accurate, for three items on the Self-Efficacy facet and a 1, Not Accurate At All, on Self-Efficacy facet of the Conscientiousness scale item #3 (CSE3). CSE3 I believed could have multiple meanings and given the pattern of answers this person gave in other ways, I ultimately decided it

would not be fair to remove them, or even remove CSE3 or the CSE facet because the data was enough in line with everything else that it did not feel justifiable, even with the outlier analysis saying I would be well within my rights to take it out.

Another way data was screened was through looking at how people answered differently for negatively worded versus positively worded items. This was a simple way, after reverse scoring negatively worded items, to see if people put 1s or 5s for everything, because averaging would not show all 1s or 5s after reverse scoring the negatively keyed items. This was considered an indirect attention check since people likely were not paying attention if they filled out the negative end of the scale with the same or very similar answers to the positive end.

Mahalanobis distance has trouble finding multivariate outliers on the extreme positive end if all personality traits are likely to have range restriction, where people nearly always answer in the top half. For example, one participant answered 4s and 5s for all multiple-choice items, despite the fact that half were negatively keyed items. This is an impossibility, even for people who struggle to mentally reverse the item in their head to the positive end. Neither studentized deleted residuals nor Mahalanobis distance were able to catch this case.

### **Six Hypothesized Facets: Reliabilities**

We found that the number of reverse scored items in a facet scale decided to a large extent the internal consistency of the subscales. Further adding support that negatively worded items add another layer of complexity to scales that can confuse participants and lead to slightly different scores than positively keyed items. Scale



reliabilities and their improvements after outlier analysis for the original hypothesized six facets are below.

*1. Reliabilities for Original Six Hypothesized Facets:  
Before and After Outlier Analysis*

<b>Title</b>	<b>Description</b>	<b>Before Alpha</b>	<b>After Alpha</b>	<b># of Items</b>	<b># of Positive Items</b>	<b># of Negative Items</b>
<b>CAS</b>	Achievement-Striving facet of General Conscientiousness	.666	.702	4	2	2
<b>CSE</b>	Self-Efficacy facet of General Conscientiousness	.826	.830	4	4	0
<b>CO</b>	Orderliness facet of General Conscientiousness	.826	.844	4	1	3
<b>CD</b>	Dutifulness facet of General Conscientiousness	.699	.720	4	2	2
<b>CSD</b>	Self-Discipline facet of General Conscientiousness	.703	.725	4	2	2
<b>CC</b>	Cautiousness facet of General Conscientiousness	.926	.926	4	0	4
<b>CON</b>	General Conscientiousness Superordinate Construct	.928	.932	24	11	13

Note. N = 358.

**CAS, Achievement-Striving facet of Conscientiousness:**

In Table 1, we see the reliability of our subscale Achievement-Striving subscale before and after outlier analysis. We see that after reverse scoring the negatively keyed

items that the internal consistency, as measured by Cronbach's alpha, of the Achievement-Striving facet of Conscientiousness is very low. After extensive study of the literature, I believe that this subscale in particular is tapping into personality using questions that would be better served to be studied under an ideal point item response theory framework. "I do more than what's expected of me" sounds like a very extreme positive item on the Achievement-Striving subscale. "I work hard" sounds like a moderately positive item. "I do just enough work to get by" sounds like a moderately negative item. "I put little time and effort into my work sounds like a very extreme negative item.

#### **CSE, Self-Efficacy facet of Conscientiousness:**

In Table 1, we see the reliability of our Self-Efficacy subscale before and after outlier analysis. We see that after reverse scoring but before doing outlier analysis that the Self-Efficacy facet of Conscientiousness (CSE) shows a fairly high internal consistency as measured by Cronbach's alpha. This was expected because all the items at first glance seem to have similar strength on the domain in question. They also all happen to be positively worded items, which gets rid of the known negatively keyed factor that arises, which some have argued comes from the increased complexity of having to mentally reverse the item in their head (van Sonderen et al., 2013). Reliability for this scale stayed mostly the same after outlier analysis. This was a positively worded scale with items that at face value seemed to be of similar strength.

### **CO, Orderliness facet of Conscientiousness:**

In Table 1 is the reliability of our Orderliness subscale before and after outlier analysis. We see that after reverse scoring but before doing outlier analysis that the Orderliness facet of Conscientiousness (CO) shows a fairly high internal consistency as measured by Cronbach's alpha. This can be expected because 3 of the 4 items are negatively keyed, meaning if there is a complexity problem that arises from negatively worded items, then the majority of items in this case all have that same problem. It helps that all 3 of the negatively worded Orderliness items all seem to be of the same moderate negative strength on the facet, with the positively worded item also being of moderate strength but in a positive direction. We have found in this study that the more negatively worded compared to positively worded items a scale has, the more internal consistency it will have. Some interpretations and implications of this will be discussed during the thesis defense as well as in the discussion section of this paper. We see after outlier analysis that the Orderliness subscale has gone up in reliability by a small amount.

### **CD, Dutifulness facet of Conscientiousness:**

We see in Table 1 that after reverse scoring but before doing outlier analysis that the Dutifulness facet of Conscientiousness (CSE) shows a fairly low internal consistency measured by Cronbach's alpha. This is consistent with other findings where there is an equal number of positively and negatively worded items that internal consistency of the scale tends to be on the lower side. We suspect that one of the reasons this is the case is because of how the items seem to span the entire range of positive/negative as well as extreme/moderate items. This means that the confusion and differing results that

accompany negatively worded items in a scale does not have high consistency with the positively worded items, hence a violation of the unidimensionality assumption has occurred (Peters, 2014; Tavakol & Dennick, 2011; Yanyun Yang & Green, 2011).

#### **CSD, Self-Discipline facet of Conscientiousness:**

In Table 1, we see that the Self-Discipline facet of Conscientiousness (CSE) shows a fairly low internal consistency as measured by Cronbach's alpha. This is consistent with other findings where there is an equal number of positively and negatively worded items that internal consistency of the scale tends to be on the lower side. As mentioned above for the Dutifulness subscale, this is likely because of the violation of unidimensionality (Peters, 2014; Yanyun Yang & Green, 2011).

#### **CC, Cautiousness facet of Conscientiousness:**

Interestingly enough, the one subscale that shows the highest internal consistency is the scale with all negatively worded items. This likely taps into the same complexity factor mentioned earlier where people have to reverse the items that have a negative meaning to see if they do the positive version of it or not. People who struggled with answering one negatively worded item or carelessly responded to one negatively worded item, very likely did the same for all negatively worded items, hence the higher internal consistency of this scale.

#### **CON, Conscientiousness superordinate construct:**

Conscientiousness superordinate construct shows a very high internal consistency as measured by Cronbach's alpha. This high internal consistency is likely due to the number of items going into the scale since Cronbach's alpha is sensitive to number of

items. This also likely means that because we have a range of moderate and extreme items, we likely have lower Cronbach's alpha since people can disagree with moderate items from both a higher and lower level of the trait (Cappelleri, Lundy, & Hays, 2014).

In summary on the reliability of our original hypothesized variables, we see decent to high reliability. The researchers though did factor analysis which showed that these original hypothesized scales are not a good fit for our data. The factor analysis results and interpretation are below.

## Six Hypothesized Facets: Descriptive Statistics and Correlations

### 2. Descriptive Statistics and Correlations of Original Hypothesized Variables

Variable	Mean	Standard Deviation	GPA	CON	CSE	CAS	CSD	CO	CD	CC
<b>GPA</b>	3.4175	.40219	<b>(1)</b>							
<b>CON</b>	3.9510	.66570	.164**	<b>(.932)</b>						
<b>CSE</b>	3.8073	.76413	.126*	.652***	<b>(.830)</b>					
<b>CAS</b>	3.9721	.79162	.155**	.817***	.557***	<b>(.702)</b>				
<b>CSD</b>	3.8156	.81320	.201***	.895***	.683***	.723***	<b>(.725)</b>			
<b>CO</b>	3.7807	1.00599	.143*	.800***	.371***	.523***	.662***	<b>(.844)</b>		
<b>CD</b>	4.1892	.67205	.092	.784***	.342***	.533***	.641***	.561***	<b>(.720)</b>	
<b>CC</b>	4.1411	1.00402	.083	.787***	.240***	.570***	.569***	.576***	.686***	<b>(.926)</b>

N = 358 for personality traits. N = 280 for undergraduate GPA (not everyone had a GPA).

CON = Conscientiousness. CSE = Self-Efficacy facet. CAS = Achievement-Striving facet. CSD = Self-Discipline facet. CO = Orderliness facet. CD = Dutifulness facet. CC = Cautiousness facet.

\* =  $p \leq .05$       \*\* =  $p \leq .01$       \*\*\* =  $p \leq .001$

Cronbach's Alpha reliabilities in parentheses on the diagonals.

We see from Table 2 that our three facets with the hypothesized significant relationship with our outcome variable all showed the expected significant relationship. The three facets hypothesized to have a significant relationship with our outcome variable of Undergraduate GPA are: Self-Efficacy ( $r = .126$ ), Achievement-Striving ( $r = .155$ ),

and Self-Discipline ( $r = .201$ ). Orderliness also showed a significant relationship with our outcome variable and this relationship was unexpected ( $r = .143$ ).

### 3. Factor Analysis General Conscientiousness: Two Factor

*Summary of Exploratory Factor Analysis Results for a General Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

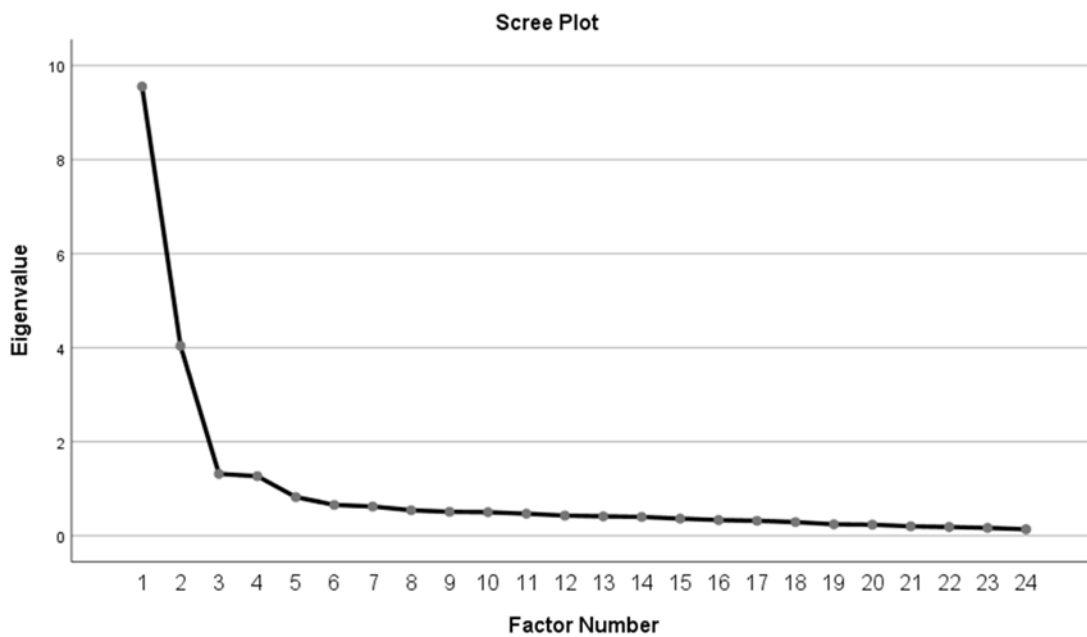
Item	Factor Loadings		Communalities Extraction
	Negative Conscientiousness	Positive Conscientiousness	
I complete tasks successfully.		.753	.579
I excel in what I do.		.744	.494
I handle tasks smoothly.		.699	.497
I know how to get things done.		.716	.521
I like to tidy up.		.535	.302
I often forget to put things back in their proper place. RK	.714		.527
I leave a mess in my room. RK	.708		.584
I leave my belongings around. RK	.683		.539
I keep my promises.		.525	.348
I tell the truth.		.372	.226
I break the rules. RK	.750		.529
I break my promises. RK	.813		.614
I do more than what's expected of me.		.696	.446
I work hard.		.686	.508
I put little time and effort into my work. RK	.782		.571
I do just enough work to get by. RK	.643		.475
I am always prepared.		.703	.488
I carry out my plans.		.747	.551
I waste my time. RK	.664		.616
I have difficulty starting tasks. RK	.666		.574
I jump into things without thinking. RK	.856		.691
I make rash decisions. RK	.842		.677
I rush into things. RK	.854		.684
I act without thinking. RK	.801		.633
<b>Eigenvalues</b>	9.553	4.037	
<b>% of variance</b>	39.8%	16.8%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor 1 had a correlation with factor 2 of .368. This is extremely bad for what is supposed to be the relationship between conscientiousness and itself. This poor convergent validity is showing that the negatively worded factor is measuring a separate construct. The assumptions of the researchers are that this is because of the nuisance factor mentioned elsewhere in the manuscript.

*1. Scree Plot for Generalized Conscientiousness Exploratory Factor Analysis*





#### 4. Factor Analysis Academic Conscientiousness: Two Factor

*Summary of Exploratory Factor Analysis Results for an Academic Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings		Communalities Extraction
	Negative Conscientiousness	Positive Conscientiousness	
I like to have a structured approach for my school assignments.		.682	.465
I follow a schedule to work on school assignment.		.769	.572
I follow a schedule at school.		.710	.514
I take organized notes in class.		.667	.457
I cheat on tests. RK	.786		.571
I break rules at school. RK	.778		.563
I cheat on homework assignments. RK	.784		.570
I finish the school assignments by their deadline.		.513	.365
I do more than what is expected of me at school.		.713	.469
I work hard at school.		.755	.606
I put little time and effort into my schoolwork. RK	.783		.643
I aim to get a perfect grade on assignments.		.686	.459
I do just enough work to get my school assignments done. RK	.605		.430
I delay finishing my school assignments. RK	.725		.680
I find it difficult to get started on schoolwork. RK	.702		.665
I find it difficult to get rid of distractions and get my school assignments done. RK	.701		.569
I get school assignments done quickly.		.666	.424
I get distracted when doing assignments. RK	.686		.559
I follow a schedule to work on assignments.		.732	.560
I make myself study during specific hours.		.689	.439
I allocate my time for school work.		.777	.623
I can focus enough to get my school assignments done on time.		.623	.495
I postpone studying for an exam. RK	.662		.557
I find it difficult to complete my assignments on time. RK	.755		.659
I waste my time when I am supposed to work on academic projects. RK	.726		.646
I am easily distracted when studying. RK	.704		.625
I am easily distracted when doing assignments. RK	.718		.589
I jump into things without thinking at school. RK	.866		.704

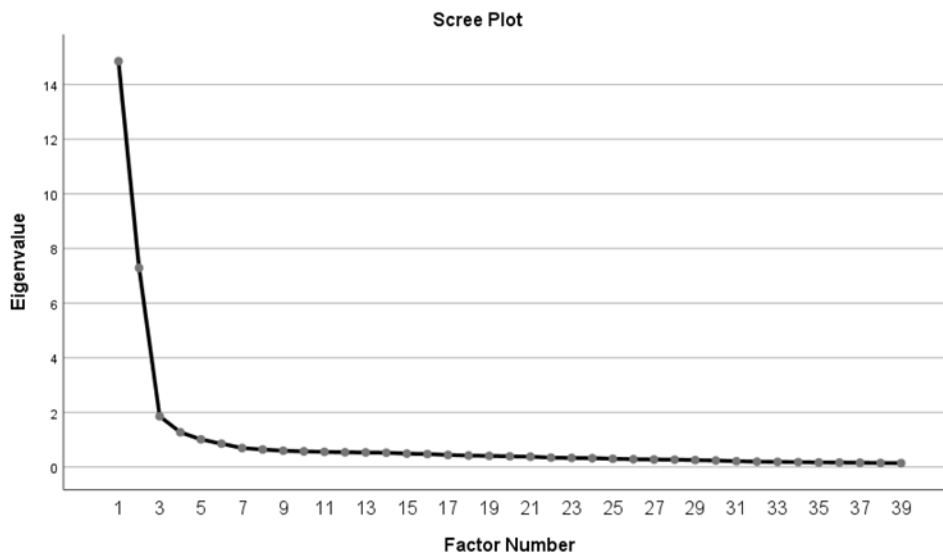
I rush into things at school. RK	.848		.677
I act without thinking at school. RK	.848		.677
I ask questions in class without first thinking it through. RK	.720		.478
I complete my school assignments successfully.		.669	.554
I excel in what I do at school.		.726	.528
I feel it's easy to keep up with all the school assignments I have.		.628	.372
I know how to get things done at school.		.696	.526
I can focus enough to get my school assignments done on time.		.609	.466
I perform below expectation on tests or assignments at school. RK	.666		.436
I complete my tasks at school successfully.		.718	.540
I am prepared to take on any school assignment.		.721	.518
Eigenvalues	14.853	7.286	
% of variance	38.1%	18.7%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor 1 had a correlation with factor 2 of .283. This is extremely bad for what is supposed to be the relationship between conscientiousness and itself. This poor convergent validity is showing that the negatively worded factor is measuring a separate construct. The assumptions of the researchers are that this is because of the nuisance factor mentioned elsewhere in the manuscript.

## 2. Scree Plot for Academic Conscientiousness Exploratory Factor Analysis



### Four Factor Analysis Derived Variables: Reliabilities

Given our exploratory factor analysis results, it was reasonable to look at the reliabilities and validities of our four-factor analysis derived variable. The two-factor analysis derived variables for academic conscientiousness were named positively and negatively worded academic conscientiousness. The two-factor analysis derived variables for our general conscientiousness scale were named positively and negatively worded general conscientiousness. The variable names were chosen because all the negatively worded items for each scale loaded onto their own factor and all the positively worded items loaded onto their own factor as well.

### 5. Reliabilities for Factor Analysis Derived Variables:

Title	Description	Cronbach's Alpha
<b>PAcaCon</b>	Positively Worded Academic Conscientiousness	.948
<b>NAcaCon</b>	Negatively Worded Academic Conscientiousness	.961
<b>PCon</b>	Positively Worded General Conscientiousness	.892
<b>NCon</b>	Negatively Worded General Conscientiousness	.946

**N= 358**

We see above in Table 4 that our four-factor analysis derived variables all have high reliabilities. We also see that the negatively worded factor reliabilities are higher than the positively worded reliabilities for both academic and general conscientiousness. Between the reliabilities and factor analysis results we can conclude that our results need to be interpreted using the positively and negatively worded factors instead of the originally hypothesized factors.

### Four Factor Analysis Derived Variables: Descriptives and Correlations

Since we have concluded from the reliabilities and exploratory factor analysis results that these 4 factors are reasonable to use for interpretive purposes, we now go into looking at the descriptive statistics and correlations of our 4 factors as well as our outcome variable, Undergraduate GPA.

## 6. Descriptive Statistics and Correlations of GPA and Factor Analysis Derived Variables

Variable	Mean	Standard Deviation	GPA	PAcaCon	NAcaCon	PCon	NCon
<b>GPA</b>	3.4175	.40219	<b>(1)</b>				
<b>PAcaCon</b>	3.6813	.75470	.269***	<b>(.948)</b>			
<b>NAcaCon</b>	4.0986	.87574	.178**	.358***	<b>(.961)</b>		
<b>PCon</b>	3.8103	.67460	.163**	.728***	.313***	<b>(.892)</b>	
<b>NCon</b>	4.0700	.88970	.124*	.289***	.886***	.387***	<b>(.946)</b>

**N= 358 for personality traits. N = 280 for undergraduate GPA (not everyone had a GPA).**

**CON = Conscientiousness. CSE = Self-Efficacy facet. CAS = Achievement-Striving facet. CSD = Self-Discipline facet. CO = Orderliness facet. CD = Dutifulness facet. CC = Cautiousness facet.**

**\* $\leq$ .05      \*\* $\leq$ .01      \*\*\* $\leq$ .001**

**Cronbach's Alpha reliabilities in parentheses on the diagonals.**

Results of bivariate correlational analysis of our four-factor analysis derived variables, found in Table 6, supports that the positively and negatively worded factors are not measuring the same construct, since the relationship between the positively and negatively worded items is so low. Positively worded academic conscientiousness has an  $r = .358$  relationship with the negatively worded academic conscientiousness variable. Positively worded academic conscientiousness though has a strong  $r = .728$  relationship with the positively worded general conscientiousness variable. An  $r$  of  $.728$  is still on the lower end of convergent validity results.

What we see from the relationship between these two constructs and our outcome variable of interest, Undergraduate GPA, is that the differences between these two predictor constructs explains extra variance in our outcome variable. This shows at the correlational level that positively worded academic conscientiousness, while having

moderate convergent validity with the positively worded general conscientiousness variable, can be interpreted to have incremental validity in the variance that positively worded academic conscientiousness does not have in common with positively worded general conscientiousness.

The same can be said for the negatively worded academic and general conscientiousness convergent validity results and correlational relationship with our outcome variable of interest. The negatively worded items have more in common with each other than they do the positively worded items of their respective academic and general conscientiousness scales, but the variance that negatively worded academic conscientiousness does not have in common with negatively worded general conscientiousness explains extra variance in our outcome variable of interest, Undergraduate GPA.

### **General Conscientiousness Confirmatory Factor Analysis: Models and Interpretations**

With all the exploratory factor analysis, reliabilities, descriptive statistics, and correlations showing the expected results for general conscientiousness, given the presence of a nuisance factor, we can conclude that confirmatory factor analysis is our next necessary step for interpreting our results. Confirmatory factor analysis restricts cross-loadings at the expense of having higher correlations between factors. Given the low correlations between our factors in earlier analysis, higher factor correlations are not

expected to be much of a problem though we will see models that try to account for that by having factors covary.

### 7. Chi-Square of All Estimated General Conscientiousness Models

<b>Title</b>	<b>Description</b>	$\chi^2$	$\Delta \chi^2$	<b>Df</b>	$\Delta$ <b>Df</b>	<b>P</b>
<b>M01</b>	Independence Model 1	4957.370		300		
<b>M02</b>	General Conscientiousness, 1 factor, all items in model	2007.627	2949.743	275	25	<.00001
<b>M03</b>	General Conscientiousness, 2 factors uncorrelated, all items in model	1076.444	931.183	274	1	<.00001
<b>M04</b>	General Conscientiousness, 2 factors correlated, all items in model	1038.208	38.236	273	1	<.00001
<b>M05</b>	General Conscientiousness 2 factors and higher order factor all items in model	1038.204	.004	270	3	<.00001
<b>M06</b>	Independence Model 2	1333.108		66		
<b>M07</b>	General Conscientiousness 1 factor solution no negatively worded items	127.794	1205.314	54	12	<.00001
<b>M08</b>	General Conscientiousness, 1 factor solution with no negatively worded items, 2 covariances (E11, 10) and (E15, 14)	94.556	33.238	52	2	.00028

Note. N = 358.  $\chi^2$  = Chi-square. Df = degrees of freedom.

Because chi-square is so dependent on sample size and complexity of models, chi-square by itself is not a good measure of model fit. The above information is the basis for our model fit indices though, so they are presented for the sake of both clarity and thoroughness. We see continued decrement of chi square with the addition of new parameters and the reduction of degrees of free.

## 8. Model Fit of All Estimated General Conscientiousness Models

<b>Title</b>	<b>Description</b>	<b>RMSEA [95% CI]</b>	<b>ΔRMSEA</b>	<b>SRMR</b>	<b>ΔSRMR</b>	<b>CFI</b>	<b>ΔCFI</b>
<b>M01</b>	Independence Model 1						
<b>M02</b>	General Conscientiousness, 1 factor, all items in model	0.150 (0.144, 0.156)		0.156		0.628	
<b>M03</b>	General Conscientiousness, 2 factors uncorrelated, all items in model	0.102 (0.096, 0.109)	-.048	0.157	+.001	0.828	+.200
<b>M04</b>	General Conscientiousness, 2 factors correlated, all items in model	0.100 (0.094, 0.107)	-.002	0.072	-.085	0.836	+.008
<b>M05</b>	General Conscientiousness, 2 factors and higher order factor, all items in model	0.101 (0.094, 0.107)	+.001	0.072	0	0.835	-.001
<b>M06</b>	Independence Model 2						
<b>M07</b>	General Conscientiousness 1 factor solution with no negatively worded items	0.070 (0.054, 0.085)	-.031	0.047	-.025	0.942	+.107
<b>M08</b>	General Conscientiousness, 1 factor solution with no negatively worded items, 2 covariances (E11, 10) and (E15, 14)	0.054 (0.036, 0.071)	-.016	0.040	-.007	0.966	+.025

Note. N = 358. CFI = Confirmatory Fit Index. SRMR = Standardized Root Mean Square Residual. RMSEA = Root Mean Square Error of Approximation.

Hu & Bentler put the cutoff for CFI at .95 or higher, for SRMR at .08 or lower, and RMSEA at .06 or lower (Hu & Bentler, 1999). Hu & Bentler (1999) also recommend sample sizes of greater than 250 in order to have confidence that those model fit indices are accurate. Since the sample here is 280, we are within the sample size necessary to be confident that those model fit rules are accurate.



Tables 8 shows us our general conscientiousness models and their fits starting with the most simplistic of models, 1 factor without removing any items (M02). The comparative fit index (CFI) = 0.628, the Standardized Root Mean Square Residual (SRMR) = 0.156, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.150 (0.144, 0.156). We can see given the suggested cutoff scores from Hu & Bentler that this model is an extremely poor fit.

Our CFA results then show that two factors, a positively and negatively worded factor for the positively and negatively worded indicators, without a higher order factor (M03), also shows poor model fit. The comparative fit index (CFI) = 0.828, the Standardized Root Mean Square Residual (SRMR) = 0.157, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.102 (0.096, 0.109). Model 4, is our model with 2 factors correlated and all items still in the model. The comparative fit index (CFI) = 0.836, the Standardized Root Mean Square Residual (SRMR) = 0.072, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.100 (0.094, 0.107). Our final model before removing items is model 5, which has 2 factors and a higher order factor, while retaining all items in the model. The comparative fit index (CFI) = 0.835, the Standardized Root Mean Square Residual (SRMR) = 0.072, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.101 (0.094, 0.107). This model fit also does not reach the levels recommended.

Because our assumption is that the negatively worded items represent a nuisance factor, we suspect that they are harming our model fit. The reason they are suspected of being a nuisance factor is because they do not have the relationships with variables that

they are supposed to predict based on findings in other studies. Prior literature has also considered all of the big five personality traits in terms of the positive end, aside from neuroticism, which is sometimes discussed in terms of the negative end because of the negative end being considered the adaptive ends of the scale. In our next model, all negatively worded items are removed.

Model 7 is our general conscientiousness 1 factor model, now with no negatively worded items. The comparative fit index (CFI) = 0.942, the Standardized Root Mean Square Residual (SRMR) = 0.047, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.070 (0.054, 0.085). We can see clearly that using rational thinking about what changes to make, before going to a mechanical/statistical solution, can significantly increase model fit. Our model fit for model 7 is almost to the point where it can be considered acceptable, so now we turn to mechanical/statistical solutions to see what our best possible solutions are to increase model fit. For this, we are using something called the Lagrange Multiple test (LMTest).

9. LMTest for Adding Parameters to General Conscientiousness Scale										
CUMULATIVE MULTIVARIATE STATISTICS					UNIVARIATE INCREMENT					
STEP	PARAMETER	CHI-SQUARE	D.F.	PROB.	CHI-SQUARE	PROB.	HANCOCK'S SEQUENTIAL		PREDICTED	
							D.F.	PROB.	RMSEA	CFI
1	E11,E10	19.702	1	0.000	19.702	0.000	54	1.000	0.061	0.957
2	E15,E14	32.779	2	0.000	13.077	0.000	53	1.000	0.054	0.966

As we can see from Tables 7, 8, and 9, model fit across the board increases once we add a covariance for the errors of the dutifulness items (E11,E10) and for the achievement-striving errors (E15, 14). The comparative fit index (CFI) = 0.966, the

Standardized Root Mean Square Residual (SRMR) = 0.040, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.054 (0.036, 0.071). This means that our model without negatively worded items and 1 factor solution worked the best; after making a few adjustments using covariances for two of the hypothesized facets. Because of the nuisance factor that was created for the negatively worded items, we only had two items for each the dutifulness and achievement striving facets. The covariance between the two dutifulness items and between the two achievement-striving items is interpreted to mean that if we had more positively worded items for each of these facets we might have found the predicted facets.

### **Academic Conscientiousness Confirmatory Factor Analysis: Models and Interpretations**

With all the exploratory factor analysis, reliabilities, descriptive statistics, and correlations showing the expected results for academic conscientiousness, given the presence of a nuisance factor, we can conclude that confirmatory factor analysis is our next necessary step for interpreting our results. Confirmatory factor analysis restricts cross-loadings at the expense of having higher correlations between factors. Given the low correlations between our factors in earlier analysis, higher factor correlations are not expected to be much of a problem though we will see models that try to account for that by having factors covary.

## 10. Chi-Square of All Estimated Academic Conscientiousness Models (GCM)

Title	Description	$\chi^2$	$\Delta \chi^2$	Df	$\Delta Df$	P
<b>GCM01</b>	Independence Model 1	9429.091		780		
<b>GCM02</b>	Academic Conscientiousness 1 factor all items in model	4848.284	4580.807	740	-40	<.00001
<b>GCM03</b>	Academic Conscientiousness 2 factors uncorrelated all items in model	2662.199	-2186.085	739	-1	<.00001
<b>GCM04</b>	Academic Conscientiousness 2 factors correlated all items in model	2620.856	-41.343	738	-1	<.00001
<b>GCM05</b>	Academic Conscientiousness 2 factors and higher order factor all items in model	2620.854	-.002	735	-3	<.00001
<b>GCM06</b>	Independence Model 2	3513.642		210		
<b>GCM07</b>	Academic Conscientiousness 1 factor solution with no negatively worded items	662.378	-2851.264	189	-21	<.00001
<b>GCM08</b>	Independence Model 3	3081.770		171		
<b>GCM09</b>	Academic Conscientiousness 1 factor solution with no negatively worded items, V26 and V45 removed, covariance (E44, E27) and (E28, E27)	403.026	-2678.744	150	-21	<.00001

Note. N = 358.  $\chi^2$  = Chi-square. Df = degrees of freedom.

Because chi-square is so dependent on sample size and complexity of models, chi-square by itself is not a good measure of model fit. The above information is the basis for our model fit indices though, so they are presented for the sake of both clarity and thoroughness. We see continued decrement of chi square with the addition of new parameters and the reduction of degrees of free.

## 11. Model Fit of All Estimated Academic Conscientiousness Models (ACM)

Title	Description	RMSEA [95% CI]	ΔRMSEA	SRMR	ΔSRMR	CFI	ΔCFI
ACM01	Independence Model 1						
ACM02	Academic Conscientiousness 1 factor all items in model	0.141 (0.137, 0.145)		0.188		0.525	
ACM03	Academic Conscientiousness 2 factors uncorrelated all items in model	0.097 (0.092, 0.100)	-.044	0.167	-.021	0.778	+.253
ACM04	Academic Conscientiousness 2 factors correlated all items in model	0.096 (0.092, 0.099)	-.001	0.099	-.068	0.782	+.004
ACM05	Academic Conscientiousness 2 factors and higher order factor all items in model	0.096 (0.092, 0.100)	0	0.099	0	0.782	0
ACM06	Independence Model 2						
ACM07	Academic Conscientiousness 1 factor solution with no negatively worded items	0.095 (0.087, 0.102)	-.001	0.060	-.039	0.857	+.075
ACM08	Independence Model 3						
ACM09	Academic Conscientiousness 1 factor solution with no negatively worded items, V26 and V45 removed, covariance (E44, E27) and (E28, E27)	0.078 (0.068, 0.087)	-.017	0.049	-.011	0.913	+.056

Note. N = 358. CFI = Confirmatory Fit Index. SRMR = Standardized Root Mean Square Residual. RMSEA = Root Mean Square Error of Approximation.

Table 11 shows us our academic conscientiousness models and their fits starting with the most simplistic of models, a one factor solution with all items in the model (ACM02). The comparative fit index (CFI) = 0.525, the Standardized Root Mean Square Residual (SRMR) = 0.188, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.141 (0.137, 0.145). This model fit is not considered acceptable given the standard that Hu & Bentler have stated in their 1999 article. Model 3 looks at a two-factor solution with all items uncorrelated. The comparative fit index (CFI) = 0.778, the Standardized Root Mean Square Residual (SRMR) = 0.167, and the Root Mean Square

Error of Approximation (RMSEA) [95% CI] = 0.097 (0.092, 0.100). This has significantly better fit than the prior model but not nearly good enough to suggest it is close to the true model. Our fourth model looks at 2 factors again but now with a covariance between them and includes all items in the model. The comparative fit index (CFI) = 0.782, the Standardized Root Mean Square Residual (SRMR) = 0.099, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.096 (0.092, 0.099). This model shows slightly better fit than the prior model but not good enough. Model 5 is our final model before removing the negatively worded items. Model 5, instead of having a covariance between the 2 factors, has a higher order factor and has all items in the model. The comparative fit index (CFI) = 0.782, the Standardized Root Mean Square Residual (SRMR) = 0.099, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.096 (0.092, 0.100).

Model 7 is our 1 factor solution model with no negatively worded items. The comparative fit index (CFI) = 0.857, the Standardized Root Mean Square Residual (SRMR) = 0.060, and the Root Mean Square Error of Approximation (RMSEA) [95% CI] = 0.095 (0.087, 0.102). This model fit was still extremely poor. Because of this still extremely poor model fit, we decided to look at item level solutions going forward since this is an under-construction scale of academic conscientiousness. It was decided based on LMTest, regression, and mediation analysis, that V26 and V45 would be removed and also that the error variances of items 44 and 27 would covary, as well as the error variances of items 28 and 27 would covary. The comparative fit index (CFI) = 0.913, the Standardized Root Mean Square Residual (SRMR) = 0.049, and the Root Mean Square

Error of Approximation (RMSEA) [95% CI] = 0.078 (0.068, 0.087). This model fit is still somewhat poor in terms of CFI and RMSEA, this model was accepted though because of the level of model complexity and because Hu & Bentler have stated for more complex models that lower model fit is reasonable (1999). Continued testing in upcoming research will be done to lower scale length and reach a greater degree of parsimony in our model.

## 12. LMTest for Adding Parameters to Academic Conscientiousness Scale

CUMULATIVE MULTIVARIATE STATISTICS					UNIVARIATE INCREMENT					
STEP	PARAMETER	CHI-SQUARE	D.F.	PROB.	CHI-SQUARE	PROB.	HANCOCK'S SEQUENTIAL D.F.	PROB.	PREDICTED RMSEA	CFI
1	V27,E44	41.850	1	0.000	41.850	0.000	189	1.000	0.091	0.869
2	V45,E27	84.757	2	0.000	42.907	0.000	188	1.000	0.087	0.882
3	V27,E26	127.601	3	0.000	42.844	0.000	187	1.000	0.082	0.894
4	V27,E28	166.279	4	0.000	38.678	0.000	186	1.000	0.078	0.906
5	V44,E26	198.799	5	0.000	32.521	0.000	185	1.000	0.074	0.915
6	E45,E44	233.266	6	0.000	34.466	0.000	184	1.000	0.069	0.926
7	V46,E27	257.790	7	0.000	24.525	0.000	183	1.000	0.066	0.933
8	V45,E57	281.071	8	0.000	23.281	0.000	182	1.000	0.063	0.939
9	V44,E63	302.353	9	0.000	21.282	0.000	181	1.000	0.060	0.946
10	V33,E61	322.721	10	0.000	20.368	0.000	180	1.000	0.057	0.951
11	V33,E47	342.658	11	0.000	19.936	0.000	179	1.000	0.053	0.957
12	V45,E26	359.719	12	0.000	17.062	0.000	178	1.000	0.050	0.962
13	V28,E26	375.093	13	0.000	15.374	0.000	177	1.000	0.048	0.966
14	E45,E28	390.392	14	0.000	15.298	0.000	176	1.000	0.045	0.971
15	V44,E46	405.154	15	0.000	14.763	0.000	175	1.000	0.041	0.975
16	V45,E46	420.974	16	0.000	15.820	0.000	174	1.000	0.038	0.979
17	V47,E61	433.849	17	0.000	12.875	0.000	173	1.000	0.034	0.983
18	V27,E63	446.527	18	0.000	12.677	0.000	172	1.000	0.031	0.986
19	V63,E57	457.027	19	0.000	10.500	0.001	171	1.000	0.027	0.989
20	V57,E33	467.312	20	0.000	10.284	0.001	170	1.000	0.024	0.992

Table 12 shows the original LMTest that we used to decide on what to look at, at the item level. Since Academic Conscientiousness is a scale that is a work in progress,

items were all test cases. This means that when the LMTest is showing the same five items as causing problems, that we decided it was necessary to look at these indicators at the item level. We found items that these items did not have the type of predictive validity we would expect to find. Examples of items that the LMTest found were related are: “I follow a schedule to work on assignments” and “I follow a schedule to work on school assignments”. These were considered partially and fully contextualized versions, respectively, of a generalized type of question such as “I follow a schedule to work on tasks”. Assignments is usually related to work or school, whereas tasks is a term considered to be more general in nature. School assignments is an even more contextualized version of assignments, so it was expected to fully explain the relationship between the partially contextualized version of the item and the outcome variable. Linear regression below shows that the fully contextualized version does explain all of the relationship with the outcome variable, with the partially contextualized version of the item having an insignificant relationship with undergraduate GPA when the fully contextualized version is in the equation.

### 13. Multiple Regression for Academic Conscientiousness Items: Part 1

Items	B	SE	$\beta$	t	p
I follow a schedule to work on school assignment.	.077	.031	.208	2.510	.013
I follow a schedule to work on assignments.	.014	.030	.040	.479	.633

**N = 280**

Multiple regression analysis of our indicators showed that “I follow a schedule to work on assignments added nothing extra to the relationship our predictor, academic conscientiousness, has with the outcome variable, Undergraduate GPA. The relationship



between “I follow a schedule to work on assignments” and Undergraduate GPA is ( $B = .014, p = .633$ ). However, the fully contextualized item shows a relationship of .077 with the outcome variable ( $p = .013$ ), so we can say that the fully contextualized item adds incremental validity to the relationship between conscientiousness and college GPA. The partially contextualized item was kept because while it added nothing unique to the relationship with the outcome variable, it was still a significant variable in a sample of very low-quality data.

LMTTest showed that the item “I follow a schedule to work on school assignments” was also related to 3 other items, so they were all looked at. For reasons unknown, “I follow a schedule to work on school assignments” and “I make myself study during specific hours” were considered to be highly related. While the former item could be suspected of being on a potential orderliness or self-discipline facet, the latter item is very clearly self-discipline related.

#### 14. Multiple Regression for Academic Conscientiousness Items: Part 2

Items	B	SE	$\beta$	t	p
I make myself study during specific hours.	-.036	.025	-.106	-1.415	.158
I follow a schedule to work on school assignment.	.112	.028	.304	4.041	.000

**N = 280**

“I make myself study during specific hours” was an item that in retrospect probably has little relationship with the outcome variable despite potentially being related to the construct of self-discipline. For criterion validity reasons, this item was removed. We see in direct comparisons of “I follow a schedule to work on school assignments” and “I make myself study during specific hours” that the latter has a trending towards

negative relationship with our outcome variable. Because of these concerns raised by the LMTest and regression analysis which shows lower criterion validity and ultimately construct validity of our measures because of these two items, “I make myself study during specific hours”, named V45, was removed for our final CFA model.

### 15. Multiple Regression for Academic Conscientiousness Items: Part 3

Items	B	SE	$\beta$	t	p
I follow a schedule to work on school assignment.	.088	.029	.239	3.022	.003
I follow a schedule at school.	.002	.031	.004	-.054	.957

**N = 280**

It is clear that a wording effect is arising from having “I follow a schedule at school” and “I follow a schedule to work on school assignments”. They both have very similar wording. Linear regression showed that “I follow a schedule to work on school assignments” fully explained the relationship with our outcome variable and that “ I follow a schedule at school” added nothing to the relationship. The linear regression results are below. Note that these items also have another data set showing a relationship with the outcome variable and items will be looked at to see if they had the same lack of relationship with the outcome variable when controlling for this item. Some items that might be tapping into the same facet but from slightly different ways may be kept in the next round of the academic conscientiousness scale even if they have no relationship with

the outcome variable after controlling for “I follow a schedule to work on school assignments”.

#### 16. Multiple Regression for Academic Conscientiousness Items: Part 4

Items	B	SE	$\beta$	t	p
I follow a schedule to work on school assignment.	.112	.029	.303	3.853	.000
I like to have a structured approach for my school assignments.	-.036	.029	-.099	-1.257	.210

**N = 280**

We see above that “I like to have a structured approach for my school assignments” was starting to be negatively related to our outcome variable when controlling for “I follow a schedule to work on school assignments”. Because of these concerns raised by the LMTest and regression analysis which shows lower criterion validity and ultimately construct validity of our measures because of these two items, “I like to have a structured approach for my school assignments”, named V26, was removed for our final CFA model.

#### Academic Conscientiousness: Summary Findings for Item-Level Scale Changes

Because of the low-quality data problem and the lack of specifically looking for facets in our data, items with overly significant covariance that are trending negatively with our outcome variable of interest are going to be removed from the scale. This adds construct validity to our scale that is under-construction by increasing one aspect of construct validity: criterion validity. This benefits us further because the academic conscientiousness scale had too many items compared to the general conscientiousness

scale. We want the scales to be as comparable as possible in terms of number of items in order to make comparisons between more equivalent scales.

“I like to have a structured approach for my school assignments”, named V26 and “I make myself study during specific hours”, named V45, were removed because they were both causing problems with model fit as well as having no relationship with the outcome variable of interest. This is potentially because of the low-quality data problem, so the items will only be fully removed from the scale after looking at other data sets. In summary, by adding and removing parameters for variables in the first five LMTTest results, our model fit indices went from CFI = .857 to .913, SRMR = .06 to .049, and RMSEA = .095 to .078. This was considered acceptable for reasons described earlier, so this was the final academic conscientiousness model used for the rest of our CFA and SEM analyses.

### **Conscientiousness Scales: Summary Findings for EFA and CFA**

Exploratory and Confirmatory Factor Analysis (EFA/CFA) was done both on general and academic conscientiousness scales. We found that the two-factor structure represented the data best when running exploratory factor analysis in SPSS. This represented the data best for both the general and the academic conscientiousness scales. After running confirmatory factor analysis in EQS, we found that the negatively worded items, which were a nuisance factor, disrupted the model fit to such an extent that removing them was the only way to achieve good model fit. This was expected because the factors were only created because of low quality data.

## Academic Conscientiousness: Item-Level Incremental Validity and Factor Structure

### Interpretation

While stepwise regression is not a type of regression that should normally be done because it is a mechanical solution to what is a theoretical problem that should have theoretical and methodological solutions, because of the low-quality data problem, we wanted to see if we could find anything in our data showing a facet structure. Stepwise regression essentially says which item adds the most to the prediction, and then which item adds something above and beyond that, and then so on and so forth. This type of analysis was predicted to offer us some evidence of at the very least a two-factor solution for academic conscientiousness without a nuisance factor given the results of our exploratory factor analyses, which can be found in the appendix under tables.

The exploratory factor results for more than two factors of conscientiousness seemed to imply that industriousness and orderliness domains or something similar could be found. This follows with results by others on facets potentially being subsumed by “domains” (DeYoung, Quilty, & Peterson, 2007). Of our original hypothesized six facets, three were expected to be subsumed by the industriousness domain: achievement-striving, self-efficacy- and self-discipline. These were the three facets that were expected to be significantly related to our outcome variable. The other three facets: dutifulness, orderliness, and cautiousness, are considered facets that would fall under the orderliness domain.

### 17. Stepwise Regression for Academic Conscientiousness Items

Model	Items	B	SE	$\beta$	t	Significance
1	I excel in what I do at school.	.114	.024	.269	4.651	.000
2	I excel in what I do at school.	.082	.028	.193	2.907	.004
	I take organized notes in class.	.056	.025	.152	2.288	.023
N = 280						

What we found, as can be seen in Table 17, was that two items were very representative of our items and they seem to represent our belief of a two-facet structure in our data despite all the noise from the nuisance factor. The first item “I excel in what I do at school” a partially contextualized item that has a tag of “at school” onto a generalized conscientiousness item and likely represents something like an industrious domain in a two-factor structure is one item. The second item is “I take organized notes in class” which likely represents more of an orderliness facet. These two items explain all the prediction at the item level of our variables, with the item that is suspected of being on the orderliness domain offering some extra prediction but the item on the industriousness domain is our best predictor. This partially supports our original hypothesized relationships between facets and our outcome variable, Undergraduate GPA.

### Results for Original Hypothesis:

Measurement invariance analysis, comparing groups to see if there are differences in the data, has been used traditionally for subgroup differences like gender and race (Millsap, 2011). It also has a less well-known use to look at how different methods of looking at a construct compare to each other (Maul, 2013). Our primary original hypothesis, which is still testable to a degree, is that we should try to constrain general conscientiousness and academic conscientiousness paths to our outcome variable of undergraduate GPA to see if there is incremental validity to using academic conscientiousness for prediction of undergraduate GPA.

18. Path Constraints for Academic and General Conscientiousness Relationship with College GPA							
NUMBER	PARAMETER	CHI-SQUARE	PROBABILITY	CHANGE	D.F.	RMSEA	CFI
1	(1,V1,F1)-(2,V1,F1)	1.720	0.190	-0.046	203	0.072	0.929
CONSTRAINT TO BE RELEASED: (1,V1,F1)-(2,V1,F1)=0;							

Constraining our paths was not shown to lead to a significant change in chi-square. This was likely because of how large chi-square is because we are comparing two entire factor structures with all items that were kept in the final models of the CFAs. While the analysis of constraining paths was done correctly, the outcome was of little benefit because our factor structure did not turn out anywhere near what we predicted because of our nuisance factor. We believe that if we had our scales for each of the facets and was able to enter a correlation or covariance matrix instead of all of the raw data into EQS, that we would have found significant results when constraining paths. Since this analysis did not properly allow us to look at incremental validity, we will be using

mediation and multiple regression analysis to show that our academic contextualized scale shows incremental validity.

### 19. Multiple Regression of Academic Conscientiousness

Model	Items	B	SE	$\beta$	t	p
<b>1</b>	PAcaCon	.188	.049	.335	3.842	.000
	PCon	-.054	.053	-.089	-1.016	.310

**N = 280**

We see from Table 18 that positively worded general conscientiousness has no relationship with our outcome variable of interest, Undergraduate GPA, when positively worded academic conscientiousness is in the equation. Positively worded academic conscientiousness meanwhile has a relationship with our outcome variable, such that (B = .188, SE = .049, t = 3.842, P = <.001). In fact, we see a negative relationship where once the contextualized measure explains all the shared variance with the outcome variable, the other predictor, positively worded general conscientiousness, is suppressed (B = -.054, SE = .053, t = -.089, P = .310).

### 20. Mediation Analysis of Academic Conscientiousness: Part 1

Model	Predictor	Outcome	B	SE	t	P	LLCI	ULCI
<b>1</b>	PCon	PAcaCon	.8067	.0427	18.8777	.0000	.7226	.8908
<b>2</b>	PCon	UnderGPA	-.0535	.0527	-1.0161	.3105	-.1572	.0502
	PAcaCon		.1880	.0489	3.8416	.0002	.0917	.2844
<b>3</b>	PCon	UnderGPA	.0982	.0357	2.7474	.0064	.0278	.1685

**N = 280.** PAcaCon= Positively Worded Academic Conscientiousness. PCon = Positively Worded General Conscientiousness. UnderGPA = Undergraduate GPA



Here we see three models for mediation analysis. The first model is considered the “a” path in mediation analysis and it is when our x variable (PCon) is predicting our mediator variable (PAcaCon). We see a significant relationship occurring here ( $B = .8067$ ,  $SE = .0427$ ,  $t = 18.877$ ,  $P = <.0001$ ). This tells us that the two variables have a lot of shared variance. We next see in model 2, the prediction of our y variable (UnderGPA) by both our x variable and our mediator. We see that positively worded academic conscientiousness explains the entirety of the relationship with our outcome variable ( $B = .188$ ,  $SE = .0489$ ,  $t = 3.8416$ ,  $P = .0002$ ), to such an extent that suppression is occurring ( $B = -.054$ ,  $SE = .053$ ,  $t = -.089$ ,  $P = .310$ ). In model 2, the relationship between our mediator and the outcome variable is our “b” path. This same result was found in our multiple regression results, so model 2 is the same as Table 18. Model 3 is our total effects model for our x variable (PCon). Model 1 gives us our a path and model 2 where the mediator to outcome variable coefficient is, gives us our b path. If we multiply a and b paths, we get the indirect effect which is  $a*b = .1517$ , if we add this to the model 2 path from x variable to outcome variable, we would get the total effect model, which is model 3. This equation is  $.1517 - 0.0535 = .0982$ . The different signs between our direct and indirect effect tells us that inconsistent mediation is going on. This model shows us the results we would get if we multiply a and b paths then add c’ path. The results show that positively worded general conscientiousness does have a relationship with our outcome variable ( $B = .0982$ ,  $SE = .0357$ ,  $t = 2.7474$ ,  $P = .006$ ), but model 2 shows us that the relationship is entirely mediated by context (PAcaCon).

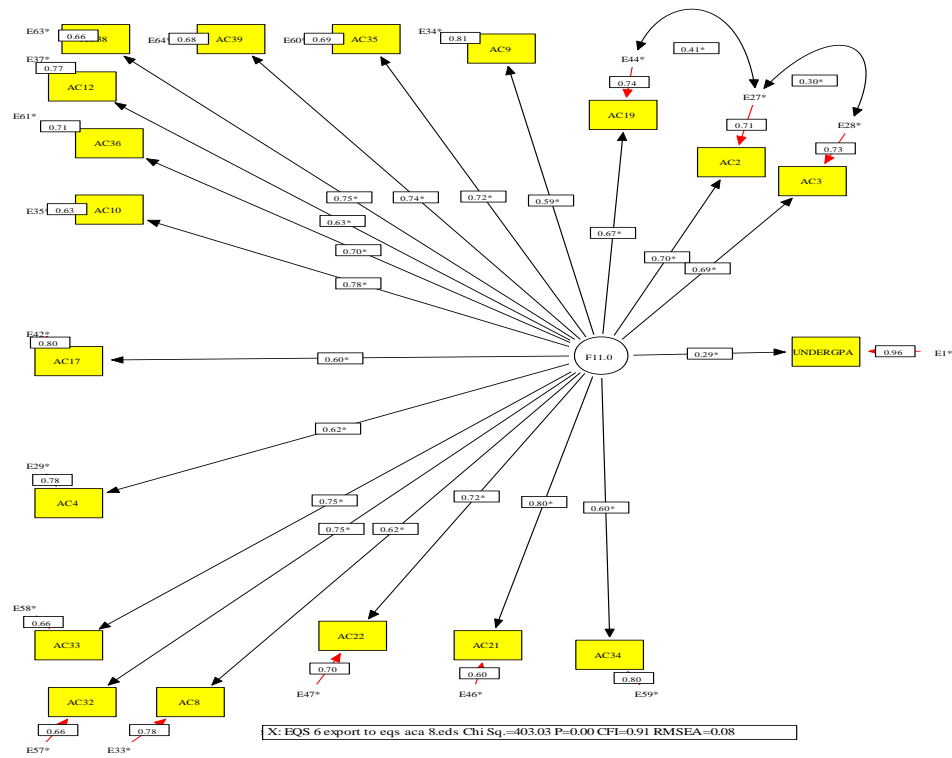
## 21. Mediation Analysis of Academic Conscientiousness: Part 2

Model	Predictor	Outcome	R	R-sq	F	P
<b>1</b>	PCon	PAcaCon	.7495	.5618	356.3664	.0000
<b>2</b>	PCon and PAcaCon	UnderGPA	.2751	.0757	11.3398	.0000
<b>3</b>	PCon	UnderGPA	.1626	.0264	7.5484	.0064

N = 280. PAcaCon= Positively Worded Academic Conscientiousness. PCon = Positively Worded General Conscientiousness. UnderGPA = Undergraduate GPA

In Table 20, we see our significance test (F) as well as our effect size (r) and percent of variance explained ( $r^2$ ). We can see in the difference in  $r^2$  between models 2 and 3 that the contextualized scale adds a lot to our prediction. We see in model 3 that our general conscientiousness scale explains 2.5% of the variance in Undergraduate GPA. Model 2 with both predictors in the equation, shows 7.5% of the variance in Undergraduate GPA explains. A full 5% absolute increase overall from 2.5 to 7.5 and a 200% increase in relative terms since 7.5 is three times the size of 2.5. These results further support the importance of context, despite nuisance factors and low quality data likely attenuating our correlations.

*3. Academic Conscientiousness, 1 factor solution with no negatively worded items, V26 and V45 removed, covariance (E44, E27) and (E28, E27)*



4. General Conscientiousness, 1 factor solution with no negatively worded items, 2 covariances (E11, 10) and (E15, 14)

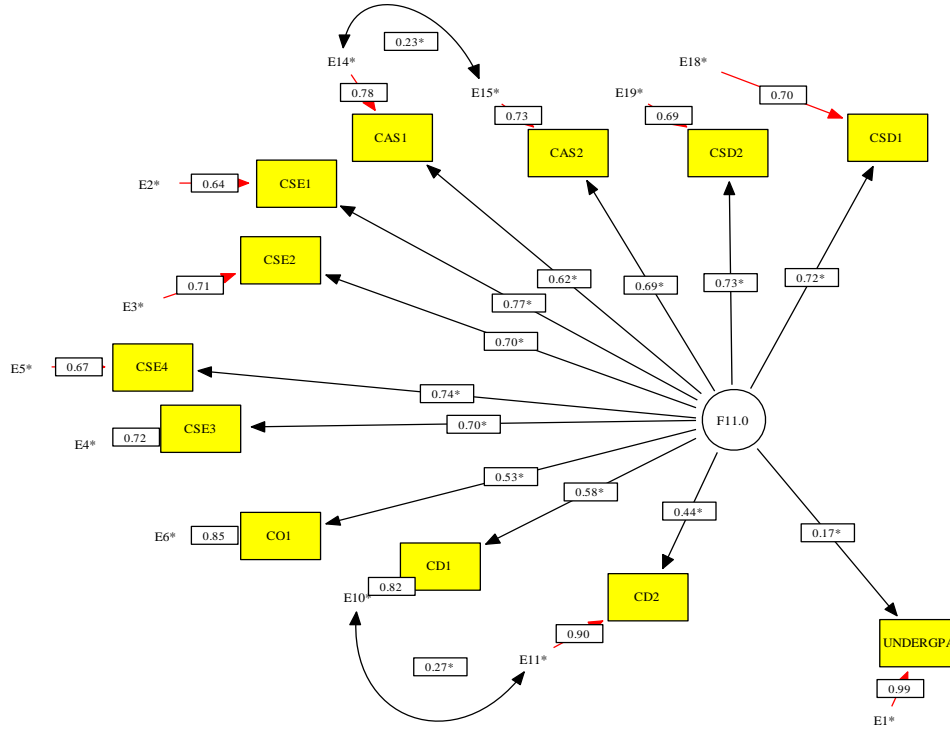


Figure X: EQS 6 export to eqs 5.ed5 Chi Sq.=94.56 P=0.00 CFI=0.97 RMSEA=0.05

## DISCUSSION

### Implications for Theory, Research, and Practice

#### Summary Findings

Our results suggest that our academic conscientiousness scale has a lot of potential to improve the prediction of outcome variables of interest. We found that the contextualized measure did have incremental validity over and above the general measure. We did technically find that specific facets matter, though the factors found were positively and negatively worded factors, instead of the factors hypothesized. Our findings, given the covariance between certain items, also shows that some of the factors

that we hypothesized would have likely been found if we had not had the nuisance factors.

Another significant outcome of our findings is that we showed that the factor structure for a contextualized scale is the same as that of generalized conscientiousness. Though this seemed to be the case because of the nuisance factor, some results like exploratory factor analysis and stepwise regression seemed to also support similar factor structures, even ignoring the negative factors. These findings offer partial support for the notion that situational personality forms distinct constructs but with potentially similar factor structures.

### **Factor Structure.**

An important implication of this study is understanding how personality constructs, which cannot be directly measured, are measured indirectly. Two issues arise when talking about this subject. One major problem that has arisen in this study and in prior studies is that personality is said to span the entire gamut from positive to negative. The continuum problem stems from the possibility that the positive end of a trait might form a separate factor from what is considered the opposite personality trait. Let's take orderliness for example: orderliness is said to span from extremely orderly to extremely messy. This makes a lot of intuitive sense and any scales that have only looked at the positive or only looked at the negative have been merged, both conceptually and methodically, to incorporate both ends.

However, the major problem ends up arising when we try to study personality and find that negatively worded items do not necessarily have the relationship with outcome

variables that we would expect from the constructs. In the case of this study, these negatively worded items, when reverse scored, do not even seem to have much in common with the positively worded items of the same construct. Some studies support the notion that this is caused by mental confusion and fatigue because participants have to mentally reverse word the items (van Sonderen et al., 2013). Other researchers argue that something like orderliness might not be on the same continuum as messiness. It might be that orderliness is on a continuum from high levels of orderliness to low levels of orderliness and that messiness is on a continuum from high to low levels. Essentially the debate is over whether orderliness and messiness form two continuums or one.

A second issue is the boundary conditions of a personality trait. We see here that personality when contextualized is a better predictor than a noncontextualized version of the same trait. But this begs the question, if a personality trait and a context are combined, is that not a violation of the assumption of unidimensionality? We also have talked about facets/narrow traits as well as broad traits but also included in domains in our results section that are under broad traits but subsume narrow traits. I believe, and recent research has supported, that personality should not only have broad and narrow traits, but also once broad and narrow traits are understood, compound traits should be created, which are statistically created traits made up of broad and narrow traits and likely further violates unidimensionality assumptions (Credé, Harms, Blacksmith, & Wood, 2016). Compound traits are the linear combination of narrow traits and potentially even some of the broader traits. These linear combinations are a way of piecing together relevant personality traits to the context and outcome variables involved.

As will be explained in the upcoming discussion on broad, narrow, and compound traits, this would allow the proper combination of personality to be studied in a variety of contexts. This would allow personality to be specific to the context, by combining traits, and in combination with contextualized measures in certain circumstances, allow for incremental validity over and above just contextualized measures, combinations of traits, or just general measures on their own.

Broad traits are traits such as the Big Five, which span a range of related constructs that are relatively unidimensional. Narrow traits, are parts of the Big Five that are the reason we preface unidimensional with the word “relatively”, since each of the Big Five have components that share things in common within that component that they do not share between components. Another word for narrow traits is “facets”. We see in this study, the attempt to study six potential facets of conscientiousness in achievement-striving, self-discipline, self-efficacy, dutifulness, orderliness, and cautiousness.

One of the reasons that contextual scales might increase prediction of outcome variables is by lowering social desirability bias. Contextual scales are considered by the researchers to be more cognitively demanding, which has been shown in prior studies to reduce social desirability bias (Stodel, 2015). Social desirability bias has been shown to potentially attenuate correlations between predictors and outcome variables (Ganster, Hennessey, & Luthans, 1982, 1983).

## Further Interpretation of Negative Factors

“Although a set of items may have been designed to measure the same construct, it is not uncommon for all the positively worded items to load on one factor and all the negatively worded items to load on another. Positively worded items present statements tapping strong expression of the construct directly and ask participants to rate the extent to which the construct applies to them. These are phrased in desirable terms, for example, in the SWBS: “I feel good about my future” is a positively worded item. Negatively worded items reflect the opposite ends of the construct and are phrased in undesirable terms, for example, in the SWBS: “I don’t enjoy much about life” is a negatively worded item. Thus, where a single substantive factor is hypothesized on theoretical grounds, two factors might arise in practice.” (Murray, Johnson, Gow, & Deary, 2015, p. 121)

Negatively worded items loading onto a factor made up solely of negatively worded items is considered by some to be a nuisance factor that is created through a wording effect, which is a type of method effect (DiStefano & Motl, 2009; Greenberger, Chen, Dmitrieva, & Farruggia, 2003; Huang & Dong, 2012; Tomás & Oliver, 1999). Method effects are “said to occur when any characteristic of a measurement procedure contributes variance to scores beyond what is attributable to variance in the attribute of interest” (Maul, 2013). Researchers who support this interpretation believe that because the items are similarly worded, that positively worded items will necessarily group together because they have the positive wording in common. This would also mean the



negatively worded items would group together because of their similar wording. This method effect is one of the justifications for using a mechanical solution to partial out variance associated with the negatively worded items. This assumes that there is no methodological solution aside from outright tossing out negatively worded items, and instead the solution is to get rid of this method effect after the data has been collected.

Even other studies say this is simply a matter of careless responding. Some research shows that if even 10% of respondents are careless, then a negatively worded factor will arise from all the negatively worded items (Schmitt & Stuits, 1985). People of this viewpoint believe that participants do not go through the effort of fully reading questions out and just believe that they are above or below average on a particular trait and answer the same after they believe they have figured out what trait is being measured.

Unfortunately, because of how many participants already had to be taken out of our data set because of problems with low quality data, any further attempt to fix the problem by removing cases were ignored. The researchers will go back before doing any follow up research and find ways to filter out this particular type of careless responding for upcoming data sets, to see if this will make a difference in whether a nuisance factor is found in factor analysis. Research, discussion, and going to presentations/workshops on careless responding are ongoing, with an ad hoc MTurk user group being created to troubleshoot problems. At least one presentation at the Society of Industrial and Organizational Psychologists (SIOP) will be attended by the lead researcher on this project. The lead researcher has also reviewed numerous articles which have been cited

throughout this paper on careless responding and related concepts like nuisance factors and method effects.

The results suggest that nuisance factors are more problematic than we had initially expected but that contextualized scales might be better at ameliorating even those problems, since we found that negatively worded academic conscientiousness was still a better predictor of college GPA than either positively worded conscientiousness, which would be conscientiousness without the nuisance factor, or conscientiousness as a whole with the nuisance factor.

There are two ways we have gone about fixing this problem. One is to just disregard negatively worded items entirely (van Sonderen et al., 2013). This has a problem of ignoring the theoretical implications of being unable to properly measure personality from positive to negative.

Another solution that seems more like a shortcut, is that we have mechanically partialled out the variance specific to the negatively worded items (Zhang et al., 2016). I disagree with both methods. Tossing out negatively worded items does not solve the problem unless we are willing to concede that personality does not go from positive to negative but is instead from positive to none and a separate construct exists from negative to none. The mechanical solution of partialling out the variance is unfortunately a solution people seem to believe solves the problem. It does not solve the problem because the implications are still that we cannot elicit knowledge from the negative end of any of our personality scales. Mechanical solutions are NOT true solutions to theoretical and methodological problems.

## Limitations

One problem with having general and academic questions on one questionnaire is that if participants get the academic questions first or even are told they are required to be a current or prior college student to fill out the survey, they are likely to be at least partially primed to think of an academic context even for the general questions. Another related problem is we simply do not have much of any research showing exactly what people are thinking of when they answer personality questions. What is the reference group they are comparing themselves to? Do things like identity salience and centrality play a part in which personality traits or contexts they use to answer a question? In what ways can we manipulate knowledge elicitation to increase the predictive validity of our measures?

Another limitation we ended up finding was that negatively worded items ended up forcing a certain factor structure to arise. While we are not sure why negatively worded items are so strongly related to each other without being strongly related to positively worded items, we can say that contextualized items handled the negatively worded attenuation to a better extent than the noncontextualized measure did. This has implications for ways to combat careless responding and social desirability.

Another thing I could have done and will do in upcoming surveys, would be to filter out people who put down a graduate school GPA without having attended graduate school, as a way to find people who are not paying attention. The reason this could not be done with this data set is that I did not put currently attending options for graduate school, so I cannot know if the degree is pending completion.

As a learning experience, I discovered that it is important to do some outlier analysis before reverse scoring items. There are a few issues with reverse scored items which the literature talks about and I discussed in some detail as well (Colosi, n.d.; van Sonderen, Sanderman, & Coyne, 2013; Zhang, Noor, & Savalei, 2016). Looking at descriptive statistics after reverse scoring can lead me to assuming the person answered around average for all items if using a scale balanced around equal numbers of positive and negative items because half the items have been reversed score. When in reality they just answered 1s or 5s for all items.

### **Future Research**

A number of future research directions immediately make themselves known based on the results, implications, and limitations of this study. This study in no way refutes the notion of facets being a potential benefit to incremental validity of personality measures. In fact, we saw that even with nuisance factors, the nuisance factors themselves could be thought of as facets and choosing one over the other was of benefit. We even found that despite the completely different factor structure than what was hypothesized, that technically looking at facet-level and contextual-level personality together did add more to the prediction than either alone or just general conscientiousness. This means that further research needs to be done into facet-level personality. As a short-term measure, coming up with more positively worded items in order to find the desired factor structure, regardless of if we come up with a solution to the negatively worded item problem. A more long-term solution and an avenue for future

research is to try to come up with a methodological solution to the negatively worded item problem.

If that proves to be untenable, then coming up with a theoretical solution is an option for future research. What I mean by this is that if there is no methodological way to show that personality exists on a continuum from positive to negative, then creating scales from positive to none and a separate scale from negative to none for personality traits would be necessary. A way to support this over the continuum from positive to negative would be to find differing relationships between the positive side of the scale and the negative side of the scale, this would add criterion validity support to the notion of them being different scales if they are able to predict outcomes differentially than if we had simply reversed the scores of the positive or negative scales.

As mentioned earlier in the discussion section, contextualized/situationally specific personality does seem to have incremental validity over and above generalized self-report measures. Further studies need to look at whether the incremental validity that contextualized/situationally specific measures have over generalized self-report measures is the same extra explained variance that observer ratings have over self-reports or not. As discussed earlier, there are reasons to suspect that observer ratings of general personality traits potentially explain the same extra variance as contextualized/situationally specific self-report measures of personality traits.

Another potential reason for a negative worded nuisance factor to arise is because of avoidance behavior. Avoidance behavior is the occurrence of withdrawal behavior that arises from a fear of aversive stimuli (DiStefano & Motl, 2009). This was not something

we looked at in our study, but it has potential for future research. We did find an odd positive bivariate correlation between age and how strongly people responded to negatively worded items. This is interpreted by the researchers to potentially mean that as a person gets older, they care less about the potential negative evaluations of others and therefor answer negatively worded questions more honestly. Age's relationship with negatively worded items was not studied any further than this but the researchers might come back to this after the main project is done to see if age has had any relationship with avoidance behaviors in other studies.

Whether situationally specific measures offer enough added value to organizations or others who might be interested in situationally specific personality is likely to at least partially depend on how many people they plan on measuring. The benefit of situationally specific measures is also dependent on the ease of potentially getting observer ratings, which also show incremental validity over self-reports of generalized measures. More research needs to be done to show whether observer ratings have incremental validity over and above contextualized/situationally specific measures of personality or if the incremental validity over and above general self-report measures is the same extra explained variance.

Other potential avenues for research include ideal point measures of personality and conditional reasoning tests. Ideal point item response theory is being potentially used since extreme ends of personality might be maladaptive (Carter, Miller, & Widiger, 2018). Conditional reasoning tests are being explored because of their potential to ameliorate faking. Some evidence shows that the reason faking occurred less frequently

in conditional reasoning tests is because people are unused to the format and do not realize what the socially desirable responses are. When participants are asked to fake and told what the test is measuring, they are frequently able to fake (LeBreton, Barksdale, Robin, & James, 2007). This could potentially mean that the newness of conditional reasoning tests makes them less susceptible to faking because people are not familiar with the test format. It is also possible that advances in conditional reasoning tests could make them more easily resistant to faking. However, we cannot know for sure which is the case since conditional reasoning tests are still new and unexplored, further research should be done on conditional reasoning tests as implicit measures of personality.

With another round of item generation, a new understanding about knowledge elicitation and how people respond to items, and a few more rounds of data collection, we believe that we can further improve prediction of our outcome variables of interest. Our scale in upcoming research needs to be tested with other academically related outcome variables like academic satisfaction, academic stress, academic adjustment, etc. We also need to test that it differentially predicts, which ideally would be done by creating some other contextual scale, such as a work conscientiousness scale, and showing that the work conscientiousness scale is a better predictor of work performance, work satisfaction, work stress, etc. In summary, we plan on making sure our contextualized scale for one context, best predicts that context, while the contextualized scale for another context, best predicts variables related to that context. We expect the general scale will be an average predictor across the board because of its creation with the emphasis on validity generalization.





## APPENDICES

### Example Questionnaire Items

3 Age:

---

4 Gender:

Male (1)

Female (2)

Other (3) \_\_\_\_\_

5a Do you consider yourself to be Hispanic/Latino?

Yes (23)

No (24)

5b Race/Ethnicity (Can choose multiple)

White (1)

Black or African American (2)

American Indian or Alaska Native (3)

Asian (4)

Native Hawaiian or Pacific Islander (5)

Other (6) \_\_\_\_\_

6 Average Annual Income of Caregivers During Childhood

- Less than \$10,000 (1)
- \$10,000 - \$19,999 (2)
- \$20,000 - \$29,999 (3)
- \$30,000 - \$39,999 (4)
- \$40,000 - \$49,999 (5)
- \$50,000 - \$59,999 (6)
- \$60,000 - \$69,999 (7)
- \$70,000 - \$79,999 (8)
- \$80,000 - \$89,999 (9)
- \$90,000 - \$99,999 (10)
- \$100,000 - \$149,999 (11)
- More than \$150,000 (12)

7a Please enter role of first caregiver (examples: Mother, Father, Aunt, Grandmother, etc)

**NOTE:** Remember choice, will be asked about a second primary caregiver.

- Mother (1)
- Father (2)
- Other (3) \_\_\_\_\_

7b First Primary Caregiver: Highest Education Attained

- Less than high school (1)
- High school graduate (2)
- Some college (3)
- 2 year degree (4)
- 4 year degree (5)
- Masters Degree (6)
- Professional degree (7)
- Doctorate (8)

Q21 Please enter role of second caregiver (examples: Mother, Father, Aunt, Grandmother, etc)

- Mother (1)
  - Father (2)
  - Other (3) \_\_\_\_\_
  - No Second Caregiver (4)
- 

Q29 Second Primary Caregiver: Highest Education Attained

- Less than high school (1)
- High school graduate (2)
- Some college (3)
- 2 year degree (4)
- 4 year degree (5)
- Masters Degree (6)
- Professional degree (7)
- Doctorate (8)

Q24 What was your average family size in household during Childhood? The number below includes you.

- 2 (1)
- 3 (2)
- 4 (3)
- 5 (4)
- 6 (5)
- 7 (6)
- 8 (7)
- 9 (8)
- 10 or more (9)

8 Your Highest Education Completed

- Less than high school (1)
- High school graduate (2)
- Some undergrad, did not finish and not currently attending (3)
- Some undergrad, currently attending (4)
- 2 year degree (5)
- 4 year degree (6)
- Masters Degree (7)
- Professional Degree (9)
- Doctorate (10)

9 Please write in your undergraduate GPA below:

NOTE: Leave blank if not applicable. Has numerical entry restriction so don't write anything in if not applicable.

---

10 Please write in your graduate GPA below:

**NOTE:** Leave blank if not applicable. Has numerical entry restriction so don't write anything in if not applicable.

---

## Generalized Conscientiousness Items

<b>Generalized Conscientiousness</b>	
<b>C1: SELF-EFFICACY (.77)</b>	
+ keyed	Complete tasks successfully.
	Excel in what I do.
	Handle tasks smoothly.
	Know how to get things done.
<b>C2: ORDERLINESS (.83)</b>	
+ keyed	Like to tidy up.
- keyed	Often forget to put things back in their proper place.
	Leave a mess in my room.
	Leave my belongings around.
<b>C3: DUTIFULNESS (.67)</b>	
+ keyed	Keep my promises.
	Tell the truth.
- keyed	Break rules.
	Break my promises.
<b>C4: ACHIEVEMENT-STRIVING (.79)</b>	
+ keyed	Do more than what's expected of me.
	Work hard.
- keyed	Put little time and effort into my work.
	Do just enough work to get by.
<b>C5: SELF-DISCIPLINE (.71)</b>	
+ keyed	Am always prepared.
	Carry out my plans.
- keyed	Waste my time.
	Have difficulty starting tasks.
<b>C6: CAUTIOUSNESS (.88)</b>	
- keyed	Jump into things without thinking.
	Make rash decisions.
	Rush into things.
	Act without thinking.

Alphas based on an Internet sample of  $N = 619,150$ .

(Johnson, 2014)



## Academic Conscientiousness Items

Below are the contextualized questions that will be used to find a factor structure. First we will tie them to outcome variables, then we will eliminate questions that do not have high loadings on the latent variable and the criterion of interest. If the contextualized scale ends up having a distinctly dissimilar factor structure from non-contextualized conscientiousness, we will discuss the implications of situations altering factor structures.

<b>Academic Orderliness</b>	
1.	I like to have a structured approach for my school assignments. (O) or (D) or (SD)
2.	I follow a schedule to work on school assignment. (D) or (O)
3.	I follow a schedule at school. (D or O)
4.	I take organized notes in class.

<b>Academic Dutifulness</b>	
5.	I cheat on tests. (D)
6.	I break rules at school. (D) PC
7.	I cheat on homework assignments
8.	I finish the school assignments by their deadline. (D) or (O) or (SD)

<b>Academic Achievement-Striving</b>	
9.	I do more than what is expected of me at school. (AS) <b>P</b>
10.	I work hard at school. (AS) PC <b>P</b>
11.	I put little time and effort into my schoolwork. (AS) N
12.	I aim to get a perfect grade on assignments. (AS) <b>P</b>
13.	I do just enough work to get my school assignments done. (AS) <b>P</b>

<b>Academic Self-Discipline</b>	
---------------------------------	--

14.	I delay finishing my school assignments.			N
15.	I find it difficult to get started on schoolwork.	(SD)		N
16.	I find it difficult to get rid of distractions and get my school assignments done.	(SD)		N
17.	I get school assignments done quickly.	(SD)	<b>P</b>	
18.	I get distracted when doing assignments.	(SD)		N
19.	I follow a schedule to work on assignments.	(SD)	<b>P</b>	
20.	I make myself study during specific hours.	(SD)		<b>P</b>
21.	I allocate my time for school work.	(SD)	<b>P</b>	
22.	I can focus enough to get my school assignments done on time.	(SD)		<b>P</b>
23.	I postpone studying for an exam.	(SD)		N
24.	I find it difficult to complete my assignments on time.	(SD)		N
25.	I waste my time when I am supposed to work on academic projects.	(SD)		N
26.	I am easily distracted when studying.	(SD)		N
27.	I am easily distracted when doing assignments.	(SD)		N

#### Academic Cautiousness

28.	I jump into things without thinking at school.	(C)	PC		N
29.	I rush into things at school.	(C)	PC		N
30.	I act without thinking at school.	(C)	PC		N
31.	I ask questions in class without first thinking it through.	(C)			N

#### Academic Self-Efficacy

32.	I complete my school assignments successfully.	(SE)			<b>P</b>
33.	I excel in what I do at school.	(SE)	PC		<b>P</b>
34.	I feel it's easy to keep up with all the school assignments I have.	(SE)			<b>P</b>
35.	I know how to get things done at school.	(SE)	PC		<b>P</b>
36.	I can focus enough to get my school assignments done on time.	(SE)			<b>P</b>
37.	I perform below expectation on tests or assignments at school.	(SE)			N
38.	I complete my tasks at school successfully.	(SE)		PC	<b>P</b>
39.	I am prepared to take on any school assignment.	(SE)			<b>P</b>

#### Abbreviation Key

PC = Partially Contextualized	SE = Self-Efficacy facet
SD = Self-Discipline facet	AS = Achievement-Striving facet
D = Dutifulness facet	O = Orderliness
C = Cautiousness	P = Positively Worded

N = Negatively Worded
-----------------------

## Likert Scale Points

### Indicate how much the following statements describe you:

1 = Not Accurate At All	2 = Slightly Accurate	3 = Moderately Accurate
4 = Very Accurate		5 = Extremely Accurate

The 5-point Likert scale was used based on research showing people struggle to make distinctions between Likert scale points and it is thought that having fewer point on the scale might make it easier for people to make meaningful decisions instead of carelessly deciding between two points (Jamieson, 2004).

The decision to use an odd or even number Likert scale was thought to be up for debate. The research is undecided on the issue. To compound the problem of making a decision, our “mid-point” is not entirely a neutral option, which is how most of the research has viewed mid-points (Garland, 1991; Moors, 2008).

## Tables

### Factor Analysis General Conscientiousness: One Factor

*Summary of Exploratory Factor Analysis Results for a General Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings	Communalities
	Conscientiousness	Extraction
I complete tasks successfully.	.518	.268
I excel in what I do.	.358	.128

I handle tasks smoothly.	.479	.229
I know how to get things done.	.490	.240
I like to tidy up.	.395	.156
I often forget to put things back in their proper place. RK	.688	.473
I leave a mess in my room. RK	.750	.562
I leave my belongings around. RK	.719	.517
I keep my promises.	.484	.234
I tell the truth.	.429	.184
I break the rules. RK	.647	.419
I break my promises. RK	.687	.473
I do more than what's expected of me.	.377	.142
I work hard.	.519	.269
I put little time and effort into my work. RK	.668	.446
I do just enough work to get by. RK	.674	.454
I am always prepared.	.456	.208
I carry out my plans.	.483	.234
I waste my time. RK	.788	.621
I have difficulty starting tasks. RK	.756	.571
I jump into things without thinking. RK	.739	.546
I make rash decisions. RK	.739	.547
I rush into things. RK	.732	.535
I act without thinking. RK	.733	.537
Eigenvalues	9.553	
% of variance	39.8	

*Note:* Factor Loadings below .30 are hidden. Factor Matrix factor loadings were used because one factor solution only provides a factor matrix.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

## Factor Analysis General Conscientiousness: Three Factors

*Summary of Exploratory Factor Analysis Results for a General Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings			Communalities Extraction
	Negative Industriousness	Positive Conscientiousness	Negative Orderliness	
I complete tasks successfully.		.776		.601
I excel in what I do.		.732		.495
I handle tasks smoothly.		.676		.495
I know how to get things done.		.711		.524
I like to tidy up.		.411	-.501	.463
I often forget to put things back in their proper place. RK	.415		-.531	.628
I leave a mess in my room. RK			-.774	.870
I leave my belongings around. RK			-.758	.815
I keep my promises.		.534		.354
I tell the truth.		.396		.235
I break the rules. RK	.648			.524
I break my promises. RK	.791			.633
I do more than what's expected of me.		.702		.455
I work hard.		.692		.515
I put little time and effort into my work. RK	.741			.580
I do just enough work to get by. RK	.629			.490
I am always prepared.		.677		.487
I carry out my plans.		.733		.551
I waste my time. RK	.555			.611
I have difficulty starting tasks. RK	.502			.578
I jump into things without thinking. RK	.867			.735
I make rash decisions. RK	.806			.692
I rush into things. RK	.881			.740
I act without thinking. RK	.833			.688
Eigenvalues	9.553	4.037	1.315	
% of variance	39.8%	16.8%	5.5%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor correlation between factor 1 and 3 was -.467. Factor correlation between 1 and 2 was .276. Factor correlation between 2 and 3 was -.370.

**Factor Correlation Matrix**

Factor	1	2	3
1	1.000	.276	-.467
2	.276	1.000	-.370
3	-.467	-.370	1.000

We see here the first evidence of a bit of separation between an orderliness and industriousness facets within the negatively worded conscientiousness items.

## Factor Analysis General Conscientiousness: Four Factors

*Summary of Exploratory Factor Analysis Results for a General Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings				Communalities Extraction
	Negative Industriousness	Positive Industriousness	Negative Orderliness	Positive Orderliness	
I complete tasks successfully.		.774			.603
I excel in what I do.		.765			.529
I handle tasks smoothly.		.676			.500
I know how to get things done.		.685			.520
I like to tidy up.		.332	-.552		.487
I often forget to put things back in their proper place. RK	.318		-.591		.625
I leave a mess in my room. RK			-.846		.866
I leave my belongings around. RK			-.834		.817
I keep my promises.		.416		.414	.510
I tell the truth.				.371	.364
I break the rules. RK	.633				.554
I break my promises. RK	.778				.646
I do more than what's expected of me.		.714			.465
I work hard.		.716			.534
I put little time and effort into my work. RK	.657				.636
I do just enough work to get by. RK	.540	.312		-.471	.703
I am always prepared.		.592			.527
I carry out my plans.		.683			.553
I waste my time. RK	.456	.320			.675

I have difficulty starting tasks. RK	.388		-.340		.640
I jump into things without thinking. RK	.844				.737
I make rash decisions. RK	.793				.713
I rush into things. RK	.864				.743
I act without thinking. RK	.835				.706
Eigenvalues	9.553	4.037	1.315	1.266	
% of variance	39.8%	16.8%	5.5%	5.3%	

Note: Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor correlation between factor 1 and 2 was .247. Factor correlation between 1 and 3 was -.523. Factor correlation between 1 and 4 was -.088. Factor correlation between 2 and 3 was -.393. Factor Correlations between 2 and 4 was .166. Factor correlations between 3 and 4 was -.019.

#### Factor Correlation Matrix

Factor	1	2	3	4
1	1.000	.247	-.523	-.088
2	.247	1.000	-.393	.166
3	-.523	-.393	1.000	-.019
4	-.088	.166	-.019	1.000



## Factor Analysis General Conscientiousness: Five Factors

*Summary of Exploratory Factor Analysis Results for a General Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings					Communalities Extraction
	Negative Cautiousness	Positive Industriousness	Negative Orderliness	Positive Orderliness	Negative Industriousness	
I complete tasks successfully.		.708				.604
I excel in what I do.		.840				.592
I handle tasks smoothly.		.648				.508
I know how to get things done.		.684				.540
I like to tidy up.			-.553			.484
I often forget to put things back in their proper place. RK			-.613			.627
I leave a mess in my room. RK			-.867			.864
I leave my belongings around. RK			-.860			.825
I keep my promises.				.614		.556
I tell the truth.				.633		.447
I break the rules. RK	.368					.568
I break my promises. RK	.409			.362	.326	.696
I do more than what's expected of me.		.576				.467

I work hard.		.526			.307	.554
I put little time and effort into my work. RK	.400				.444	.632
I do just enough work to get by. RK					.699	.728
I am always prepared.		.575				.539
I carry out my plans.		.536				.555
I waste my time. RK					.567	.718
I have difficulty starting tasks. RK			-.367		.442	.645
I jump into things without thinking. RK	.842					.792
I make rash decisions. RK	.726					.729
I rush into things. RK	.783					.762
I act without thinking. RK	.828					.751
Eigenvalues	9.553	4.037	1.315	1.266	.821	
% of variance	39.8%	16.8%	5.5%	5.3%	3.4%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor correlation between factor 1 and 2 was .033. Factor correlation between 1 and 3 was -.493. Factor correlation between 1 and 4 was .294. Factor correlation between 1 and 5 was .549. Factor correlation between 2 and 3 was -.322. Factor correlation between 2 and 4 was .484. Factor correlation between 2 and 5 was .178. Factor correlation between

3 and 4 was  $-.378$ . Factor correlation between 3 and 5 was  $-.379$ . Factor correlation between 4 and 5 was  $.169$ .

**Factor Correlation Matrix**

Factor	1	2	3	4	5
1	1.000	.033	-.493	.294	.549
2	.033	1.000	-.322	.484	.178
3	-.493	-.322	1.000	-.378	-.379
4	.294	.484	-.378	1.000	.169
5	.549	.178	-.379	.169	1.000

## Factor Analysis General Conscientiousness: Six Factors

*Summary of Exploratory Factor Analysis Results for a General Conscientiousness 24 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings						Communalities Extraction
	Negative Cautiousness	Positive Industriousness	Negative Orderliness	Positive Orderliness	Negative Industriousness		
I complete tasks successfully.		.719					.624
I excel in what I do.		.780					.591
I handle tasks smoothly.		.572					.506
I know how to get things done.		.772					.586
I like to tidy up.			-.565				.491
I often forget to put things back in their proper place. RK			-.606				.646
I leave a mess in my room. RK			-.823				.866
I leave my belongings around. RK			-.829				.823
I keep my promises.				.616			.551
I tell the truth.				.664			.459
I break the rules. RK	.315						.569
I break my promises. RK				.347	.427	-.377	.775
I do more than what's expected of me.						.398	.530
I work hard.		.301			.315		.579
I put little time and effort into	.388				.457		.629

my work. RK							
I do just enough work to get by. RK					.724		.762
I am always prepared.		.540					.537
I carry out my plans.		.437		.307			.553
I waste my time. RK					.632		.728
I have difficulty starting tasks. RK			-.318		.499		.658
I jump into things without thinking. RK	.854						.793
I make rash decisions. RK	.731						.729
I rush into things. RK	.827						.776
I act without thinking. RK	.850						.754
Eigenvalue s	9.553	4.037	1.315	1.266	.821	.656	
% of variance	39.8%	16.8%	5.5%	5.3%	3.4%	2.7 %	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor correlation between factor 1 and 2 was .111. Factor correlation between 1 and 3 was -.479. Factor correlation between 1 and 4 was .304. Factor correlation between 1 and 5 was .620. Factor correlation between 1 and 6 was -.282. Factor correlation between 2 and 3 was -.353. Factor correlations between 2 and 4 was .525. Factor correlation between 2 and 5 was .230. Factor correlation between 2 and 6 was .445. Factor correlation between 3 and 4 was -.363. Factor correlation between 3 and 5 was -

.412. Factor correlation between 3 and 6 was .028. Factor correlation between 4 and 5 was .186. Factor correlation between 4 and 6 was .121. Factor correlation between 5 and 6 was -.035.

**Factor Correlation Matrix**

Factor	1	2	3	4	5	6
1	1.000	.111	-.479	.304	.620	-.282
2	.111	1.000	-.353	.525	.230	.445
3	-.479	-.353	1.000	-.363	-.412	.028
4	.304	.525	-.363	1.000	.186	.121
5	.620	.230	-.412	.186	1.000	-.035
6	-.282	.445	.028	.121	-.035	1.000

## Factor Analysis Academic Conscientiousness: One Factor

*Summary of Exploratory Factor Analysis Results for an Academic Conscientiousness 39 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings	Communalities
	Academic Conscientiousness	Extraction
I like to have a structured approach for my school assignments.	.521	.272
I follow a schedule to work on school assignment.	.545	.297
I follow a schedule at school.	.564	.318
I take organized notes in class.	.538	.289
I cheat on tests. RK	.508	.258
I break rules at school. RK	.524	.275
I cheat on homework assignments. RK	.518	.268
I finish the school assignments by their deadline.	.569	.323
I do more than what is expected of me at school.	.394	.156
I work hard at school.	.638	.407
I put little time and effort into my schoolwork. RK	.688	.474
I aim to get a perfect grade on assignments.	.499	.249
I do just enough work to get my school assignments done. RK	.605	.366
I delay finishing my school assignments. RK	.785	.616
I find it difficult to get started on schoolwork. RK	.785	.616
I find it difficult to get rid of distractions and get my school assignments done. RK	.690	.477
I get school assignments done quickly.	.458	.210
I get distracted when doing assignments. RK	.692	.479
I follow a schedule to work on assignments.	.603	.364
I make myself study during specific hours.	.405	.164
I allocate my time for school work.	.627	.393
I can focus enough to get my school assignments done on time.	.644	.415
I postpone studying for an exam. RK	.707	.500
I find it difficult to complete my assignments on time. RK	.741	.550
I waste my time when I am supposed to work on academic projects. RK	.752	.565
I am easily distracted when studying. RK	.747	.558
I am easily distracted when doing assignments. RK	.698	.487
I jump into things without thinking at school. RK	.608	.370
I rush into things at school. RK	.600	.359
I act without thinking at school. RK	.601	.361
I ask questions in class without first thinking it through. RK	.449	.201
I complete my school assignments successfully.	.672	.451
I excel in what I do at school.	.557	.311
I feel it's easy to keep up with all the school assignments I have.	.413	.170
I know how to get things done at school.	.606	.367
I can focus enough to get my school assignments done on time.	.621	.385
I perform below expectation on tests or assignments at school. RK	.531	.282
I complete my tasks at school successfully.	.595	.354
I am prepared to take on any school assignment.	.550	.302
Eigenvalues	14.853	

% of variance	38%	
---------------	-----	--

*Note:* Factor Loadings below .30 are hidden. Factor Matrix factor loadings were used since one factor extraction only provides a Factor Matrix.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor 1 had a correlation with factor 2 of .283



## Factor Analysis Academic Conscientiousness: Three Factors

*Summary of Exploratory Factor Analysis Results for an Academic Conscientiousness 39 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings		Communalities	
	Academic Conscientiousness			Extraction
I like to have a structured approach for my school assignments.			.573	.475
I follow a schedule to work on school assignment.			.617	.601
I follow a schedule at school.			.678	.513
I take organized notes in class.			.636	.456
I cheat on tests. RK		-.756		.671
I break rules at school. RK		-.637		.599
I cheat on homework assignments. RK		-.680		.624
I finish the school assignments by their deadline.			.607	.407
I do more than what is expected of me at school.			.660	.467
I work hard at school.			.762	.614
I put little time and effort into my schoolwork. RK	.371	-.527		.650
I aim to get a perfect grade on assignments.			.722	.478
I do just enough work to get my school assignments done. RK	.397	-.312		.429
I delay finishing my school assignments. RK	.723			.735
I find it difficult to get started on schoolwork. RK	.758			.742
I find it difficult to get rid of distractions and get my school assignments done. RK	.733			.641
I get school assignments done quickly.			.555	.435
I get distracted when doing assignments. RK	.788			.665
I follow a schedule to work on assignments.			.608	.576
I make myself study during specific hours.	.308	.439	.478	.521
I allocate my time for school work.			.739	.621
I can focus enough to get my school assignments done on time.			.635	.501
I postpone studying for an exam. RK	.728			.634
I find it difficult to complete my assignments on time. RK	.486	-.397		.656
I waste my time when I am supposed to work on academic projects. RK	.755			.718
I am easily distracted when studying. RK	.799			.728
I am easily distracted when doing assignments. RK	.815			.702
I jump into things without thinking at school. RK	.327	-.659		.729
I rush into things at school. RK		-.671		.712
I act without thinking at school. RK		-.681		.717
I ask questions in class without first thinking it through. RK		-.658		.538
I complete my school assignments successfully.		-.302	.825	.661
I excel in what I do at school.			.785	.560
I feel it's easy to keep up with all the school assignments I have.			.580	.370
I know how to get things done at school.			.786	.576

I can focus enough to get my school assignments done on time.			.631	.475
I perform below expectation on tests or assignments at school. RK			-.521	.458
I complete my tasks at school successfully.			.859	.638
I am prepared to take on any school assignment.			.737	.529
<hr/>				
Eigenvalues	14.853	7.286	1.856	
% of variance	38%	18.7%	4.8%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

**Factor Correlation Matrix**

Factor	1	2	3
1	1.000	-.434	.443
2	-.434	1.000	-.022
3	.443	-.022	1.000

## Factor Analysis Academic Conscientiousness: Four Factors

*Summary of Exploratory Factor Analysis Results for an Academic Conscientiousness 39 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings				Communalities Extraction
	Academic Conscientiousness				
I like to have a structured approach for my school assignments.				.584	.538
I follow a schedule to work on school assignment.				.768	.737
I follow a schedule at school.				.579	.594
I take organized notes in class.			.454		.455
I cheat on tests. RK		-.843			.701
I break rules at school. RK		-.781			.652
I cheat on homework assignments. RK		-.792			.662
I finish the school assignments by their deadline.			.632		.429
I do more than what is expected of me at school.			.544		.474
I work hard at school.			.624		.614
I put little time and effort into my schoolwork. RK	.365	-.468			.652
I aim to get a perfect grade on assignments.			.547		.478
I do just enough work to get my school assignments done. RK	.399				.432
I delay finishing my school assignments. RK	.642				.736
I find it difficult to get started on schoolwork. RK	.729				.740
I find it difficult to get rid of distractions and get my school assignments done. RK	.825				.690
I get school assignments done quickly.			.362		.433
I get distracted when doing assignments. RK	.839				.690
I follow a schedule to work on assignments.				.693	.683
I make myself study during specific hours.				.626	.574
I allocate my time for school work.			.468	.398	.630
I can focus enough to get my school assignments done on time.			.570		.511
I postpone studying for an exam. RK	.644				.636
I find it difficult to complete my assignments on time. RK	.545				.685
I waste my time when I am supposed to work on academic projects. RK	.732				.717
I am easily distracted when studying. RK	.798				.732
I am easily distracted when doing assignments. RK	.889				.743
I jump into things without thinking at school. RK		-.729			.745
I rush into things at school. RK		-.745			.730

I act without thinking at school. RK		-.767			.740
I ask questions in class without first thinking it through. RK		-.615			.534
I complete my school assignments successfully.			.763		.663
I excel in what I do at school.			.761		.588
I feel it's easy to keep up with all the school assignments I have.			.577		.414
I know how to get things done at school.			.718		.583
I can focus enough to get my school assignments done on time.			.625		.502
I perform below expectation on tests or assignments at school. RK		-.418			.470
I complete my tasks at school successfully.			.898		.686
I am prepared to take on any school assignment.			.599		.529
Eigenvalues	14.853	7.286	1.856	1.275	
% of variance	38%	18.7%	4.8%	3.3%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor	1	2	3	4
1	1.000	-.590	.404	.258
2	-.590	1.000	-.146	.080
3	.404	-.146	1.000	.600
4	.258	.080	.600	1.000

## Factor Analysis Academic Conscientiousness: Five Factors

*Summary of Exploratory Factor Analysis Results for an Academic Conscientiousness 39 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings					Communalities Extraction
	Academic Conscientiousness					
I like to have a structured approach for my school assignments.				.653		.551
I follow a schedule to work on school assignment.				.823		.744
I follow a schedule at school.				.641		.598
I take organized notes in class.					.393	.462
I cheat on tests. RK		-.862				.710
I break rules at school. RK		-.779				.651
I cheat on homework assignments. RK		-.823				.683
I finish the school assignments by their deadline.			.640			.540
I do more than what is expected of me at school.					.684	.560
I work hard at school.					.463	.616
I put little time and effort into my schoolwork. RK	.364		-.484			.659
I aim to get a perfect grade on assignments.					.558	.514
I do just enough work to get my school assignments done. RK	.402					.455
I delay finishing my school assignments. RK	.628					.736
I find it difficult to get started on schoolwork. RK	.722					.749
I find it difficult to get rid of distractions and get my school assignments done. RK	.817					.696
I get school assignments done quickly.				.318		.433
I get distracted when doing assignments. RK	.829					.689
I follow a schedule to work on assignments.				.734		.683
I make myself study during specific hours.				.630		.572
I allocate my time for school work.				.485		.642
I can focus enough to get my school assignments done on time.			.542			.590
I postpone studying for an exam. RK	.628					.635
I find it difficult to complete my assignments on time. RK	.539					.702
I waste my time when I am supposed to work on academic projects. RK	.718					.718
I am easily distracted when studying. RK	.790					.737

I am easily distracted when doing assignments. RK	.879					.741
I jump into things without thinking at school. RK		-.725				.744
I rush into things at school. RK		-.730				.731
I act without thinking at school. RK		-.754				.740
I ask questions in class without first thinking it through. RK		-.584				.554
I complete my school assignments successfully.			.436		.357	.669
I excel in what I do at school.					.794	.674
I feel it's easy to keep up with all the school assignments I have.					.360	.412
I know how to get things done at school.					.517	.585
I can focus enough to get my school assignments done on time.			.529			.559
I perform below expectation on tests or assignments at school. RK		-.439				.480
I complete my tasks at school successfully.			.421		.530	.679
I am prepared to take on any school assignment.					.533	.547
Eigenvalues	14.853	7.286	1.856	1.275	1.012	
% of variance	38%	18.7%	4.8%	3.3%	2.6%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

#### Factor Correlation Matrix

Factor	1	2	3	4	5
1	1.000	-.602	.358	.290	.280
2	-.602	1.000	-.245	.033	-.040
3	.358	-.245	1.000	.412	.545
4	.290	.033	.412	1.000	.618
5	.280	-.040	.545	.618	1.000

## Factor Analysis Academic Conscientiousness: Six Factors

*Summary of Exploratory Factor Analysis Results for an Academic Conscientiousness 39 Item Measure Using Principal Axis Factoring Extraction and Oblimin Rotation (N = 358)*

Item	Factor Loadings					Communalities Extraction
	Academic Conscientiousness					
I like to have a structured approach for my school assignments.				.656		.553
I follow a schedule to work on school assignment.				.824		.746
I follow a schedule at school.				.639		.598
I take organized notes in class.				.309	.378	.479
I cheat on tests. RK		-.852				.713
I break rules at school. RK		-.756				.650
I cheat on homework assignments. RK		-.858				.710
I finish the school assignments by their deadline.			.664			.551
I do more than what is expected of me at school.					.664	.558
I work hard at school.					.453	.694
I put little time and effort into my schoolwork. RK	.348		-.370			.724
I aim to get a perfect grade on assignments.					.542	.519
I do just enough work to get my school assignments done. RK	.390					.542
I delay finishing my school assignments. RK	.634					.747
I find it difficult to get started on schoolwork. RK	.722					.751
I find it difficult to get rid of distractions and get my school assignments done. RK	.818					.697
I get school assignments done quickly.						.491
I get distracted when doing assignments. RK	.829					.690
I follow a schedule to work on assignments.				.713		.688
I make myself study during specific hours.				.647		.581
I allocate my time for school work.				.490		.644
I can focus enough to get my school assignments done on time.			.538			.598
I postpone studying for an exam. RK	.624					.635

I find it difficult to complete my assignments on time. RK	.542						.703
I waste my time when I am supposed to work on academic projects. RK	.714						.721
I am easily distracted when studying. RK	.789						.737
I am easily distracted when doing assignments. RK	.874						.743
I jump into things without thinking at school. RK			-707				.744
I rush into things at school. RK			-717				.732
I act without thinking at school. RK			-734				.740
I ask questions in class without first thinking it through. RK			-531				.560
I complete my school assignments successfully.			.448		.337		.667
I excel in what I do at school.					.765		.669
I feel it's easy to keep up with all the school assignments I have.					.356		.460
I know how to get things done at school.					.502		.599
I can focus enough to get my school assignments done on time.			.531				.558
I perform below expectation on tests or assignments at school. RK			-.445				.485
I complete my tasks at school successfully.			.442		.507		.680
I am prepared to take on any school assignment.					.521		.560
Eigenvalues	14.853	7.286	1.856	1.275	1.012	.854	
% of variance	38%	18.7%	4.8%	3.3%	2.6%	2.2%	

*Note:* Factor Loadings below .30 are hidden. Pattern Matrix factor loadings were used.

RK = Reverse keyed items, meaning they are on the negative end of the trait continuum.

Factor	1	2	3	4	5	6
1	1.000	-.600	.371	.305	.275	-.156
2	-.600	1.000	-.268	.006	-.051	.293
3	.371	-.268	1.000	.418	.539	.046
4	.305	.006	.418	1.000	.605	.175
5	.275	-.051	.539	.605	1.000	.142
6	-.156	.293	.046	.175	.142	1.000



## Figures

Figure 1.

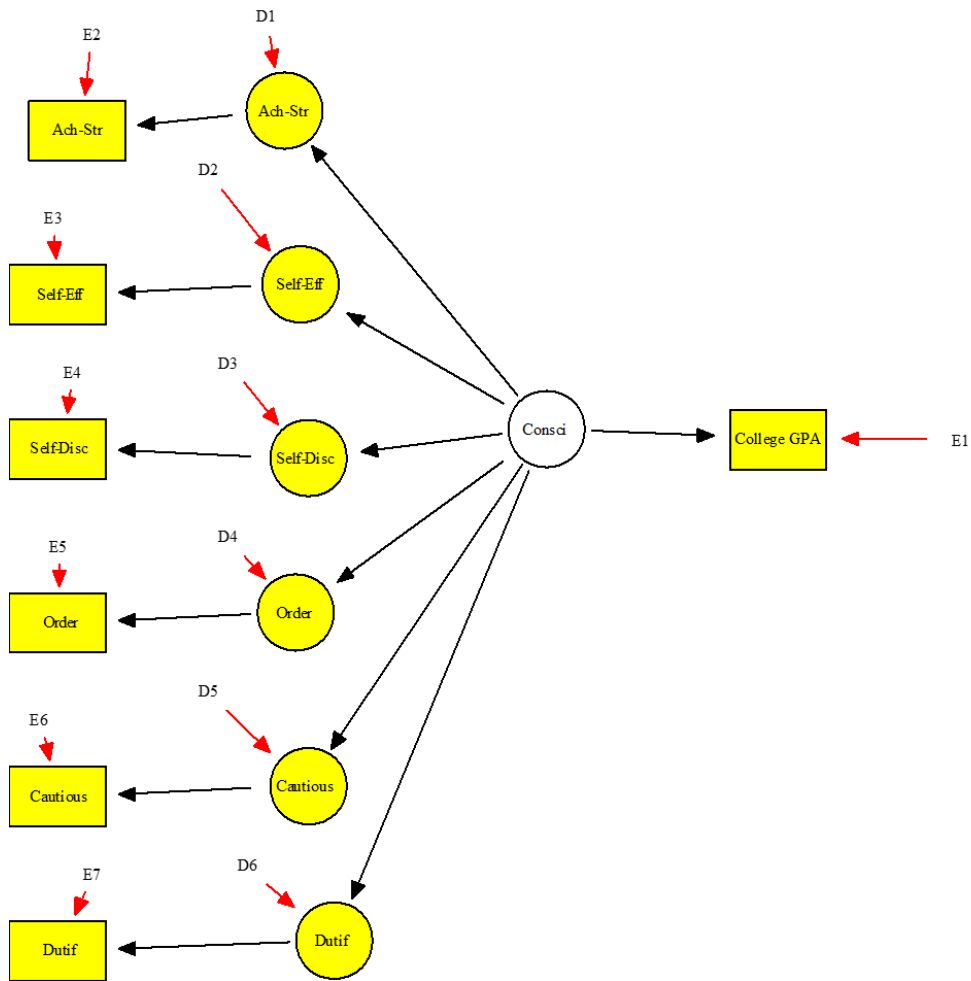
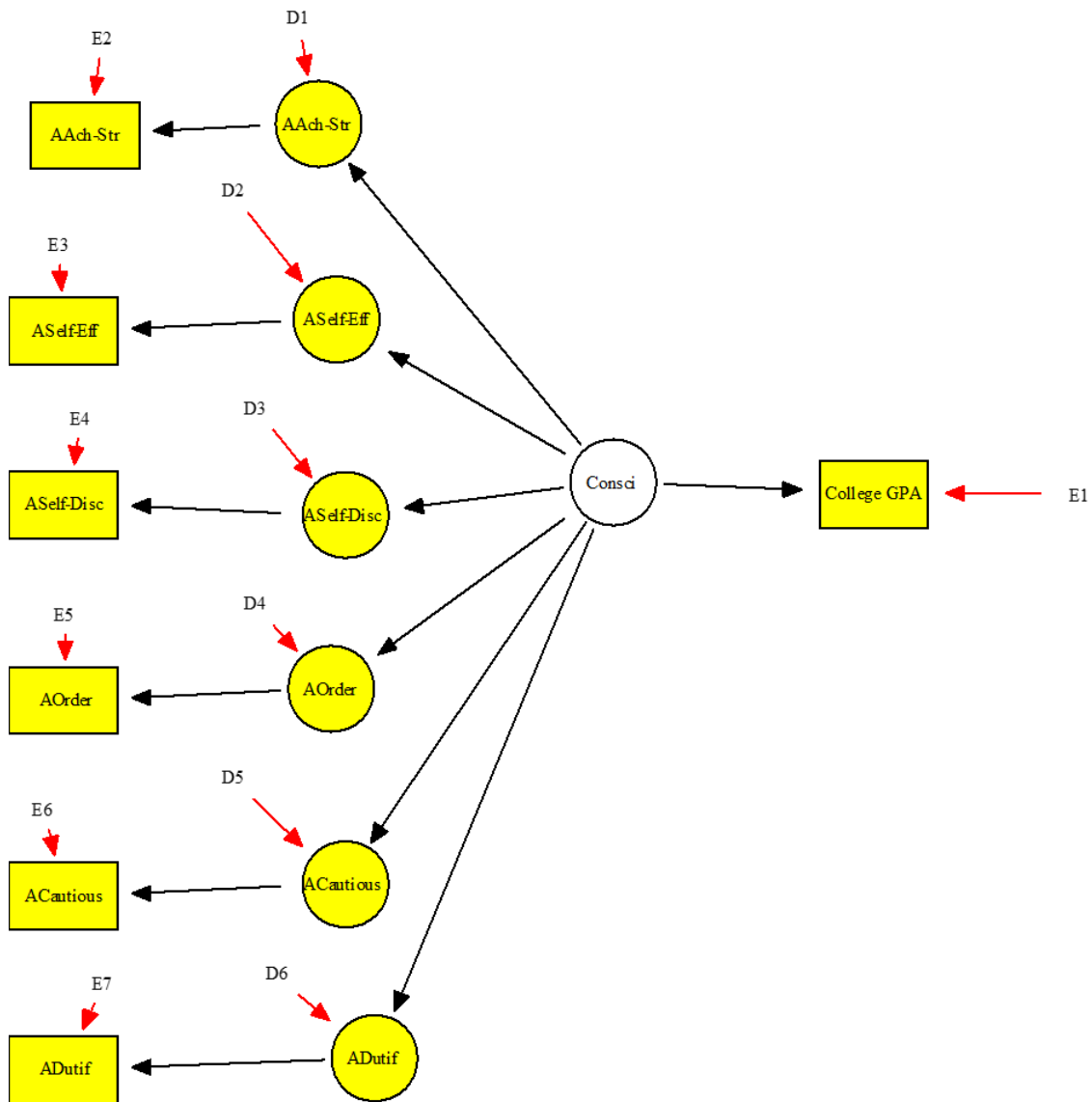


Figure 1. Non-Contextualized Conscientiousness as a second order factor structure. Each item representative of multiple items that will actually be in the diagram corresponding to each facet. Its Non-Contextualized Conscientiousness-College GPA path has been constrained to be the same as Academic Conscientiousness-College GPA.

**Figure 2**



*Figure 2.* Academic Conscientiousness as a second order factor structure. Each item representative of multiple items that will actually be in the diagram corresponding to each facet. Its Academic Conscientiousness-College GPA path has been constrained to be the same as Non-Contextualized Conscientiousness-College GPA.

## REFERENCES

- America's College Students Infographic | Postsecondary Success. (2014, December 12). Retrieved October 4, 2018, from <https://postsecondary.gatesfoundation.org/areas-of-focus/incentives/policy-advocacy/advocacy-priorities/america-100-college-students/>
- Baird, B. M., & Lucas, R. E. (2011). "... And How About Now?": Effects of Item Redundancy on Contextualized Self-Reports of Personality: Effects of Item Redundancy. *Journal of Personality, 79*(5), 1081–1112. <https://doi.org/10.1111/j.1467-6494.2011.00716.x>
- Bowling, N. A., & Burns, G. N. (2010). A comparison of work-specific and general personality measures as predictors of work and non-work criteria. *Personality and Individual Differences, 49*(2), 95–101. <https://doi.org/10.1016/j.paid.2010.03.009>
- Buss, D. M. (2009). An evolutionary formulation of person–situation interactions. *Journal of Research in Personality, 43*(2), 241–242. <https://doi.org/10.1016/j.jrp.2008.12.019>
- Cappelleri, J. C., Lundy, J. J., & Hays, R. D. (2014). Overview of Classical Test Theory and Item Response Theory for Quantitative Assessment of Items in Developing Patient-Reported Outcome Measures. *Clinical Therapeutics, 36*(5), 648–662. <https://doi.org/10.1016/j.clinthera.2014.04.006>
- Caprara, G. V., Alessandri, G., Barbaranelli, C., & Vecchione, M. (2013). The longitudinal relations between self-esteem and affective self-regulatory efficacy. *Journal of Research in Personality, 47*(6), 859–870. <https://doi.org/10.1016/j.jrp.2013.08.011>

- Carter, N. T., Miller, J. D., & Widiger, T. A. (2018). Extreme Personalities at Work and in Life. *Current Directions in Psychological Science*, 27(6), 429–436.  
<https://doi.org/10.1177/0963721418793134>
- Christensen, C. M. (2001). The Past and Future of Competitive Advantage. Retrieved from <https://www.hbs.edu/faculty/Pages/item.aspx?num=6215>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Routledge.
- Colosi, R. (2004). *Negatively worded questions cause respondent confusion*.
- Credé, M., Harms, P. D., Blacksmith, N., & Wood, D. (2016). Assessing the Utility of Compound Trait Estimates of Narrow Personality Traits. *Journal of Personality Assessment*, 98(5), 503–513. <https://doi.org/10.1080/00223891.2016.1170023>
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93(5), 880–896.  
<https://doi.org/10.1037/0022-3514.93.5.880>
- Dickerson, A., & Taylor, M. A. (2000). Self-limiting behavior in women self-esteem and self-efficacy as predictors. *Group & Organization Management*, 25(2), 191–210.
- DiStefano, C., & Motl, R. W. (2009). Self-Esteem and Method Effects Associated With Negatively Worded Items: Investigating Factorial Invariance by Sex. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(1), 134–146.  
<https://doi.org/10.1080/10705510802565403>
- Dreyfuss, E. (2018, August 17). A Bot Panic Hits Amazon’s Mechanical Turk. *Wired*. Retrieved from <https://www.wired.com/story/amazon-mechanical-turk-bot-panic/>

- 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. <https://doi.org/10.1037/0021-9010.91.1.40>
- Dunlop, W. L. (2015). Contextualized Personality, Beyond Traits: Contextualized personality. *European Journal of Personality, 29*(3), 310–325. <https://doi.org/10.1002/per.1995>
- Fleeson, W. (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science, 13*(2), 83–87.
- Fleeson, W., & Leicht, C. (2006). On delineating and integrating the study of variability and stability in personality psychology: Interpersonal trust as illustration. *Journal of Research in Personality, 40*(1), 5–20. <https://doi.org/10.1016/j.jrp.2005.08.004>
- Fleeson, W., & Nofle, E. E. (2009). In favor of the synthetic resolution to the person–situation debate. *Journal of Research in Personality, 43*(2), 150–154.  
<https://doi.org/10.1016/j.jrp.2009.02.008>
- Funder, D. C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality, 40*(1), 21–34.  
<https://doi.org/10.1016/j.jrp.2005.08.003>
- Furr, R. M., & Funder, D. C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centered, and person-centered approaches. *Journal of Research in Personality, 38*(5), 421–447. <https://doi.org/10.1016/j.jrp.2003.10.001>

- Ganster, D. C., Hennessey, H. W., & Luthans, F. (1982). The Role of Social Desirability in Organizational Behavior Research. In *Academy of Management Proceedings (00650668)* (pp. 141–145). Academy of Management. <https://doi.org/10.5465/AMBPP.1982.4976494>
- Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social Desirability Response Effects: Three Alternative Models. *Academy of Management Journal*, *26*(2), 321–331. <https://doi.org/10.5465/255979>
- Geiser, C., Litson, K., Bishop, J., Keller, B. T., Burns, G. L., Servera, M., & Shiffman, S. (2015). Analyzing person, situation and person  $\times$  situation interaction effects: Latent state-trait models for the combination of random and fixed situations. *Psychological Methods*, *20*(2), 165–192. <https://doi.org/10.1037/met0000026>
- Geukes, K., Mesagno, C., Hanrahan, S. J., & Kellmann, M. (2012). Testing an interactionist perspective on the relationship between personality traits and performance under public pressure. *Psychology of Sport and Exercise*, *13*(3), 243–250. <https://doi.org/10.1016/j.psychsport.2011.12.004>
- Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: do they matter? *Personality and Individual Differences*, *35*(6), 1241–1254. [https://doi.org/10.1016/S0191-8869\(02\)00331-8](https://doi.org/10.1016/S0191-8869(02)00331-8)
- Griffo, R., & Randall Colvin, C. (2009). A brief look at interactionism: Past and present. *Journal of Research in Personality*, *43*(2), 243–244. <https://doi.org/10.1016/j.jrp.2008.12.038>

- Guillaume, E., Kumagai, S., Kawamoto, S., Sato, T., & others. (2012). The person-situation debate and the assessment of situations. *The Japanese Journal of Personality*, 21(1), 1–11.
- Haney, C., Banks, C., & Zimbardo, P. (1973). Interpersonal dynamics in a simulated prison. *International Journal of Criminology & Penology*, 1(1), 69–97.
- Harris, D. (2018, August 30). People fear Mechanical Turk has a bot problem, which would be bad for AI. Retrieved October 1, 2018, from <https://architect.io/people-fear-mechanical-turk-has-a-bot-problem-which-would-be-bad-for-ai-1e2027209864>
- Heggestad, E. D., & Gordon, H. L. (2008). An Argument for Context-Specific Personality Assessments. *Industrial and Organizational Psychology*, 1(03), 320–322.  
<https://doi.org/10.1111/j.1754-9434.2008.00056.x>
- Heine, S. J., & Buchtel, E. E. (2009). Personality: The Universal and the Culturally Specific. *Annual Review of Psychology*, 60(1), 369–394.  
<https://doi.org/10.1146/annurev.psych.60.110707.163655>
- Heller, D., Watson, D., Komar, J., Min, J.-A., & Perunovic, W. Q. E. (2007). Contextualized Personality: Traditional and New Assessment Procedures. *Journal of Personality*, 75(6), 1229–1254. <https://doi.org/10.1111/j.1467-6494.2007.00474.x>
- Herndon, J. S. (2011). The Effects Of Delay Of Gratification On The Academic Achievement, Substance Abuse, And Violent Behavior Of Middle-school Students In Alternative Learning Settings, 127.
- Hill, P. L., & Lapsley, D. K. (2009). Persons and situations in the moral domain. *Journal of Research in Personality*, 43(2), 245–246. <https://doi.org/10.1016/j.jrp.2008.12.034>

- Hogan, R. (2009). Much ado about nothing: The person–situation debate. *Journal of Research in Personality*, 43(2), 249.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Huang, C., & Dong, N. (2012). Factor Structures of the Rosenberg Self-Esteem Scale: A Meta-Analysis of Pattern Matrices. *European Journal of Psychological Assessment*, 28(2), 132–138. <https://doi.org/10.1027/1015-5759/a000101>
- Johnson, J. A. (2009). Wrong and right questions about persons and situations. *Journal of Research in Personality*, 43(2), 251–252. <https://doi.org/10.1016/j.jrp.2008.12.022>
- Johnson, J. A. (2014). Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality*, 51, 78–89. <https://doi.org/10.1016/j.jrp.2014.05.003>
- Judge, T. A., & Zapata, C. P. (2015). The person–situation debate revisited: Effect of situation strength and trait activation on the validity of the Big Five personality traits in predicting job performance. *Academy of Management Journal*, 58(4), 1149–1179.
- Krishna, A., & Peter, S. M. (2018). Questionable research practices in student final theses – Prevalence, attitudes, and the role of the supervisor’s perceived attitudes. *PLOS ONE*, 24.
- Kuppens, P. (2009). The legacy of the person–situation debate for understanding variability in emotional experience. *Journal of Research in Personality*, 43(2), 255–256. <https://doi.org/10.1016/j.jrp.2008.12.027>



- LeBreton, J. M., Barksdale, C. D., Robin, J., & James, L. R. (2007). Measurement issues associated with conditional reasoning tests: indirect measurement and test faking. *The Journal of Applied Psychology, 92*(1), 1–16. <https://doi.org/10.1037/0021-9010.92.1.1>
- Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology, 93*(2), 268–279. <https://doi.org/10.1037/0021-9010.93.2.268>
- Lievens, F., & Sackett, P. R. (2017). The effects of predictor method factors on selection outcomes: A modular approach to personnel selection procedures. *The Journal of Applied Psychology, 102*(1), 43–66. <https://doi.org/10.1037/apl0000160>
- Lucas, R. E., & Donnellan, M. B. (2009). If the person–situation debate is really over, why does it still generate so much negative affect? *Journal of Research in Personality, 43*(2), 146–149. <https://doi.org/10.1016/j.jrp.2009.02.009>
- Mastroianni, G. R. (2011). THE PERSON–SITUATION DEBATE: IMPLICATIONS FOR MILITARY LEADERSHIP AND CIVILIAN–MILITARY RELATIONS. *Journal of Military Ethics, 10*(1), 2–16. <https://doi.org/10.1080/15027570.2011.561636>
- Maul, A. (2013). Method Effects and the Meaning of Measurement. *Frontiers in Psychology, 4*. <https://doi.org/10.3389/fpsyg.2013.00169>
- McCabe, K. O., Van Yperen, N. W., Elliot, A. J., & Verbraak, M. (2013). Big Five personality profiles of context-specific achievement goals. *Journal of Research in Personality, 47*(6), 698–707. <https://doi.org/10.1016/j.jrp.2013.06.003>
- McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. *American Psychologist, 52*(5), 509.

- Michael Furr, R. (2009). Profile analysis in person–situation integration. *Journal of Research in Personality*, 43(2), 196–207. <https://doi.org/10.1016/j.jrp.2008.08.002>
- Milgram, S. (1963). Behavioral Study of obedience. *The Journal of Abnormal and Social Psychology*, 67(4), 371–378. <https://doi.org/10.1037/h0040525>
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY, US: Routledge/Taylor & Francis Group.
- Mischel. (2004). Toward an Integrative Science of the Person. *Annual Review of Psychology*, 55(1), 1–22. <https://doi.org/10.1146/annurev.psych.55.042902.130709>
- Mischel, & Shoda, Y. (1995). A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics. *Psychological Review (PsyncARTICLES)*, 102, 2.
- Morf, C. C. (2002). Personality at the hub: Extending the conception of personality psychology. *Journal of Research in Personality*, 36(6), 649–660.
- Murray, A. L., Johnson, W., Gow, A. J., & Deary, I. J. (2015). Disentangling wording and substantive factors in the Spiritual Well-Being Scale. *Psychology of Religion and Spirituality*, 7(2), 120–129. <https://doi.org/10.1037/a0038001>
- Nisbett, R. E., & Ross, L. (1985). *Human inference : strategies and shortcomings of social judgment*. Englewood Cliffs, N.J.: Prentice-Hall.
- Open Peer Commentary and Author’s Response: Discussion. (2015). *European Journal of Personality*, 29(3), 326–362. <https://doi.org/10.1002/per.2006>

- Paunonen, S. V., & Ashton, M. C. (2013). On the prediction of academic performance with personality traits: A replication study. *Journal of Research in Personality, 47*(6), 778–781. <https://doi.org/10.1016/j.jrp.2013.08.003>
- Paunonen, S. V., & Hong, R. Y. (2013). The many faces of assumed similarity in perceptions of personality. *Journal of Research in Personality, 47*(6), 800–815. <https://doi.org/10.1016/j.jrp.2013.08.007>
- Peters, G.-J. Y. (2014). The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality, *16*(2). <https://doi.org/10.31234/osf.io/h47fv>
- Petska, K. S. (2006). Using personality variables to predict academic success in personalized system of instruction. Retrieved from <http://digitalcommons.unl.edu/cehsdiss/3/>
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322–338. <https://doi.org/10.1037/a0014996>
- Revelle, W., & Condon, D. M. (2015). A model for personality at three levels. *Journal of Research in Personality, 56*, 70–81. <https://doi.org/10.1016/j.jrp.2014.12.006>
- Roberts, B. W. (2009). Back to the future: Personality and Assessment and personality development. *Journal of Research in Personality, 43*(2), 137–145. <https://doi.org/10.1016/j.jrp.2008.12.015>
- Roberts, B. W., & Caspi, A. (2001). Personality Development and the Person-Situation Debate: It's D7éjà Vu All Over Again. *Psychological Inquiry, 12*(2), 104–109. [https://doi.org/10.1207/S15327965PLI1202\\_04](https://doi.org/10.1207/S15327965PLI1202_04)

- Robie, C., & Risavy, S. D. (2016). A comparison of frame-of-reference and frequency-based personality measurement. *Personality and Individual Differences, 92*, 16–21.  
<https://doi.org/10.1016/j.paid.2015.12.005>
- Rosenthal, R. (1994). Science and Ethics in Conducting, Analyzing, and Reporting Psychological Research. *Psychological Science, 5*(3), 127–134.  
<https://doi.org/10.1111/j.1467-9280.1994.tb00646.x>
- Rosnow, R. L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44*(10), 1276–1284.  
<https://doi.org/10.1037/0003-066X.44.10.1276>
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY, England: McGraw-Hill Book Company. Retrieved from <http://libproxy.clemson.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1991-97382-000>
- Schilling, M. A. (2000). Toward a General Modular Systems Theory and Its Application to Interfirm Product Modularity. *The Academy of Management Review, 25*(2), 312.  
<https://doi.org/10.2307/259016>
- Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology, 80*(5), 607–620. <https://doi.org/10.1037/0021-9010.80.5.607>
- Schmitt, N., & Stuits, D. M. (1985). Factors Defined by Negatively Keyed Items: The Result of Careless Respondents? *Applied Psychological Measurement, 9*(4), 367–373.  
<https://doi.org/10.1177/014662168500900405>

- Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology, 65*(3), 445–494.
- Shiner, R. L. (2009). Persons and situations act together to shape the course of human lives. *Journal of Research in Personality, 43*(2), 270–271.  
<https://doi.org/10.1016/j.jrp.2008.12.023>
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A Proposed Addition to All Empirical Papers. *Perspectives on Psychological Science, 12*(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95*(2), 334–344. <https://doi.org/10.1037/0033-2909.95.2.334>
- Stodel, M. (2015). But What Will People Think?: Getting beyond Social Desirability Bias by Increasing Cognitive Load. *International Journal of Market Research, 57*(2), 313–322.  
<https://doi.org/10.2501/IJMR-2015-024>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Pearson Education.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tett, R. P., & Guterman, H. A. (2000). Situation Trait Relevance, Trait Expression, and Cross-Situational Consistency: Testing a Principle of Trait Activation. *Journal of Research in Personality, 34*(4), 397–423. <https://doi.org/10.1006/jrpe.2000.2292>

- Timothy Church, A., Katigbak, M. S., Reyes, J. A. S., Salanga, M. G. C., Miramontes, L. A., & Adams, N. B. (2008). Prediction and cross-situational consistency of daily behavior across cultures: Testing trait and cultural psychology perspectives. *Journal of Research in Personality*, 42(5), 1199–1215. <https://doi.org/10.1016/j.jrp.2008.03.007>
- Tomás, J. M., & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling*, 6(1), 84–98. <https://doi.org/10.1080/10705519909540120>
- Tracy, J. L., Robins, R. W., & Sherman, J. W. (2009). Will the real personality researcher and the real social researcher please stand up? *Journal of Research in Personality*, 43(2), 272–273. <https://doi.org/10.1016/j.jrp.2008.12.020>
- TurkPrime. (2018, September). After the Bot Scare: Understanding What's Been Happening with Data Collection on MTurk and How to Stop it. Retrieved October 1, 2018, from <https://blog.turkprime.com/after-the-bot-scare-understanding-whats-been-happening-with-data-collection-on-mturk-and-how-to-stop-it>
- van Oers, K. (2005). Context dependence of personalities: risk-taking behavior in a social and a nonsocial situation. *Behavioral Ecology*, 16(4), 716–723. <https://doi.org/10.1093/beheco/ari045>
- van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of Reverse Wording of Questionnaire Items: Let's Learn from Cows in the Rain. *PLoS ONE*, 8(7). <https://doi.org/10.1371/journal.pone.0068967>
- Vannette, D. L. (2016). Testing the Effects of Different Types of Attention Interventions on Data Quality in Web Surveys. Experimental Evidence From a 14 Country Study.

Presented at the 71st Annual Conference of the American Association for Public Opinion Research, Austin, TX.

Wang, Q., Bowling, N. A., & Eschleman, K. J. (n.d.). *A Meta-Analytic Examination of Work and General Locus of Control*.

Webster, G. D. (2009). The person–situation interaction is increasingly outpacing the person–situation debate in the scientific literature: A 30-year analysis of publication trends, 1978–2007. *Journal of Research in Personality*, *43*(2), 278–279.  
<https://doi.org/10.1016/j.jrp.2008.12.030>

Yanyun Yang, & Green, S. B. (2011). Coefficient Alpha: A Reliability Coefficient for the 21st Century? *Journal of Psychoeducational Assessment*, *29*(4), 377–392.  
<https://doi.org/10.1177/0734282911406668>

Zhang, X., Noor, R., & Savalei, V. (2016). Examining the Effect of Reverse Worded Items on the Factor Structure of the Need for Cognition Scale. *PLOS ONE*, *11*(6), e0157795.  
<https://doi.org/10.1371/journal.pone.0157795>